In Silico Representation of the Liver - Connecting Function to Anatomy, Physiology & Heterogeneous Microenvironments

by

Li Yan

B.S. (Jilin University) 1995
M.S. (Chinese Academy of Sciences) 1998
M.A. (University of Colorado, Boulder) 2000

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Joint Doctor of Philosophy
with University of California, San Francisco

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor C. Anthony Hunt, Chair
Professor Adam P. Arkin
Professor Lee W. Schruben

Fall 2007

The dissertation of Li Yan is approved:

Chair	Date
	Date
	Date

University of California, Berkeley

Fall 2007

In Silico Representation of the Liver - Connecting Function to Anatomy, Physiology & Heterogeneous Microenvironments

Abstract

In Silico Representation of the Liver - Connecting Function to Anatomy, Physiology & Heterogeneous Microenvironments

by

Li Yan

Joint Doctor of Philosophy with University of California, San Francisco

in Bioengineering

University of California, Berkeley

Professor C. Anthony Hunt, Chair

I report development, validation and application of an in silico liver (ISL). The ISL is a synthetic, discrete event, discrete space, agent-oriented model. ISL design is consistent with the vision of physiologically based modeling. Autonomous software objects representing hepatic components such as molecules, cells, and hepatic microarchitectural details were plugged together to form a functioning analogue of an isolated, perfused rat liver. The mappings between model components and their in situ counterparts are realistic and intuitive. Aspects of lobule microarchitecture and sinusoid structure are represented separately from drug metabolic and excretion functions. The hepatic heterogeneity that is characteristic of both normal and diseased livers can be easily represented. In silico experiments begin with co-administration of cationic drug and extracellular space marker sucrose. Measures of ISL outflow profiles are taken in a manner analogous to that in the referent PK experiments. ISL hepatic disposition of simulated sucrose, antipyrine, atenolol, labetalol, and diltiazem were validated separately against in situ data. By

following an iterative parameter refinement procedure, a single ISL structure was selected and parameterizations were held constant for all compounds. Parameters that respond to drug-specific physicochemical properties were tuned so that ISL outflow profiles and clearances properties matched the four pairs of in situ outflow profiles. The consequences of parameter changes on outflow profiles were explored. Systematic parameter changes altered outflow profiles in ways that were consistent with expectations based on knowledge of hepatic anatomy and physiology and drug physicochemical properties. The data have provided further verification and validation evidences. A Fuzzy c-Means (FCM) classification algorithm was used to determine the degree of similarity between the four matched drugs and the two new drugs, and to predict ISL parameter values for the latter. Predicted parameter values were then used in ISL to predict hepatic disposition of those two drugs. FCM predicted parameter values gave acceptable outflow profiles for prazosin and propranolol. The ISL is designed to adapt easily to new situations. It can become a component of larger, whole-organism models. This new class of models is expected to increase the pharmaceutical research productivity at all levels that make use of modeling and simulation.

_____

Professor C. Anthony Hunt

Dissertation Committee Chair

Dedicated to my parents, Qihou Yan and Shuqin Shi,

my husband, Guoqiang Chen, and my son, Patrick Chen.

Acknowledgement

It is finally time to write my acknowledgments. The people I want to acknowledge first, and to whom I owe my eternal gratitude, are my parents. Because of their endless love, I have been able to come a long way to where I am today. My father missed me immensely in his final days and kept hoping that I would return home soon. But he did not allow my family members to call me back home, because his dearest daughter was taking her entrance examinations for graduate school far from her hometown. His daughter's bright future was more important to him than seeing her one last time. He has been with me in my heart every step of the way. He was so proud of me when I received the highest score on the National College Entrance Examination in our hometown. I hope this work makes him proud, as well. My mother has contributed all of her time and love to me and to my brothers. She pays attention to all the news happening in the United States just because her beloved daughter is here. My life journey has been accompanied by her deep blessing. Without her selfless love and understanding, I would not have finished my seven years of study at Berkeley. Thank you, Mom and Dad. I love you both so very much.

I also wish to express my extreme gratitude to my advisor, Dr. C. Anthony Hunt. He is always ready to help and has been a great source of knowledge and advice during every stage of my graduate study. He equipped me with essential knowledge and skills on computational biology during the early years of my graduate study. He trained me to develop independent thinking and research skills throughout my doctoral work. His

tremendous help on proposal review and oral presentation rehearsal allowed me to successfully pass my qualifying exam. In the midst of the fog of my PhD research, he continually provided invaluable information and advice to guide my research towards its completion. During the final few months when I was wrapping up my research papers, he reviewed my manuscripts on weekends and even when he was away on vacation. Without his support, the work presented here would not have been possible. I also feel fortunate to have received from him many helpful suggestions and advice on my career and professional development.

I owe a special debt of gratitude to Dr. Adam P. Arkin for being my advisor since I came to Berkeley and for guiding my first-ever biological model using partial differential equations. I also thank him for being on my qualifying and dissertation committee and for his keen advice and constructive criticism on my research project. I am inspired and my horizons are broadened by every conversation I have with him. Thank you also to Dr. Lee W. Schruben, whose modeling and simulation course motivated me greatly. I am very grateful for his helpful advice on my project, and for being so considerate and making a special trip to sign my dissertation while he was on sabbatical. Thank you to my qualifying committee member and dissertation reader, Dr. Donna Hudson, for always replying to my e-mails promptly and for making a great effort to read my long manuscripts. Appreciation also goes out to my other qualifying committee members, Dr. Robert Spear, Dr. Betty-Ann Hoener, and Dr. Song Li, for taking the time to read my qualifying proposal and for providing many valuable comments. I also thank everyone in the BioSystems Group at UCSF for their support, cooperation, and discussions. In particular, I thank Glen E. P. Ropella for answering my numerous questions about

agent-based modeling and the swarm framework, Shahab Sheikh-Bahaei for help with the FCM classification method and manuscript review, Tai Ning Lam for hepatic clearance and PK discussion, G. Cosmo Haun for developing the ISL visualizations, Dr. Suman Ganguli for manuscript review, Dr. Sunwoo Park and Jesse Engelberg for FURM system maintenance, Mark Grant for helpful discussions, and the late Pearl Johnson for office support while she was at the Hunt Lab. I also thank the members of the Arkin group for their help and discussions during my seven years of graduate study. I especially thank Gwyneth Terry for her help and for the joyful chats, Janet Jacobsen for helping me with my English writing, and Patrick Flaherty for answering my statistics questions. .

I am deeply indebted to my son, Patrick, who brings me so much happiness, makes our house a cozy home, and turns our love into a family. I am so sorry that I sometimes had to refuse to play with him during the busy days. Finally, I am immensely grateful to my husband, Guoqiang. He shared my happiness when I made a little bit of progress; he eased my stress and encouraged me when I faced difficulties; he tolerated my bad temper when I was moody; and he took care of our son so that I could concentrate on writing my dissertation. His patience and understanding during the many long nights and weekends of work were selfless and are appreciated more than he knows. His unconditional love and support are unfailing and empower me to face life's ups and downs. I feel so lucky to have met him and am so grateful to be able to share this life-long journey with him.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANN: artificial neural network

CC: correlation coefficient

CL_{int} : intrinsic elimination clearance

CV: central vein

FST: Fuzzy Set Theory

FCMA: Fuzzy C-Means algorithm

f_{uB}: unbound drug fraction

HBA: hydrogen bound acceptor

HBD: hydrogen bound donor

ISL(s): in silico liver(s)

MW: molecular weight

P_{app}: octanol/water partition coefficient for pH 7.4

PB: physiologically based

PCs: properties and characteristics

PCP: physicochemical property

PK: pharmacokinetic

PS: permeability-surface area product

PV: portal vein

RB: number of rotatable bonds

$\boldsymbol{S}_1$ and $\boldsymbol{S}_2$: two classes of SS

SD: standard deviation

SM(s): similarity measure(s)

SS(s): sinusoidal segment(s)

TPSA: topological polar surface area

1 Introduction

1.1 The PBPK Vision

A key goal in developing pharmacokinetic (PK) models is to use simulation to facilitate drug development and clinical pharmacology. PK modeling has undergone considerable evolution in its efforts to achieve this goal. Nonparametric analysis makes no model assumptions about the body but only provides moment data, which may be used to generate PK parameter values. More commonly, data is analyzed assuming the body may be described as an abstract series of compartments or, in physiologically based pharmacokinetics (PBPK), by representing organs as compartments connected to represent a vascular system (1, 2). These and related model types are by nature *inductive*: they are a cognitive reduction of the biology to describe PK data, using representations of the hypothesized essential determinants governing the measured phenomena. Almost all of these models rely on systems of equations[1] (typically differential equations) and/or probabilistic networks to represent or describe essential features of PK data.

[1] Inductive models can take any form. Equations fit to data are just one example.

PBPK models recognize anatomical and physiological realities, and attempt to account for the role of differential distribution within and between organs as well as their varying blood flows (3). The expected heuristic value of PBPK modeling in research, drug development and regulatory science, and toxicological risk assessment is evident in the variety of envisioned model uses discussed in several recent reviews (3-5). Several desired and anticipated advantages of PBPK models have been cited (4, 5), but not yet realized, of which the following are just a few. It should be straightforward to reuse a parameterized PBPK model to explore the expected behaviors of different compounds. Models should be capable of reflecting whatever physiological detail is relevant to the problem. Updating needs to be facile. One should be able to replace low-resolution components with more detailed ones that will enable the simulation of mechanistically based pharmacodynamics (3, 6) or to account for changes in rate-limiting steps or relevant heterogeneity within tissues and cells (7, 8). Realization of these uses is expected to allow researchers to answer important questions such as: is it possible to provide increasingly confident explanations of events within target sites, and is it possible to adjust those events to take into account patient specific knowledge?

1.2 Inductive and Synthetic Methods

The inductive method for creating models dominates the PK literature. Inductive models, by definition, abstract away the very detail required by heuristic PBPK modeling. This has leaded to top-down attempts to graft details onto the highly abstract inductive models. Augmenting PBPK modeling with the more flexible synthetic method provides a middle-out strategy with two primary benefits: 1) a deep physiological

model–to–referent mapping and 2) access to unpredictable systemic phenomena in the model (e.g. those that may result from nonlinearities). Inductive methods, especially those dependent on continuum mathematics like ordinary differential equations, often rely on assumptions like linearity for their mathematical basis. Hence, any model developed with induction exhibits only the behaviors present, a priori, in the model *family* used for induction. Synthetic PBPK modeling involves using independent component to build a functioning analogue of the *mechanisms* (form and function; component interactions) of which an inductive PBPK model is an abstraction. Being composed of separate sub-models (9-13), synthetic models provide more flexibility in representation and underlying assumptions[1]. I posit that this flexibility is critical to the satisfaction of the core requirements of PBPK modeling and thereby achieving the vision.

1.3 Models of Hepatic Elimination

How will changes in hepatic details alter the hepatic disposition of two compounds administered simultaneously? Traditionally, answers have been obtained using appropriately designed *in vitro* and *in vivo* experiments. Sometimes they are not enough: the appropriate experiment can be impossible, too challenging, too costly, take too long to set up and complete, or the researcher may be unable to uncover the precise intrahepatic events. It would be desirable to be able to conduct experiments in silico that would provide useful answers to questions, coupled with a useful measure of uncertainty. The following three questions are examples of what can be addressed using in silico experimentation on synthetic models:

[1] Ideas of *inductive* and *synthetic* modeling are expanded in Chapter 2.

1. Can a synthetic model parameterized for one compound be used to represent the outflow profile of new compounds without compromising the model's ability to interact as it did before with the previous compound?
2. How might the change in space partition capability, solute intrahepatic binding probability and metabolization probability influence the hepatic elimination and metabolic profiles of compounds?
3. Is it possible to predict the hepatic disposition of new, unstudied compounds by studying synthetic models validated against other compounds with different PCPs and different PK behavior?

In this report, an ISL (in silico liver) is developed that can answer such questions. It is envisioned that this ISL will have many properties and behaviors that partially overlap with those of the current inductive models of hepatic elimination. They will also be capable of exhibiting properties and behaviors that are not achievable by those inductive models. During early development of a new class of models, it is best to have trusted, successfully used models that, for overlapping behaviors, can serve as standards, against which to contrast members the new class of models. The primary reason for focusing on the liver and hepatic elimination in this report is that the already existing rich literature provides multiple options for crossmodel validity.

1.4 Model Usage

The hepatic outflow profile of a compound is a phenotypic attribute. The greater the similarity between the measured behaviors of an ISL and known attributes of a liver, the more useful that ISL will become as a research tool and as an expression of the coalesced,

relevant knowledge of the liver. In this report, I focus on hepatic outflow profiles of sucrose and a co-administered drug, antipyrine. To simulate the complete system through a fundamental study of its component parts (2), this research will produce increasing overlap between ISL behaviors, properties, and characteristics, and measures of hepatic phenotypic attributes.

The long-term goal of this research is to produce increasing overlap between ISL behaviors, properties, and characteristics, and measures of hepatic phenotypic attributes. Achieving that goal depends on demonstrating models that achieve at least ten capabilities.

1. *Intrahepatic events*: The models are capable of accurately representing intrahepatic events.
2. *Physiological mapping:*There is clear physiological mapping between referent and model components because ISL observables are designed to be consistent with those of the referent liver.
3. *Turing test:* When dosed with a simulated compound, the ISL generates outflow data that are, to a domain expert (in a type of *Turing test*), experimentally indistinguishable from the referent wet-lab data; this requires that the ISL and its framework must be suitable for experimentation.
4. *Discrete interactions:* To enable the above capabilities and support 6-8 below, the ISL and its framework must use discrete interactions.
5. *Transparency:* The ISL must be transparent: the details of the simulation, as it progresses, need to be visualizable and measurable.
6. *Articulation:* It must be easy to join, disconnect, and replace ISL components without having to reengineering it or its components.

7. *Flexibility:* It must be relatively simple to change usage and assumptions, or increase or decrease detail in order to meet the particular needs of an experiment, without requiring significant re-engineering of the model.
8. *Reusability:* The ISL must be reusable for simulating the disposition, clearance, and metabolic properties of multiple compounds in the same experiment, not just one each in separate experiments.
9. *Adaptability:* In addition to being flexible and reusable, the ISL must be constructed so that it can adapted to function as an organ component within a larger, synthetic, physiologically based, whole organism model.
10. *Prediction:* The ISL will be able to predict new compounds' hepatic disposition based on those compounds' physicochemical properties.

1.5 In Silico Framework: Technical Detail

The primary computational tool is an experimental, small-scale Beowulf cluster, referred to as the FURM cluster, configured as a test-bed for developing and testing the ISL. It consists of 8 nodes and a Gigabit switch. To control and maintain the whole cluster system, one node is assigned as a system management node (head node). The other seven are dedicated to only computational tasks (slave nodes). All eight processors are available for performing distributed/parallel computational tasks.

Code is managed using CVS with a single HEAD branch and ChangeLogs for each commit. Experiments are conducted using the last stable version of the HEAD branch. Validation data is loaded using HDF5 (a Hierarchical Data Format product). As an experiment runs, simulation data is kept in memory (indexed by Scheme representation

of the parameter vector) until the end of the experiment, at which point it is written to comma separated files indexed by filename. A "make" target is used to move all the relevant input, output, and parameter data into a dated directory, preserving the integrity of the relationships. A Python script processes the raw output and prepares the data for analysis and plotting by R scripts. For each experimental milestone, data is archived to CD-ROM.

Development follows a loose and iterative specification, implementation, test, and commit process. The desired features are listed and the developers propose a mechanism to implement the features. When accepted, development begins. After each code change, the specific changes are documented in the ChangeLog. The code is tested against a brief experiment where inputs, parameters, and output are compared with prior experiments and examined for anomalies. The changes are then committed to CVS. A new, clean sandbox of the HEAD branch is checked out, built from scratch, and a new test experiment is run and the data checked for anomalies. The developer who made the changes then posts a high-level description of the changes, and the reasons for making them.

Coding standards follow typical Swarm patterns, using Objective-C and C++ (with Object Oriented patterns). Development adheres to the architecture of the FURM method (*http://furm.org*). There is a single CVS module for the entire framework. There are seven sub-modules: one for the experimental data; one for the Reference model; one for the Research models (ISLs); one for documentation; one for common modeling utilities (programming generics, graphs, and the statistics and data management classes); one for the input data; and one for the set of data processing scripts and other tools.

1.6 In Silico Experimental Method

The in silico experimental method is as follows. An experiment is set up by building the software into an executable, creating/editing the parameter vector, and beginning the simulation. An experiment consists of several runs that comprise a Monte Carlo style simulation, the results of which are post-processed to arrive at a nominal data set that represents the result of the experiment. The parameter vector also contains the limited capability to specify multiple experiments by iterating over some of its values. The raw data is preserved, but this is typically used only for parameter sweeps and data matching. At the end of an experiment, the data is isolated with the "make" target and one of several data reduction scripts may be run to process and analyze the simulation result. Typically, this processing involves an external (to the simulation) calculation of the similarity measure (SM) and plots made for examination by the experimenter. A new parameter vector is then designed based on the quality of similarity and the profile features as indicated by the data processing scripts.

1.7 Thesis Outline

The goal of this project is to develop a physiologically and mechanistically realistic software device, an ISL, which can be used to refine and test hypotheses about interacting mechanisms that influence the hepatic disposition of compounds of interest, therefore gaining deeper insight into the complexity of the interacting features of hepatic clearance, and can be used as a tool to predict a new drug's hepatic disposition.

This dissertation is organized as follows. The ISL is based on the synthetic

approach, rather than the inductive approach used by current PBPK models. Chapter 2 focuses on the analysis of inductive and synthetic modeling methods. Those two methods are contrasted and compared, and the validation of synthetic models is discussed.

One specific aim of this project is to develop and validate an ISL that is suitable for studying the hepatic disposition and metabolism of cationic drugs. In Chapter 3, I plugged together autonomous software objects representing hepatic components such as metabolic enzymes, cells, and hepatic microarchitectural details and developed a functioning analogue of an isolated, perfused rat liver. ISL hepatic disposition of sucrose, antipyrine, atenolol, labetalol, and diltiazem are separately validated against in situ data. All five of the simulated compounds give acceptable outflow profiles compared to in situ data. I explore the influences of ISL parameter changes on the outflow profiles. Systematic parameter changes alter outflow profiles in ways that are consistent with expectations based on knowledge of hepatic anatomy and physiology and drug physicochemical properties.

Accurate estimation of a drug's hepatic disposition is a crucial step in the development of practical clinical drugs. The other specific aim of this project is to predict hepatic disposition of new drugs using the ISL. The traditional approach to predicting the *in vitro* and *in vivo* properties of new drugs is to search for patterns within large data sets of measured biological properties and seek patterns within the set of biological property data of those correlated drugs. Knowledge about the mechanisms that generated the biological data is used indirectly. In chapter 4 I predict hepatic disposition for two cationic drugs, prazosin and propranolol, based on both the

knowledge mechanisms and the patterns found in the space of biological and physicochemical properties. Fuzzy C-mean algorithm predicted ISL parameter values that give acceptable outflow profiles for both prazosin and propranolol.

Finally, in Chapter 5, an overview is given for the whole project. The Details, Accuracy and Realism of the ISL are discussed. Possible future research directions are also addressed.

2 Inductive and Synthetic Methods

2.1 Inductive and Synthetic Methods

Inductive models are usually built by analyzing data, creating a mapping between the envisioned system structure and components of the data, and then representing the generation of those data components with mathematical equations. The method relies heavily on the researcher's knowledge combined with induction. When successful, it creates models that extrapolate somewhat beyond the original data, making the method ideal for prediction. However, the models lack particulars. They are black boxes. The detail that is washed away when an equation is induced contains heuristic value that is needed to achieve the PBPK vision. That detail may describe, for example, the mechanism by which the data were generated, whereas the mathematics only describes the abstracted properties of the data.

Most PBPK modeling is equation based and, consequently, is limited by the complexity of the equations needed to describe intricate functional details. A challenge has been to add back detail and thereby create models that more reliably describe desired and relevant features, including aspects of the very functional and spatial details that were

abstracted away in equation-based models. The convection-dispersion model of the liver and its extension to two compartments in the extended convection-dispersion model (7) is an example. Logically, an even more detailed model would be based on an assembly of objects representing cells in a configuration that recognizes the heterogeneity in cell function. Such higher resolution phenomena cannot easily be simulated within a larger system model using the continuous state plus continuous time class of models to which most current PBPK models belong. Consequently, there is a need for additional modeling and simulation methods that are more responsive to the preceding challenges.

There are now a variety of systems and behaviors emphasizing discrete events that are being simulated using object-oriented (OO) programming, with the extent of the modeling being more than what can be represented using the more mature continuous state, continuous time models (13). Discrete event, discrete space, discrete time models offer theoretical and practical advantages because they allow, even facilitate, the composition (plugging together separate components) and decomposition of models (e.g., removing components without breaking the model). Moreover, in the case where a continuous model is desirable, a discrete model can simulate or approximate it to arbitrary precision.

The constructivist or synthetic method (9-12, 14) consists of proposing and constructing building blocks that can be assembled to create a transparent (easily understood or seen through), artificial system that functions in the real world. The experimental *in vitro* models that are the mainstay of biomedical research (cell cultures, isolated, perfused organs and tissues, etc.) are synthetic models where some of the building blocks are laboratory items and others are living parts. Using OO

programming it is now feasible to build synthetic models computationally. Synthetic PBPK modeling involves using component parts to build, in silico, a functioning analogue of the *mechanisms* (form and function) of which an inductive PBPK model is an abstraction. This approach is especially suitable for representing spatial and discrete event phenomena and has the potential to be developed and merged with the current PBPK modeling approach. The resolution of the model dictates whether a mechanism is best modeled as discrete or continuous and a general discrete modeling approach allows the flexibility to choose the appropriate resolution according to the particular mechanism.

2.2 Contrasting Inductive and Synthetic Models

The advancements in OO programming make it easier to construct large system models where the modular components are inductive models or synthetic models. What follows are brief descriptions of examples of five *different types* of constructed, composite system models. In the paragraphs that follow, I draw on the second and third types to briefly compare and contrast inductive and synthetic PKPB models.

1. Synthetic models of engineered systems (e.g., a FedEx distribution center, a Boeing 787) that are detailed and articulated: the underlying mechanisms are known. Each component is very well understood and can be accurately represented in a number of ways including validated abstract equations and more detailed, mechanistic, hierarchical representations. The components are integrated into a system model in order to *study the referent system as a whole and each component in its systemic context.* The system model is transparent because at any level of resolution there is enough knowledge about the system to

replace any component at any time with a more detailed, less abstract, representation, should that be needed.

2. Assembled, equational PBPK models (e.g., a physiologically based toxicokinetic model as in (14): there is incomplete knowledge about the underlying mechanisms. Consequently, aspects of system function are separately represented using equations that map to abstract representations of known or hypothesized mechanistic components. The biological components are not transparent because their underlying mechanisms (although basically understood) are not fully *known*, and/or because there is inadequate data on important deeper features. Each PBPK model component is an equation; that equation is a black box because it is an induced abstraction (one can not look into the component to see how it works). More refined wet-lab experiments can supply additional detailed data to make the functioning of a component less opaque. However, inductive components in a PBPK model can never become transparent, as in type 1, because making them so would require *complete* knowledge of biology. When each component of a system model is an inductive model of its biological counterpart, the composite model, at its core, remains an inductive model: it is a large, composite inductive model. Robert Rosen (15) eloquently details the dangers of extrapolations that rely too strongly on the parameter values of such models.
3. Synthetic physiologically based models of mammalian systems of the type presented herein: there is incomplete knowledge about the underlying mechanisms. Nevertheless, a model is built using OO analogues to recognizable components (as in the ISL). As in type 2, each *biological* component lacks transparency.

However, unlike type 2, each component *in the synthetic model* is transparent. The analogue components in the OO model, when they behave like those in the biological system, provide a specific instantiation for how the corresponding biological components *might actually work*.

4. Of course, depending on intended use, hybrids of type 2 and 3 are feasible. Neither of the methods in 2 and 3 is superior. Both should be used when the particular strengths of both are needed, which is the case for the PBPK vision.
5. To actually close the gap between both types of models and their biological referents, synthetic models will need the ability to fundamentally change during simulations. They must be able to evolve and undergo localized model morphogenesis in response to events within the simulation.

It is appreciated that all large, multi-component models have synthetic and inductive aspects, depending on the perspective from which they are viewed. At the bottom of every synthetic model lie inductively defined atomic components, and every inductively defined computer model is synthetic at its highest systemic level.

Inductive models are induced from data and synthetic models are *constructed* from pre-existing components. Any component (like a LOBULE) could be an ODE model. As an example, I could replace one of the articulated ISL LOBULES with the two-phase stochastic kinetic model (2PSK model) in (1). The resulting hybrid ISL (1 [2PSK model] + 49 LOBULES) would be a synthetic model. In fact, I could replace all 50 LOBULES with varying instantiations of the 2PSK model and the resulting ISL would *still* be a synthetic model because it is a composite of 2PSK models and the result of executing all 50 together is not induced from any data set (even though the 2PSK model is).

2.3 Comparing Inductive and Synthetic Models

Most current PBPK modeling methods have both inductive components and synthetic features, as described above for type 2. Whereas the components themselves are inductive models, the combination of the components is synthetic. However, the combinations remain mathematical summaries of the component phenomena being modeled and, consequently, their predictive capabilities are limited to variants that preserve the induced parameterizations and I/O characteristics of the components. The conditions under which such models can be used reliably also depend on the characteristics of the data from which the model components were induced. Combining the model components together raises the potential for component discordance, especially with an increase in the variety of components. Component discordance can arise from inconsistencies or lack of conformity between parameterized components at the model or simulation level.

The components of the type 3 model described above are designed to be concordant (they interact smoothly over the full range of conditions). Such a model is intended to be exercised (run) and measured in the same way as its referent *in vitro* or *in vivo* experimental system and data taken from it is validated against known details of, and the data taken from the referent. An essential difference between the methods in type 2 and type 3 lies in the fact that the inductive method explicitly uses the observed phenomena (the data) as its input whereas the synthetic method starts with proposed building blocks and the relations between them (10, 11). The inductive method takes the ontology defined implicitly in the data as an inherent assumption whereas the synthetic method, employing cycles of creation, verification, and model change, attempts to construct an

ontology (and an analogue that "lives" in that ontology) that can realize data of interest.

Differences between the two modeling methods can be seen in the mappings from the space of generator mechanisms to the space of phenomena. The inductive method starts with the phenomena, PK data, and works backward to the generators in an attempt to discover an inverse mapping from range to domain. The criteria for success sit primarily in the range (the data). The synthetic method, in contrast, works forward from domain (components and generators) to range. Its constraints and criteria sit primarily in the domain. Induced models are ideally suited for exploiting discovered characteristics whereas synthetic models are ideally suited for exploring the consequences of assembled components. Which type is used depends on what one is trying to do, to what ends one intends to use the model, and which elements of trade-off one needs or can take. Equation-based models lack that detail, yet all the heuristic value lies in the details. The inductive model will likely provide a better fit to the data and a better extrapolation of that data if the same experimental conditions and protocol are retained. An inductive model allows one to make claims about the data (on that model). The synthetic model allows one to make claims about mechanisms. The synthetic model can provide a hypothesis for the "machine" that is believed to generate the data. Liu and Hunt have shown that such hypotheses can be tested in silico (16). Because of their greater detail, synthetic models can have a greater heuristic value and a larger space of possible experimental conditions and system behaviors.

A strength of inductive, equation-based models is that they are purposefully removed from the specific, mechanistic, and causal particulars of the system being modeled. Consequently, they are easily scaled. The only reason to use any other class of models

is when one's objective depends on knowledgeable use or understanding of those particulars. In that case, one must refuse to abstract away from the particulars. In the ISL, for example, I want to be able to distinguish ANTIPYRINE from SUCROSE within the same simulation, between sinusoidal segments, and between spaces such as the intra-endothelial-cell space, the Space of Dissé, and the intra-hepatocyte space. I want to be able to distinguish between normal and diseased sinusoidal segments, etc.

Prior to the advent of OO programming it was difficult to build serious biological analogues using software. Consequently, effort and energy logically fixated where meaningful progress could be made: inductive modeling and reliance on equations. The fundamental difficulty with the synthetic method is establishing requirements for building an analogue that functions somewhat like the referent. Synthetic modeling requires knowledge of the function and behavior of the referent, of plausible mechanisms for that function, and of relevant observables by which the analogue and the referent will be measured. For the ISL, considerable histological, physiological and pharmacokinetic (PK) knowledge is already available.

2.4 Validation of Synthetic Models

How do PK researchers validate their experimental *in vitro* model systems? The question is relevant because synthetic PBPK models are analogues of their wet-lab counterparts (for a discussion of analogues, see (15)). The validation of a model is defined in the (US DoD) Defense Modeling and Simulation Office's glossary as "the process of determining the degree to which a model or simulation is an accurate representation of the real world from the perspective of the intended uses of the model or

simulation." The concept of accuracy for a given purpose boils down to believability and usefulness on the part of domain experts. Experimental reproducibility within and between laboratories is a key aspect of believability for wet-lab models (17). Models are designed to meet specific requirements. If the requirements are taken from the referent biological system, then a set of referent observables will guide evaluation of the model. Those observables can be qualitative or quantitative, and in the latter case, it is feasible to have a quantitative measure of how well the observables are represented. In addition, equation based models can be induced from the simulated data to facilitate validation. Validation can be achieved, for a model, by a direct, quantitative, or semi-quantitative comparison between model outputs and the referent data. A synthetic model will not necessarily have a quantitatively precise mapping between its observables and those of the referent *if* model use requirements that were not based on quantitatively precise observables. When there is high uncertainty in a synthetic model, or when quantitatively precise referent data with which to compare is not available, other forms of validation are required.

As in the case of *in vitro* models, weaker forms of validation can be established through any process that increases the believability of the model, particularly to domain experts. These validation methods vary in their precision from precise but derived quantitative methods, to a kind of Turing test, where the model observables are shown to an expert in an attempt to convince the expert that s/he may be observing data from the referent system. Similar approaches can be used in validating synthetic in silico models. Ultimately, the techniques for validating synthetic models all descend from the basic process of measuring the model and the referent, and then comparing the behaviors and

outputs. The type of accuracy that can be obtained depends fundamentally on the type of requirements for which the model is being developed.

3 Modeling and Simulation of Hepatic Drug Disposition Using a Physiologically Based, Multi-Agent In Silico Liver[1]

3.1 Introduction

The vision of physiologically based (PB) pharmacokinetic (PK) modeling (3) is to provide a mechanistic and more realistic description of the behavior of substances in various tissues, with the intent of addressing such questions as: "Why do we see the observed behavior? Can we explain differences among compounds?" How do we better anticipate pharmacokinetics in patients, when a compound is administered alone or with other drugs "from in vitro and preclinical information?" How can we "provide increasingly confident predictions of events occurring with drugs at target and other sites … with age, in disease?" For that vision to be achieved, we need "realistic yet more complex models that take into account such factors as the various physical spaces within tissues, the existence of permeability barriers, organ heterogeneity, and active

[1] This chapter is published in Pharmaceutical Research (DOI: 10.1007/s11095-007-9494-y)

transport or metabolic processes," (3) all within a reliable and easily reused simulation framework. Progress has not yet been sufficient to enable answering these questions today. Consequently, expediting achievement of that vision was a factor motivating development of the recently described In Silico Liver (ISL) (18). It is an example of a (new) class of synthetic, discrete, componentized, physiologically based, computational, analogue models that are intended for refining, exploring, and testing hypotheses about interacting mechanisms that influence the transport, metabolism, and hepatic disposition of compounds of interest. It is intended for unraveling the complexity of the interacting features of hepatic disposition defined by *simulation* of events occurring within the PBPK context, but using a precise representation of the hepatic architecture, biochemistry and physiology. In contrast, classical PBPK models, as exemplified in (1, 7, 19-22), make parsimony paramount. By emphasizing identifiability and certainty of parameters obtained by nonlinear regression, they emphasize describing PK data in terms of the simplest global representation of physiological events. Once that has been done satisfactorily, interest turns logically to the problem of gaining a better understanding of the detailed mechanisms underlying the data. Historically, the only practical approach was to adapt analytical modeling methods and use equations to represent posited mechanistic details. Traditional PBPK models are an example of that approach. The approach, however, encounters obstacles when there is incomplete knowledge about the mechanistic details. Coincident with advances in object-oriented programming, it became feasible to focus directly on mechanisms by asking, what mechanisms can I build from software components (objects) that when measured in some way will yield data comparable to my referent data? The model that results from such building and

validation is not inductive: it is synthetic.

Other key features of the similarities and differences between traditional, induced, PK models and synthetic models of the ISL type have been detailed under Supplementary Material in (18) and in (23) and (16). It becomes clear that the two model classes have different goals and uses. If future ISL models become more predictive, it will be because they are more realistic and thus more complicated. In contrast, PBPK models defined using nonlinear regression tend to be more robust because the fewest number of parameters possible are used, recognizing that for induced, mathematical models, each additional parameter adds to model uncertainty and may compromise predictive power. For models of the ISL class, the following has been a working hypothesis. Through the combined use of discrete and synthetic methods to more realistically represent underlying 3D morphology and microevents, computational scientists will be able to better predict the PK consequence of changes in drug structure or pathology, for example, by using validated simulations like the ISL, as distinct from the more conventional PBPK regression approaches. Although this report is an important step toward that end, further research will be needed to support that working hypothesis.

Within this first generation of ISL, hepatic form and function are disjoined: different, autonomous components represent them separately at several levels. This disjunction does not occur in reality and that issue will be addressed in later ISL generations. Computational efficiency in this first generation is facilitated by having aspects of hepatic microarchitecture represented independently by components (the agents and functional objects described under Methods), rather than equations, arranged to represent spatial organization of function. Further, each ISL component can interact

uniquely with mobile objects representing drugs and other compounds of interest. Indeed, a variety of different mobile objects can be studied, simultaneously or separately. The consequences of interactions between simulated drugs and spatially confined components, controlled by known or hypothesized component-specific principles, can be measured and studied analogous to wet-lab experiments. The resulting data can be compared with (as distinct from being fit to) in vitro and in situ data to refine the ISL and help clarify posited, causal linkages thought to underlie hepatic disposition phenomena.

Research on this class of models is motivated by achieving three broad goals. 1) Experimentation with such models will help unravel complex pharmacokinetic processes. 2) Future device components can be reused and "educated," for example, as in (3) and (24), to use the physicochemical properties (PCPs) of new compounds as input to generate expected hepatic disposition data that can be mapped to the species of interest. 3) With progress on 1) and 2), expected hepatic disposition and drug metabolism properties for sets of new compounds can be explored in advance of costly wet-lab experiments or clinical trials, potentially saving considerable time. Achieving these capabilities will require pushing modeling and simulation capabilities such that I can methodically capture more multi-scale knowledge within models.

To what extent can different ISLs overcome limitations encountered by traditional PBPK models? Exactly what is needed to move toward the above expectations? Hunt et al. (18) suggested that the following may be possible. 1) Having specified an ISL, separately parameterize and hold constant its key microarchitectural features. 2) Dose with simulated counterparts of each of several previously studied compounds. 3) Enable the ISL to distinguish among the different simulated compounds. 4) Adjust only

PCP-sensitive, ISL parameters so that the in silico disposition properties of the simulated compounds—dosed alone or in combination—match the known properties of those compounds reasonably well so that a similarity criterion is met.

In this work, I apply the ISL to describe the isolated perfused rat liver systems, in an attempt to advance on traditional models that have been used to evaluate hepatic drug disposition and clearance (1, 7, 19-22). Recognizing that cationic drugs account for 70 to 80% of all drugs (25), I simulated the hepatic disposition and metabolism of four cationic compounds, atenolol, antipyrine, labetalol, and diltiazem, along with the co-administered, neutral extravascular marker sucrose. I verified that the ISL could generate in silico experimental results that were indistinguishable from those measured during in situ experiments. I followed an iterative ISL refinement procedure that led to a single core ISL structure parameterization that could be used for all five compounds. Only the parameter values of PCP-sensitive, ISL components that interact differently with different compounds were tuned to give good compound-specific results. Monte Carlo variants of the resulting ISLs were used to simulate outflow profiles for all five compounds, alone or in combination. The results provide additional ISL validation evidence.

Because of the non-deterministic nature of the ISL, the consequences of PCP-sensitive, parameter changes are neither unique nor simple. The consequences of changing four, PCP-sensitive, probabilistic parameter values are presented and discussed to illustrate specifically how a change in PCPs can be expected to alter ISL behavior. Two consider changes in the probability of simulated drug moving between sinusoidal spaces, the third reflects on changing the probability of simulated intracellular binding,

and the forth deals with changing the probability of a metabolic event. Based on that validation, I hypothesize that static and dynamic ISL mechanistic details, although abstract, map realistically to hepatic mechanistic details. The results represent an important advance in the science and methods of PBPK modeling and simulation.

3.2 Methods

To clearly distinguish in silico components and processes from corresponding hepatic structures and processes, I use SMALL CAPS when referring to the in silico counterparts.

3.2.1 ISL Structure and Design

ISL structure is illustrated in Figures 3.1 and 3.2 and detailed in Hunt et al. (18). For convenience, an abridged description is provided along with a brief rationale for ISL component design. Where most useful, I contrast the ISL with conventional PK models. See (18) and its Supplemental Material, along with (23) and (16) for additional discussion of design considerations and contrasts. The ISL is an abstract, physiologically based representation of a hepatic lobule. It is not intended to duplicate a liver, detail-for-detail. An operating guideline has been that if results of in silico experiments reasonably match wet-lab counterparts over a variety of experiments, then the in silico mechanistic details may map to corresponding in vivo or in situ details in informative ways. ISL components mimic essential form and function features of the in situ perfused rat liver system (26) used to study the hepatic disposition of sucrose, antipyrine, atenolol, labetalol, and diltiazem, the compounds on which this report focuses. In (1), the portal

vein was cannulated to enable single pass perfusion; perfusate was collected using a fraction collector. The fraction of the administered dose contained per unit of collected perfusate was measured. ISL experiments follow the same protocol. (27)

The liver is composed of secondary units, which are composed of several lobules. I assume that lobules are similar throughout the liver and within secondary units. My task thus reduces to simulating a large collection of lobules. Blood flows through the lobules via sinusoids, which provide access to hepatocytes. Blood enters lobules through portal vein tracts (PV) and vascular septa, and drains into branches of a common central vein (CV) (27). The acinar flow patterns within a lobule are represented in the ISL by a directed graph (Figure 3.1), which abstractly mimics these flow paths. To illustrate, a flow path from PV to CV can be divided into three segments and represented by a directed graph: PV to node 1, node 1 to node 2, and node 2 to CV. Following that method, hepatic function is represented within the segments. ISL graph edges represent flow connections between the graph nodes, called Sinusoidal Segments (SSs). Because traditional PK models have focused primarily on the PK data and its structure, information about paths and their interconnections has been abstracted away. Such abstractions are adequate when the primary use is systemic prediction for which those details are not needed.

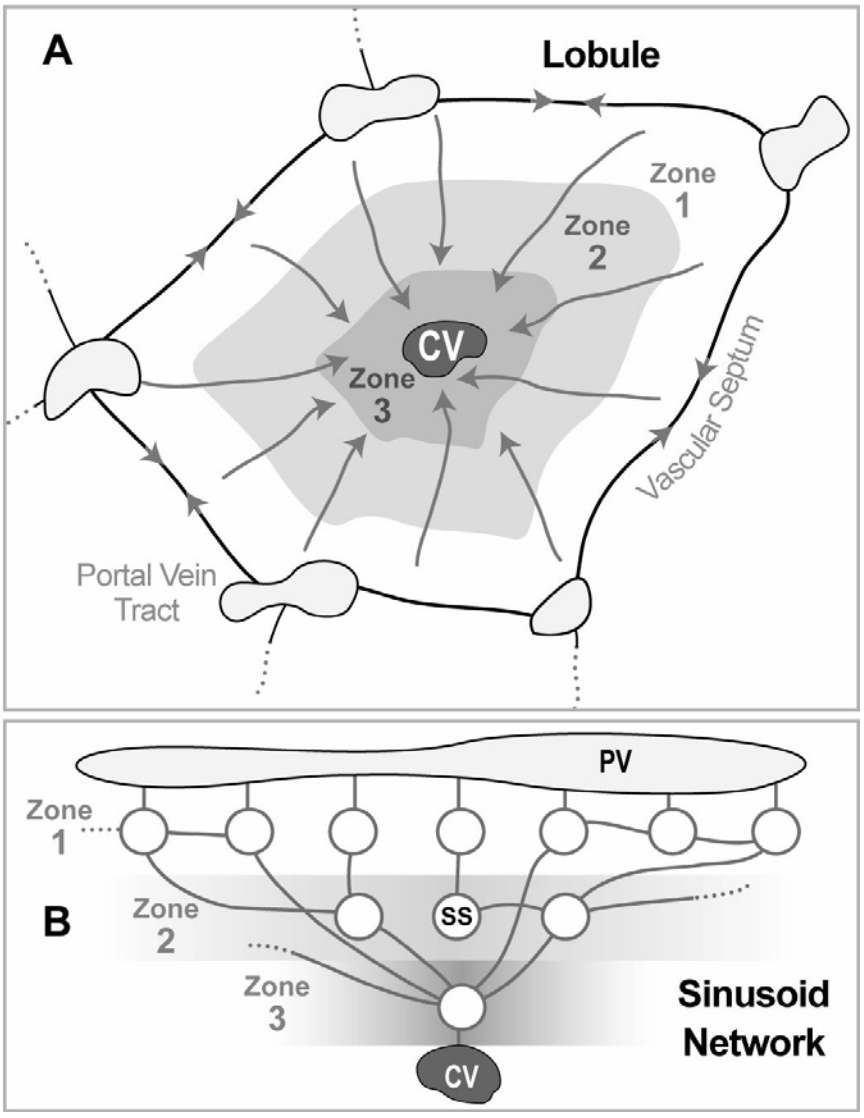

Figure 3.1 Illustrations of hepatic lobular structures and their representation within the ISL. (A) A schematic of a cross-section of a hepatic lobule showing the direction of flow from the terminal protal vein tracts (PV) through sinusoids in three concentric zones to the central hepatic vein (CV). Different zones have quantitative differences in structural characteristics and enzyme levels. (B) A portion of the sinusoid network is represented by an interconnected, three zone, directed graph. Data from the literature are used to constrain the graph size and structure. SS: sinusoidal segment.

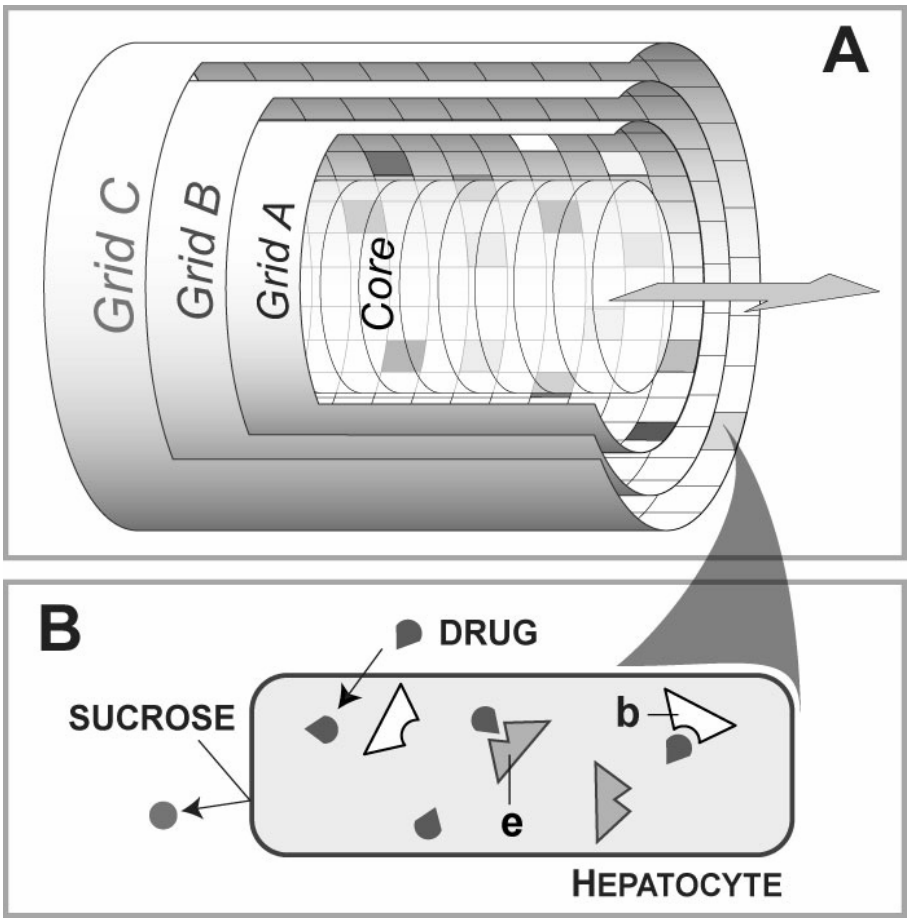

Figure 3.2 Illustrations of two key ISL components. (A) A schematic of a sinusoidal segment (SS): one SS occupies each node specified by the directed graph (Figure 3.3). Grids represent spaces and can contain objects representing the functions associated with a portion of sinusoids. Objects representing DRUG and SUCROSE enter and exit via the Core and Grid A. From Grid A, they can access the other spaces. Grid locations have properties and that govern their interaction with mobile COMPOUNDS. Different shadings of Grid A illustrate the potential for representing heterogeneous properties. Objects functioning as containers (for other objects) are used to represent cells, and can be assigned to any grid location. The Core represents blood flow; Grid A (referred to as Rim Space) represents the sinusoid rim; Grid B (referred to as the Endothelial Layer) represents endothelial cells and fenestra; Grid C (referred to as the Space of Disse and Hepatocyte Layer) represents all other spaces, including hepatocytes. (B) A HEPATOCYTE container: objects representing all needed intracellular components can be placed within. Only two types of intracellular BINDERS recognize drug: those that simply bind (b) and those that represent enzymes (e) and can metabolize. Cell containers will not allow SUCROSE to enter. Bile was not needed for these simulations but can easily be represented as a extension as described in (18).

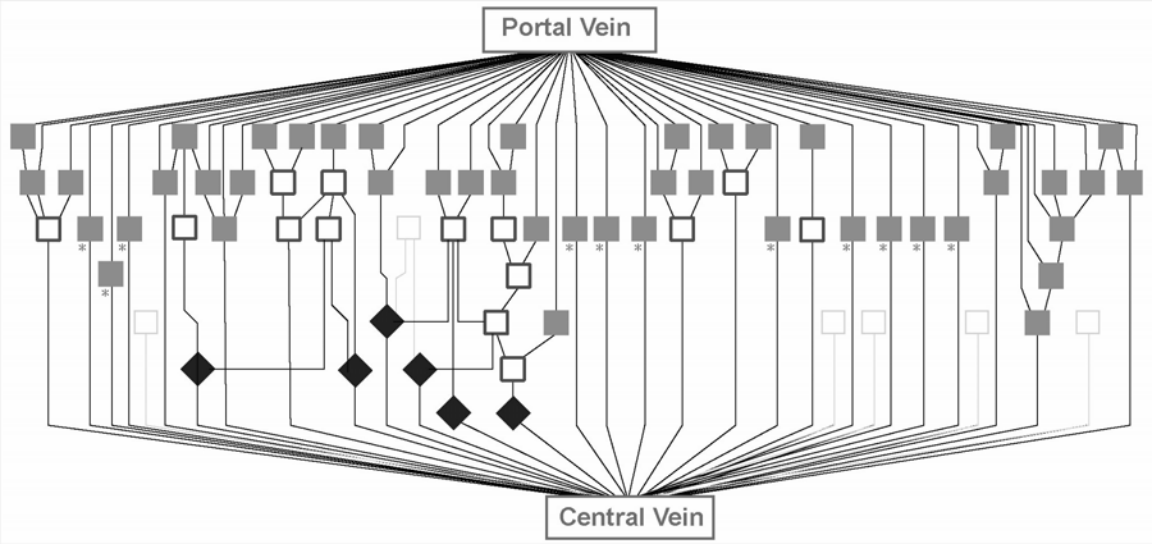

Figure 3.3 A sample ISL lobule graph structure generated for one simulation run. All Zone 1 nodes (gray squares) have one incoming edge from the PV. They have one outgoing edge; it can be connected to another SS or to the CV. Zone 2 nodes (unfilled squares) have an edge incoming from a Zone 1 SS; the outgoing edge can be connected to a SS in Zones 2 or 3. Zone 3 nodes (black diamonds) have one incoming edge from a Zone 2 SS; the outgoing edge is connected to the CV. The seven, light gray, unfilled squares represent Zone 2 nodes that, by chance, were not assigned an incoming edge. Asterisk (∗): as drawn, these Zone 1 nodes look identical. However, the SS placed at each such apparently identical node is different: they can be either S_1 or S_2; within both types, size and other details can be different. When needed, a parallel, mirror image graph structure can be added to represent biliary excretion.

Table 3.1 ISL parameters, descriptions, and values

Category	Name	Description	Value[a]
Lobule Graph	*GraphSpecFile*	Node in Zone I	45
		Node in Zone 2	21
		Node in Zone 3	6
		Intra-Zone I edges	18
		Intra-Zone II edges	10
		Intra-Zone III edges	0
		Zone I → Zone II edges	15
		Zone I → Zone III edges	0
		Zone II → Zone III edges	8
Sinusoid Structure Parameters	*DirSinRatio*	Percentage of SS that are type S_1 ("direct")	0.75
	TortSinRatio	Percentage of SS that are type S_2 ("tortuous")	0.25
	DirSinCircMin	Upper and lower bounds of the SS circumference,	50
	DirSinCircMax	generated by a pseudo-random number using the	50
	TortSinCircMin	uniform distribution	4
	TortSinCircMax		4
	DirSinLenAlpha	Length of SS, generated by a modified Gamma distribution;	3.0
	DirSinLenBeta	the modification consists of a left-right shift of the	0.215
	DirSinLenShift	distribution, allowing the user to clip off the front of the	0
	TortSinLenAlpha	distribution	10
	TortSinLenBeta		0.125
	TortSinLenShift		-40
	ECDensity	Relative ENDOTHELIAL CELL density within Grid B	0.8
	HepDensity	Relative HEPATOCYTES density within Grid C	0.6
Parameters influenced by MW	*SinusoidTurbo*	The complement of the amount of turbulence allowed in SS	0.3
	CoreFlowRate	The number spaces moved forward within the SS core during each step	2
	ISL2WetLabScaling[b]	Provides the precise validation mapping from ISL output to the wet-lab (IPRL) output fraction	7
Parameters influenced by logP_{app}	*A2BJumpProb*	Probability that a COMPOUND will jump from Grid A to Grid B when given the option	0.1
	B2AJumpProb	As above: from Grid B to Grid A	0.6
	B2CJumpProb	As above: from Grid B to Grid C	0.35
	C2BJumpProb	As above: from Grid C to Grid B	0.65
Parameters influenced by protein binding	*BindersPerCellMin*	Min and max for the number of binding agents inside each CELL.	5
	BindersPerCellMax	Simple BINDERS for ECs and ENZYMES for HEPATOCYTES.	10
	MetabolizeProb	Probability that an ENZYME will metabolize a SOLUTE	0.4
	SoluteBindingProb	Probability a SOLUTE will be bound when contact a BINDER	0.5
	SoluteBindingCycle	Number of cycles a BINDER will bind a SOLUTE	25

[a] Parameter values when SUCROSE and ANTIPYRINE were dosed in combination.
[b] In [18], *ISL2WetLabScaling* is called *SoluteScale*.

3.2.2 Representing Sinusoidal Details

The usage and aspect focus taken on hepatic sinusoidal function dictates the simplest 3D ISL component representation for a SS. My primary use has been to match current and future hepatic outflow profiles for many drugs using just one, basic ISL structure. The interesting aspects are those that interact with administered compounds moving through sinusoids. Starting simple, I used exploratory modeling protocols to assess the suitability of a variety of component structures. The first effective structure is shown in Figure 3.2A. A SS represents a unit of sinusoid function, along with its related spatial features. The design and construction validates against the considerable body of qualitative, histological data available in the literature. A SS is a discretized, tube-like structure comprised of a blood "Core" surrounded by three identically sized 2D grids, which together constitute a 3D structure. A more realistic 3D grid is not yet justified and tracking and managing events occurring in concentric 2D grids is computationally simpler than in 3D grids.

Grid A represents sinusoid edges near endothelial cells. Grid B is wrapped around Grid A to represent the endothelial layer. The parameter *ECDensity* controls the size and prevalence of FENETRATIONS within Grid B: in this report, 20% is randomly assigned to FENESTRA, and the remaining space is assigned to CELLS. Grid C is wrapped around Grid B; it represents the Space of Dissé and hepatocytes. They are separate in the liver. Here I represent the Space of Dissé and hepatocytes using one, rather than two spaces Two can be used. However, during the exploratory modeling referred to above, it became clear that conflating their representation into one space, one grid, would provide model detail that was adequate to achieve project goals. The parameter *HepDensity*

controls the relative fraction of Grid C assigned to HEPATOCYTES. I specify two SS classes: S_1 and S_2. Compared to a S_2, a S_1 has a shorter internal path length and a smaller surface-to-volume ratio. I found that it was essential to have sufficient variety of effective SS lengths and diameters. There are several ways to provide that variety. Using two classes of SS was the first strategy that was successful. The merits of alternative, possibly simpler strategies are being explored. As described in Table 7.1, SS length is given by a random draw from a modified gamma distribution having a parameter-specified mean and variance. I used that distribution because drawing from a uniform distribution failed. The same proved true for normal and unmodified gamma distributions.

3.2.3 Interconnecting Sinusoids

The blood supply for one lobule feeds into several sinusoids that merge in stages to only a small fraction of their original number as blood is fed into the lobule's outgoing CV. Interconnections between sinusoids are frequent in the periportal region but are not seen near the CV. The ISL's graph structure reflects that arrangement. Miller et al. (28) and Gumucio et al. (29) subdivide the lobule interior into three concentric zones to distinguish the quantitative difference in structural characteristics and enzyme levels among different zones. Figure 1 illustrates that I do the same. The resulting directed graph maps to a portion of a lobule's sinusoidal network and to half an acinus. In the ISL, each zone contains at least one node. There are three graph structure parameters: 1) number of nodes, 2) number of intra-zone connections, and 3) number of inter-zone connections. Graph connectivity (assignment of edges to nodes) is randomly specified

for each simulation to emulate intrahepatic lobule variability and uncertainty about sinusoidal fine structure. Consequently, each is unique. The number of nodes per zone is always in the order Zone 1 > Zone 2 > Zone 3. As specified in Table 3.1, there are more interconnections between Zone 1 nodes than between Zone 2 nodes. There are no interconnections between Zone 3 nodes. Figure 3 shows an example of a LOBULE graph structure that was used for one simulation run. In (18), I used available literature observations to narrow the variety and range of stochastically allowed graph structures.

For the intended uses, the ISL has what I believe is a minimum set of hepatic features. It is designed to enable new features and functions to be easily added (or removed) without interfering with the function of the existing components and features. It is designed to be reused. Structural and spatial parameterizations that have been validated for one set of drugs can be held constant while drug-specific parameters are tuned to match data for additional sets of drugs.

3.2.4 Compounds as Unique Objects

ISL parameters are grouped into three categories: 1) those that control LOBULE graph and 2) sinusoid structures, and 3) those that reflect the influence of the physicochemical properties (PCPs) of mobile compounds. For convenience, the parameters in the latter are segregated into three groups: those for which compound's a) molecular weight (MW) and size, b) partition coefficient ($\log P_{app}$), and c) unbound fraction are thought to exert a strong influence. Table 3.1 lists the ISL parameter names, descriptions, and values. Additional detail is available in Supplementary Material.

The COMPOUNDS studied are in silico analogues of sucrose, a neutral extracellular

space maker, and four cationic drugs: antipyrine, atenolol, labetalol, and diltiazem. The wet-lab outflow profiles against which simulations have been validated are reported in (1). Each compound is represented using objects that move through the LOBULE and interact with each SS feature encountered. A typical COMPOUND represents many drug molecules: the value and the mapping to wetlab data is controlled by the parameter *ISL2WetLabScaling*. A COMPOUND'S behavior is determined by the PCPs of its referent, along with the LOBULE and SS features encountered during its unique trek from PV to CV. During a simulation cycle, an encountered component "reads" the information carried by a COMPOUND and then uses it to customize its response, in compliance with its parameter values, following some pre-specified or learned logic (24). For example, a CELL detects a COMPOUND in an adjacent space and "reads" that $P_{app} = 0$ (because it is SUCROSE). It therefore does not allow that object to enter. The dosage function simulates the effects of catheters and large vessels; Hung et al. (1) used an inverse Gaussian function for this purpose; I used a modified gamma function.

COMPOUNDS enter the lobule via the PV. After that, they enter a SS in Zone 1 at either the Core or Grid A. Simulated flow occurs only in the Core. The parameter *SinusoidTurbo*, which represents flow turbulence, biases COMPOUND movement in the three spaces in the CV direction. Until being collected at the CV, each COMPOUND has several stochastic options. In the Core or Grid A, a COMPOUND can move within either space, jump to an adjacent space, or exit the SS. Within Grid B, it has three options: move within that space, jump back to the Grid A, or on to Grid C. All COMPOUNDS can move within the EXTRACELLULAR portions of Grids B and C. A DRUG but not SUCROSE can move into CELLS in either Grid B or C. A COMPOUND can exit a SS only from Grid

A or the Core. After a COMPOUND exits a SS in Zone 3, it enters the CV: its arrival is recorded, simulating being collected by a fraction collector.

3.2.5 Representing Cells and Subcellular Components

An unspecified number of cells are represented by objects (called CELLS) in Grids B and C: they function as containers for other objects. A grid location and its container are the limit of spatial resolution. At the start of each simulation, CELLS are placed randomly at some fraction of the available grid locations. A HEPATOCYTE is illustrated in Figure 3.2B. In this report, each CELL contains a randomly specified number of BINDERS in a well-stirred space. No additional detail was needed for validation. The objects within CELLS are below that level, but that condition too can be easily changed when needed.

BINDERS are the INTRACELLULAR components that represent transporters, enzymes, and other cellular material that binds or sequesters drug molecules. A binder within a Grid B CELL can only bind and later releases a DRUG. A binder within a Grid C HEPATOCYTE is called an ENZYME because it can bind DRUG and either release or METABOLIZE it. As done in (18) and (18), an ENZYME can be designed to bind more than one DRUG object. Additional objects or agents can be specified and added as needed without compromising either ISL function or the function of objects already present. Because of the stochastic nature of ISL simulations, each in silico experiment generates a slightly different outflow profile. Typically, a single experiment is comprised of 48 trials that were averaged to represent a referent outflow profile. Because I used an 8-node computer cluster (see below), it was convenient to run a number of experiments

that was divisible by eight. Forty experiments proved to be too few and 56 were unnecessary. Results of the 48 experiments were averaged to represent a referent outflow profile.

Drug partitioning into cells was simulated as follows. If a COMPOUND is adjacent to a CELL container and all other grid points around the COMPOUND are empty, then (when *SinusoidTurbo* = 0) there is an equal probability of it moving in any of the available directions. There is a 1-in-11 chance that it will try to move to the CELL'S location. If the solute decides to move to that location and the parameter *isMembraneCrossing* is true, because $P_{app} \neq 0$, then the COMPOUND enters the CELL. As more COMPOUNDS are successfully simulated using future ISLs, I expect to assign a probability to CELL entry when *isMembraneCrossing* is true.

3.2.6 Similarity Measure

Due to differences between rats and among experimental details, no two liver perfusion outflow profiles are identical. Assuming that experimental, quality control measures are met, there are no specific metrics for deciding when the next in situ outflow profile is sufficiently similar to those already completed to be classified as a repeat experiment. Typically, the decision is made by the researcher based on inspection of the data, guided by experience. Ideally, an ISL outflow profile should be indistinguishable from the wet-lab profile of an acceptable repeat experiment, as judged by an expert (this characteristic is one of ten, targeted ISL design capabilities listed in Supplementary Material). Whereas the precision simulation output of an inductive, mathematical model can be precisely fit to certain selected referent data, the output of an ISL

experiment cannot, because it mimics the referent, in situ experiments, where results vary among repeat experiments. It can, however, be quantitatively compared. I attempted to capture key features of expert opinion in a quantitative Similarity Measure (SM). It compares in silico with the referent outflow profiles as detailed in (18). Briefly, I assumed that the coefficients of variation of repeat observations (referent) are constant. For each referent outflow profile measure P, I create two curves, $P^l = P(1 - d)$ and $P^u = P(1 + d)$. They were the lower and upper bounds of a band around P. Here, d is the standard deviation of the relative differences between each of six replicate in situ experiments and the mean value for a given collection interval, pooled over all collection intervals. An ISL outflow profile was deemed similar to the referent if 80% or more of ISL outflow values were within the band ($SM \geq 0.8$).

3.2.7 Parameter Tuning

There are thirty-seven ISL parameters. Several represent hepatic structures that are not influenced by the PCPs of a referent compound. A major simulation task has been to locate a region of ISL parameter space that can provide parameter vectors that generate biologically realistic outflow profiles that meet the SM criteria for sucrose and the co-administered referent drug. Because SUCROSE is a marker for accessible extracellular spaces, I focused on it first. I tuned the LOBULE graph and SS structure parameter values so that the referent sucrose outflow profiles met the minimum SM value. I then held those values constant and focused on antipyrine, the least lipophilic of the four drugs. I assumed that when co-injected, the initial appearance of sucrose and antipyrine would be similar (within 1 second of each other). When needed, a small lag-time was

used to make that adjustment for each of the referent drugs. I tuned the PCP-sensitive parameters, including parameterizations of all intracellular objects, to improve SM values. I then continued the iterative tuning process to obtain a single ISL from which acceptable outflow profiles would be obtained following administration of any combination of SUCROSE, ATENOLOL, ANTIPYRINE, LABETALOL, and DILTIAZEM.

3.2.8 Consequence of Parameter Changes

In most cases, a change in an outflow profile caused by altering the value of one parameter can be compensated by adjusting a small set of other parameter values. As part of the verification process, it was important to demonstrate that a change in a COMPOUND'S parameter value produced the type of outflow profile alteration that would be expected based on knowledge of hepatic anatomy and drug disposition. With that in mind, I undertook a series of experiments that focused on four, PCP-sensitive, ISL parameters, two that control a COMPOUND'S stochastic movement between grid spaces and the two that control INTRACELLULAR binding and METABOLISM. A change in any of these four values could map to a relative change in a specific hepatic feature resulting from different genetics, differences between individuals, or the consequences of certain hepatic diseases; or it could map to differences in PCPs.

The parameter *A2BJumpProb* specifies the probability that, within each simulation cycle, a COMPOUND will jump from Grid A to Grid B, simulating moving from the sinusoid edge to the endothelial surface. When its value is smaller, the COMPOUND spends less time in Grids B and C, and is more likely to reach the CV sooner. A decrease in *A2BJumpProb* might be expected to map to increased binding to red cells or

an increase in the fraction ionized, for example. *B2CJumpProb* specifies the probability that a COMPOUND will jump from Grid B to Grid C, simulating moving from the endothelial layer to the space of Dissé and hepatocytes. When its value is large, more compounds will move to Grid C. The net result will be that COMPOUNDS tend to spend relatively more time wandering around within Grid C. A larger value might be expected to map to an increased P_{app}, for example. The parameter *SoluteBindingProb* specifies the probability that, within each simulation cycle, an unoccupied BINDER within a CELL will bind a free COMPOUND within that same CELL. A smaller value might be expected to map to a decrease in the fraction of drug that is ionized, for example, or to fewer hydrogen bond donor groups. The parameter *MetabolizeProb* specifies the probability that a bound COMPOUND within HEPATOCYTE will METABOLIZE it rather than simply releasing it. Consider a COMPOUND that is bound to a binder object at the start of a simulation cycle (t_1) and remains bound at the end of that simulation cycle (t_2). Because simulation events within the simulation cycle are not resolved, that scenario can represent several referent scenarios, including a drug being bound to one cell component at t_1, becoming free, and then being bound by another cell component at t_2. A larger value for another drug might be expected to map to increased enzyme affinity, for example, or increased relative abundance of the specific enzyme responsible for metabolism.

3.2.9 Hardware

The experiments were executed on an eight node OSCAR cluster (oscar.openclustergroup.org/) running RedHat's Fedora 5. The distribution of the runs uses MPICH 1.2.7 (www-unix.mcs.anl.gov/mpi/mpich1/). The ISL was compiled using

GCC 4.1.1 against the Swarm 2.2.3 Objective-C libraries (swarm.org/wiki/Main_Page). Each experiment consisted of 48 Monte-Carlo runs with the initial pseudo-random number seed extracted from the machine's clock. The fraction of dose flowing out of the LOBULE was averaged over all runs. The Smodulus_smoothing function from the Rwave (version 1.22) package for R was used to smooth post-peak data using a wavelet window of two or three observations.

3.3 Results

3.3.1 Outflow Profiles for Four Cationic Drugs Plus Sucrose

In the following experiments, DRUGS are always co-administrated with the same amount of SUCROSE. Following are the parameterization processes. Table 3.2 lists PCPs. The profile change induced by a modest change in one of the parameters can be reasonably compensated by adjustments in one or more of the other parameters. Increasing *A2BJumpProb* (Grid A→B) or decreasing *B2AJumpProb* (Grid B→A), for example, can reduce the peak height of an outflow profile, because more COMPOUNDS stay in Grid B: the rate of return of COMPOUNDS to Grid A and Core are reduced. Such change has been used to guide parameter tuning.

I tuned the parameter values for each of the four DRUGS, as described above, until minimally acceptable SM values for the DRUG and co-administrated SUCROSE were achieved (e.g., SM ≥ 0.5). Separately, I iteratively improved the SM values for each of the DRUGS and co-administered SUCROSE. I choose the best LOBULE parameterization for each, four total, for the next stage of improvement. Keeping each LOBULE parameterization constant, I adjusted the remaining parameters to obtain an improved SM

value (SM ≥ 0.7). I then picked one LOBULE parameterization and held it constant while iteratively searching the PCP–influenced parameter space. This iterative process resulted in the LOBULE and SS structures for all four DRUGS specified in Table 3.1. Acceptable outflow profile matches for each of the four DRUGS along with co-administrated SUCROSE are shown in Figures 3.4 and 3.5. The corresponding parameter values for drug-specific physicochemical parameters and SM values are listed in Table 3.2.

Because the molecular weight of diltiazem is much larger than that of the other three drugs, I found it convenient to set its *ISL2WetLabScaling* parameter value to 1. Because ATENOLOL and ANTIPYRINE have similar molecular weights, I used *ISL2WetLabScaling* = 7 for each—the value used for SUCROSE, and used *ISL2WetLabScaling* = 6 for LABETALOL, which has slightly higher molecular weight compared to ATENOLOL and ANTIPYRINE.

DILTIAZEM had the lowest *MetabolizeProb* (0.02 compared to 0.3–0.4 for the other three DRUGS), even though diltiazem is described in the literature being a high clearance drug. Because of DILTIAZEM'S large value of *A2BJumpProb* (0.9) and small value of *C2BJumpProb* (0.2), DILTIAZEM spent relatively more time in Grid C. Consequently, metabolism was extensive even though *MetablizeProb* was low. The high *A2BJumpProb* and low *C2BJumpProb* parameter settings map to diltiazem's physicochemical property of high lipophilicity.

Table 3.2 DRUG-specific ISL parameter values

Category	Parameter	ATENOLOL	ANTIPYRINE	LABETALOL	DILTIAZEM
Parameters influenced by MW	CoreFlowRate	2	2	2	2
	ISL2WetlabScaling	7 [a]	7	6	1
Parameters influenced by logP_{app}	A2BJumpProb	0.1	0.1	0.35	0.9
	B2AJumpProb	0.6	0.6	0.2	0.2 [b]
	B2CJumpProb	0.3	0.35	0.5	0.5
	C2BJumpProb	0.6	0.65	0.5	0.2
Parameters influenced by fraction unbound	*BindersPerCellMin*	5	5	5	10
	BindersPerCellMax	10	10	10	20
	MetabolizeProb	0.35	0.4	0.3	0.02
	SoluteBindingProb	0.35	0.5	0.6	0.35
	SoluteBindingCycle	25	25	25	20
SM Values		0.92	0.81	0.91	0.97
Physicochemical properties	MW	266.3	188.2	328.4	414.5
	logP_{app} [c]	0.14	0.33	2.69	3.53
	fraction unbound [c]	0.74	0.60	0.52	0.28
	pK_a [c]	9.60	1.45	7.40	7.70

Values are listed for the DRUG-specific ISL parameters used to generate the outflow profiles in Figures 3.4 and 3.5. The Similarity Measure (SM) values for those profiles are listed along with, and physicochemical properties of the four drugs

[a] *SoulteScale* for SUCROSE that was coadministered with all four DRUG.

[b] *B2AJumpProb* for SUCROSE that was coadministered with DILTIAZEM was 0.75 to compensate the difference between *ISL2WetLabScaling* values.

[c] From (1)

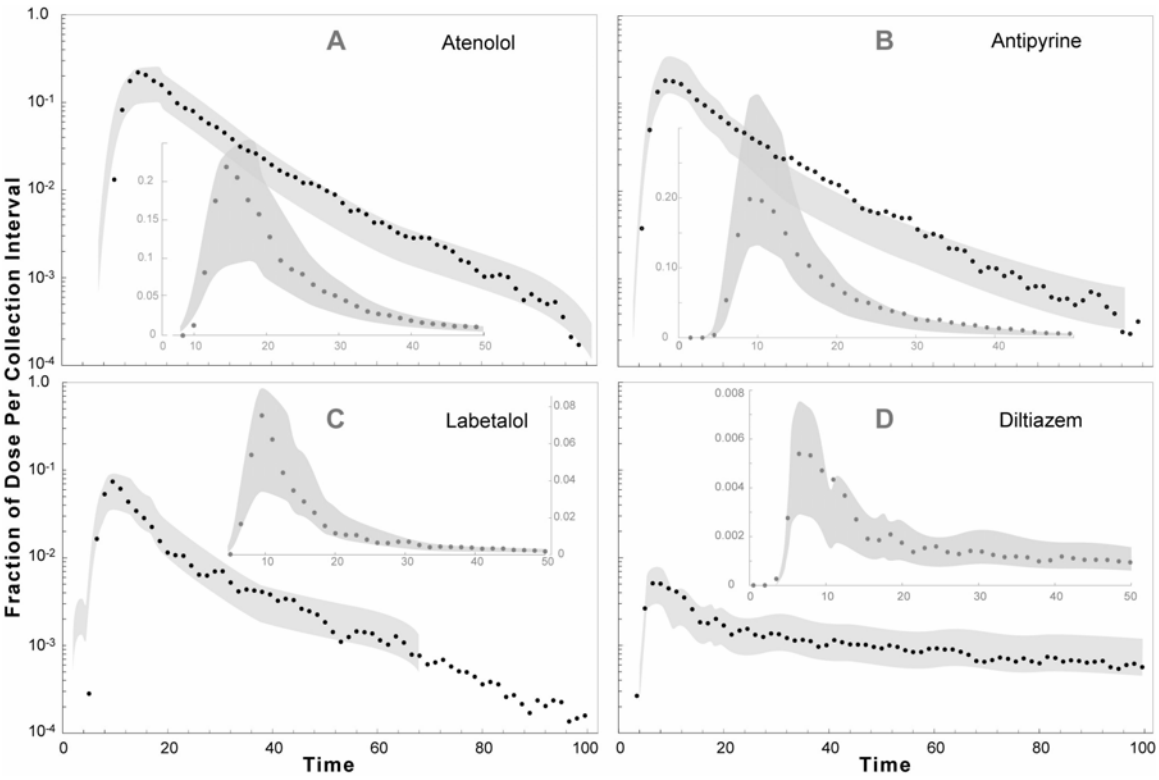

Figure 3.4 Semilog and scatter plots of acceptable outflow profile matches for the four drugs. (A) ATENOLOL, (B) ANTIPYRINE, (C) LABETALOL, and (D) DILTIAZEM were co-administrated with SUCROSE. Plots show the fraction of dose per outflow unit (per ml for the referent) as a function of time (1 unit = 1 second) after dosing with equal amounts of SUCROSE and DRUG. The gray band spans the range for the mean ± one standard deviation as specified in the text. Open circles: wet-lab outflow values for each drug. Dark gray circles: DRUG outflow values using the tuned parameter sets in Tables 3.1 and 3.2. Each ISL datum is the smoothed (window size: 3 values) mean value of 48 independent ISL runs (all stochastic parameter values are changed prior to each run). Although each independent run uses the same LOBULE graph and SSs, their arrangements are randomized prior to each run; consequently, the actual structure of the LOBULE is different for each simulation. Gray band for LABETALOL ends at $t = 70$ because the in situ experiment ended after 70 seconds. SM values: ATENOLOL: 0.92; ANTIPYRINE: 0.81; LABETALOL: 0.91; and DILTIAZEM: 0.97.

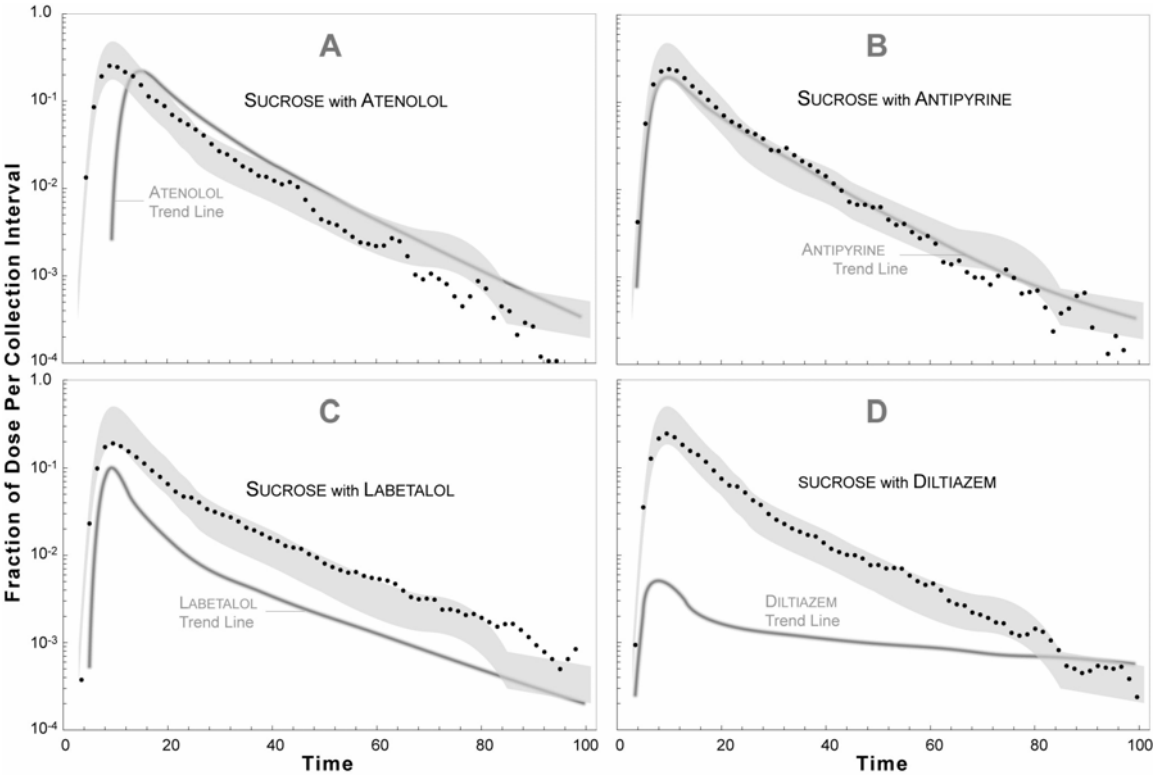

Figure 3.5 Semilog plots of outflow data for SUCROSE when co-administrated with each DRUG in Figure 3.4. Open circles: mean wet-lab outflow values for sucrose. The gray band represents the mean ± one standard deviation for the referent wet-lab sucrose data. The gray curve is the DRUG trend line from Figure 3.4.

3.3.2 Consequences of Changing DRUG-specific Parameter Values

To study how systematic changes in the values of certain PCP-sensitive parameters influenced outflow profiles, I conducted numerous experiments focused on assessing the consequences of parameter changes. I used LABETALOL as the test DRUG because its outflow profile properties were somewhat centric of the four DRUGS. All ISL structural parameter values were unchanged from values listed in Table 3.1. In each experiment, LABETALOL was co-administrated with an equal amount of SUCROSE. For each changed parameter setting, I calculated the ratio of the area under the LABETALOL outflow curve (over 100 simulated seconds) to the area under the curve following the parameter change: this ratio provided a relative, single value measure of the consequences of the change; I refer to the value as Area Ratio.

The first set of experiments studied the four space-jump-probability parameters, *A2BJumpProb*, *B2AJumpProb*, *B2CJumpProb*, and *C2BjumpPrpb* (for LABETALOL in Figure 3.4C it was 0.35, 0.2, 0.5, and 0.5, respectively). For each, I ran sets of simulations for eight parameter values ranging from 0.05 to 1.0. The results showed a consistent trend of either increased or decreased outflow profiles with increasing parameter values. Increasing *A2BJumpProb* or decreasing *B2AJumpProb* had a similar influence on LABETALOL'S outflow profile: the peak decreased. Figure 6A shows results for two *A2BJumpProb* parameter values. For *A2BJumpProb* = 0.05, Area Ratio = 2.5, whereas for *A2BJumpProb* = 0.95, Area Ratio = 0.54. The results were consistent with expectations, for ISL structures and functions. Increasing *A2BJumpProb* (or decreasing *B2AJumpProb*) caused more COMPOUNDS to move into (or less out of) Grids B; from there, more moved into Grid C and that enabling increased METABOLISM.

Increasing *B2CJumpProb* or decreasing *C2BJumpProb* had similar influences on LABETALOL'S outflow profile: the tail end of the profile was lowered. Figure 6B shows results for *B2CJumpProb* parameter values 0.05 and 0.95. Decreasing *B2CJumpProb* from 0.5 (the value for LABETALOL in Figure 3.4C) to 0.05 caused the Area Ratio to increase to 1.5; increasing *C2BJumpProb* from LABETALOL'S value of 0.5 to 0.95 caused the Area Ratio to decrease to 0.83. These results were consistent with the ISL's design and hepatic histology. Increasing *B2CJumpProb* or decreasing *C2BJumpProb* increased the amount of COMPOUND in Grid C, which enabled increased METABOLISM while slowed the return of COMPOUNDS to Grid A and the Core (and thus arrival at CV). The leading edge and the peak of the outflow profile were much more sensitive to changes in *A2BJumpProb* and *B2AJumpProb* than to changes in *B2CJumpProb* and *C2BJumpProb.*

A second set of experiments explored the consequences of changing the parameter values of two CELLULAR components: *SoluteBindingProb* and *MetabolizeProb*. I again used eight parameter values between 0.05 and 1.0. The consequences of changing *SoluteBindingProb* influenced the peak (but only slightly), the shape, and the tail end of the outflow profile. The results in Figure 3.6C are for values of 0.05 and 0.999 (for LABETALOL in Figure 3.4C it was 0.6). Increasing *SoluteBindingProb* decreased the peak, lowered the tail end of outflow profile, and increased METABOLISM: Area Ratio = 0.8. Because of the increased probability of binding in Grids B and C, fewer of the COMPOUNDS that reached Grids B and C could return quickly to Grid A and the Core. Consequently, post-peak values were lower. Increasing binding probability in Grids B and C increased the residence time in those spaces, increasing METABOLISM. Decreasing *SoluteBindingProb* to 0.05 had the opposite effects: peak height increased, profile shape

changed, and the tail end of the profile was raised: Area Ratio = 2.2.

Eight *MetabolizeProb* parameter values were studied. The consequences of change (Figure 3.6D) were essentially the same as changing *B2CJumpProb*. When the value was increased to 0.95 (from 0.3 for LABETALOL in Figure 3.4C), the tail end of the outflow profile was lowered and METABOLISM increased: Area Ratio = 0.88. Because the change only influenced COMPOUNDS that had already reached Grid C and within the HEPATOCYTES, the outflow profile's peak was not significantly influenced. When *MetabolizeProb* decreased to 0.05, the consequences were opposite those just described: Area Ratio = 1.3.

3.3.3 Visualization of Simulation Details

Because of the design and discretized features of the ISL, the details of events as COMPOUNDS trek through various SS can be easily visualized for experimental, verification, and/or validation purposes. An interactive movie of SUCROSE and ATENOLOL moving separately through the ISL following a single combined dose is provided within Appendix.

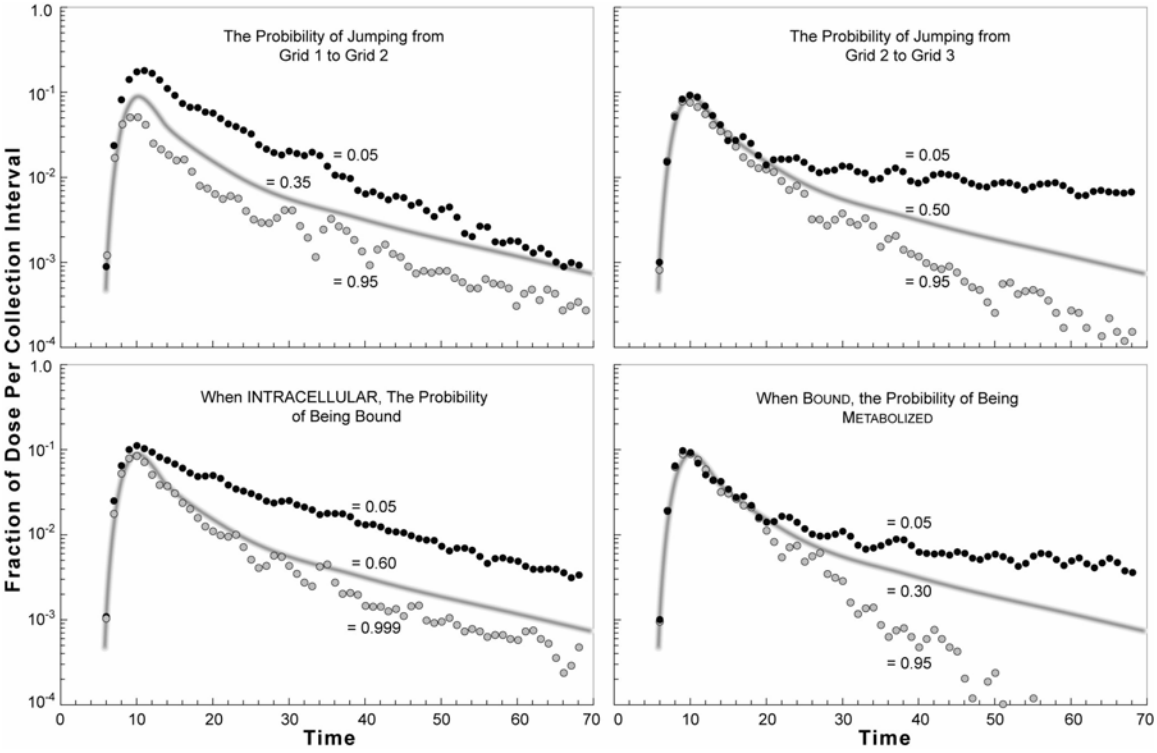

Figure 3.6 Semilog plots demonstrating the consequences of changing four parameter values. In each graph, the continuous gray curve is the trend line for the LABETALOL data in Figure 3.4C, and it provides reference to which the experimental values should be compared. Experiments: LABETALOL and SUCROSE were administered as for Figure 3.4C; however, one parameter value (indicated) was either increased or decreased relative to the control value LABETALOL (shown). Only LABETALOL data are shown because these parameter values are DRUG-specific and do not influence the behavior of co-administered SUCROSE. The parameter values changed determine the probability of one of four events occurring within any one simulation cycle: (A) jumping from Grid A to B (parameter: *A2BJumpProb*); (B) jumping from Grid B to C (parameter: *B2CJumpProb*); (C) being bound by a binder (parameter: *SoluteBindingProb*); and (D) being metabolized when bound (parameter: *MetabolizeProb*).

3.4 Discussion

3.4.1 ISLs and Traditional Models

The stated objective of this work was to instantiate a physiologically and mechanistically realistic software device, which could be used and reused to simulate hepatic outflow profiles of any of five compounds, alone or in combination, and study plausible, detailed, generative mechanisms. The results (Figures 3.4–3.6) provide the validation evidence that the objective has been achieved. The physiological and mechanistic realism of the ISL derives in part from component design and connectivity, and from how mobile components—representing drugs of interest—interact with spatially fixed components. Static and dynamic relationships within the ISL, although currently abstract and low resolution, are intended to map to corresponding relationships within a liver during perfusion. In that sense, the ISL is physiological and mechanistically realistic, although it uses parameters and values different from traditional, continuous, physiologically-based models.

It should be emphasized that Hung et al. (1) used a two-phase stochastic, PBPK model and nonlinear regression to fit the referent data in Figures 3.4 and 3.5. The authors state, "this model perfectly fitted the data from the peak to the tail." ISL outflow profiles have provided a description of the original experimental data that is inferior to that obtained by the parsimonious nonlinear regression approach. However, it is evident that ISL simulations, as previously detailed (18), are better placed to meet the constructive goal of describing the causes of PK phenomena, predicting how detailed changes in the system affect those phenomena, and that ISL simulations yield profiles that are fairly similar to those obtained by regression.

The amount of detail included in this version of the ISL is purposefully just enough to represent the outflow profiles of the four drugs and sucrose. For example, the current coarse-grained movement of COMPOUND is modeled differently within and between grids than within CELLS, and can be refined differently for each space when that is needed. For example, I could replace the two probabilities for jumping between Grids A and B with specific details. No transporters were specifically represented, even though I know several varieties exist. They, along with other cellular material, have been conflated into the abstract binder objects. I can implement a spatially detailed CELL MEMBRANE, as done in (30), and imbue it with transporter objects, as done in (31). A random walk mechanism for DRUG transport was adequate to simulate the referent disposition behaviors. More complicated transport phenomena can be added, without compromising the functionality already present when model use requires doing so.

It is too early to make claims about the relative predictive strengths of the ISL class and classical PBPK models. For that, the ISL will need to be challenged to make predictions for new drugs or make predictions of disposition events from experiments that use experimental designs different from that used for validation. Nevertheless, it is important to note that the ISL has been carefully designed specifically to meet such challenges. It has been designed to exhibit ten capabilities (see Supplementary Material) and the results demonstrate that the ISL exhibits nine of the ten. The following are the five that are most relevant to future use of ISLs for prediction.

- It must be easy to *reconfigure* an ISL to represent different histological, physiological, or experimental conditions.

- In order to represent the particular specifics of different experiments, it must be relatively simple to change ISL usage and assumptions, or increase or decrease detail, without requiring significant re-engineering.
- To facilitate the two preceding capabilities, it must be easy to join, disconnect, and replace ISL components: ISL components *articulate*. It must be straightforward to separately validate components. This capability was also demonstrated in (30) and (31).
- The ISL must be usable for simulating the disposition, clearance, and metabolic properties of a wide variety of *compounds*, separately or in the same experiment.
- It must be easy to use a validated ISL as an organ component within a larger, synthetic, physiologically based, whole organism model (not yet demonstrated).

All computational models designed to exhibit the same behavior are effectively identical. However, different models often have different goals, which manifest in differences between components, their organization, and their implementation. For instance, the classical, inductive PBPK models exemplified in (1, 7, 19-22) make parsimony paramount in order to insure identifiably and minimize parameter uncertainty. Such models seek to faithfully describe the time course of PK data in terms of physiological events occurring globally within the liver. Both an advantage and a limitation is that, as a top down approach providing that global view of events, they must ignore intricacies and system heterogeneity. Hence, classical PK models are the best option for precisely describing the global behavior of the system using a mathematically minimal representation. In contrast, the ISL works locally, middle-out, in providing a

description of the articulation of hepatic components. For the task of representing PK data appropriately using minimalist models, the ISL and their ilk will be inferior to classical PBPK models.

The ISL's real power is experimental exploration of what biochemical, morphological, and physiological alterations in subcellular and local mechanisms may have on PK phenomena. In (2) I demonstrated the consequences of simulated lobular morphological and physiological changes. The outflow profile changes in Figure 3.6 show outflow profile changes resulting from event changes within each SS. The different profiles could represent the consequence of disposition of the same drug in a different liver in which changes in subcellular and local mechanisms are caused by biochemical, morphological, and physiological alterations. Alternatively, the different profiles could be caused by changes in subcellular and local mechanistic details that are a consequence of dosing with a different drug. In the latter case, the specific ISL parameter change, the probability of jumping from Grid B to Grid C, for example, may be traceable to differences in a specific subset of PCPs. Hence, if the goal is to predict altered PK caused by altered pathophysiology or a change in drug structure, or to discover why liver disease leads to a particular PK phenomena rather than another, then the ISL class is likely to be among the best options. Further, if the researcher intends to also use a model to explore the PK consequences of specific mechanistic changes that may result, for example, from disease or co-administered drugs, then models of the ISL class are expected to have advantages. However, at this time, the ISL class has not been used for such purposes and so I must await the supporting evidence. The discussions of comparisons of inductive PBPK models and synthetic models of the ISL class in (18)

(including Supplementary Material) and (10, 11, 16, 23) address other ways that the ISL class differs from traditional PBPK models.

3.4.2 Disposition Mechanisms

Features of hepatic architecture at several levels of detail are encountered by all xenobiotics traveling through the liver. My approach (Figures 3.1 and 3.2) has been to represent some of those common features within Sinusoidal Segments. Probabilistic parameters control how each spatially fixed component interacts with mobile objects representing drugs, and these interactions can be tailored to the PCPs of different compounds. Fixed ISL components "read" PCP information carried by each COMPOUND before initiating an interaction. Consequently, several different DRUGS can be present during the same simulation. SUCROSE plus the four DRUGS can be administered together and the outflow profile of each will be experimentally indistinguishable from referent profiles.

The mechanisms within the ISL are relevant for other drug classes because the rules governing interaction between a mobile object and a fixed component can range from nil to extreme. Sheihk-Bahaei et al. (32) have provided an example of applying exactly the same methods to two very different compounds: salicylate and enkephalin; in (32) the study was extended to include taurocholate and methotrexate. They tuned in silico hepatocytes to simulate the uptake and biliary clearance of those four compounds in sandwich-cultured hepatocytes. Metabolic enzymes and transporters are specifically represented. It is at this level of detail that CYP enzyme and transporter polymorphisms could be represented within an ISL, where separate enzyme and/or ISL versions simulate

a specific polymorph.

With additional ISL validations comes the opportunity to anticipate or predict ISL parameter values for a new compound, based on its PCPs. There is a mapping between the space of ISL parameterizations and the space of compounds' PCPs. The expectation is that as the number of satisfactorily validated COMPOUNDS increases, identifiable patterns will emerge in these mappings. In that situation, the ISL can be given methods to exploit and use posited mappings to anticipate a reasonable set of parameter values for a new, previously unseen DRUG, given only its PCPs. An example of how this can be done was recently presented (24).

The conceptual separation of hepatic form and function and their subsequent instantiation within the ISL has been somewhat arbitrary. For example, extracellular access to Grid C is regulated by the density of FENESTRATIONS within Grid B combined with the probability of jumping to Grid C when a FENESTRATION is encountered. A set of different FENESTRATION densities and transition probabilities values may give the same Grid C access for a given COMPOUND. This example illustrates that different interactions between LOBULE structural components and mobile COMPOUNDS of the same type can similarly impact disposition within the ISL. Sets similar to the preceding example exist for each probabilistic parameter influenced by PCPs. Importantly, the size of these sets shrinks when the same ISL is also validated sequentially against data for a second compound. It shrinks further for each additional, successful validation.

3.4.3 Tuning and Refining ISLs

The iterative model refinement method is the heart of exploratory modeling. When

faced with the task of building a scientifically relevant model in the face of significant gaps in the body of knowledge used to guide the modeling, parameterizations and model components must strike a flexible balance between too many and too few. Too many can imply redundancy or a lack of generality; too few can make the model useless for practical research. With the ISL, I have attempted to create a minimal model, both abstract enough to be meaningful and concrete enough to provide a foundation from which future ISL descendents can develop. As the ISL evolves, I expect to refine the above, compensatory parameter sets such that they guide further experimentation on the model and the referent. For current use, the ISL is striking that balance between too many and too few components and parameters.

The structural and microarchitectural details for each simulation are nondeterministic. Graph structure has a major influence on ISL outflow profiles (18). Even though the number of nodes per zone and the number of intra- and interzone edges are specified, the connectivity pattern is determined randomly at the start of each simulation. SS structures for a given set of parameter values are also stochastic: the actual structures of the 72 SSs used for each ISL run in Figures 3.4 and 3.5 are highly constrained yet typically different, and their properties can form clusters as a consequence of the discretization within an ISL. Because of the many forms of discretization, there is variability between ISL instantiations and that causes outflow profiles to be somewhat different. That variability reflects, and to some extent simulates, some of the uncertainty in my knowledge about the exact details of the generative mechanisms and wet-lab experiments from which the data were collected.

Each 48-trail experiment models a single rat liver. Two experiments using the same

parameter vector are unlikely to vary enough to represent two different individuals. Instead, they effectively simulate repeat experiments on a single liver where, unlike with the referent livers, hepatic function is not modified a previous experiment. To effectively simulate interindividual variances, some modification of the parameter vector is needed. For variance in liver size and blood flow, I can modify *CoreFlowRate*, the number of SSs, the geometry of the SSs, and *SinusoidTurbo*. HEPATOCYTE and ENDOTHELIAL CELL density, binder, and METABOLISM parameters can be changed to vary the overall amount of HEPATIC function and that can simulate interindividual differences in age, damage to the liver, or previous exposure to compounds that induce enzymes. Various types of liver damage can also be simulated by changing the ratio between the two types of SS or by making the SSs more or less constricted and tortuous. The connectivity of the SS graph can also be changed to increase or decrease intra-lobule mixing, or increase or decrease the range of possible path lengths.

I do not measure or record frequencies of path "lengths" or path properties encountered by COMPOUNDS as they move from PV to CV. If I could obtain such data during liver perfusions, I might expect a smooth and asymmetric distribution ranging from short to long treks (21). In the ISL, if the treks taken by COMPOUNDS form clusters, then the outflow profile will evidence that heterogeneity by exhibiting post-peak bumps in the outflow profile like those seen in Figure 3.6. Such patterns are a natural consequence of how I discretized the ISL. For example, decreasing the probability of moving from Grid A to B (Figure 3.6A) reduces COMPOUND resident time within Grids B and C. The bumpy tail of the outflow profile reflects primarily that, when COMPOUNDS are confined more to Grid A and the Core, the available path options from PV to CV are

reduced, and the frequency distribution of paths taken becomes less smooth (data not shown). Having more options smoothes the curve; decreasing the probability of moving from Grid A to B reduces options. Similarly, setting the probability of moving from Grid B to C (Figure 3.6B) to a high value increases the fraction of resident time spent by COMPOUNDS within Grid C. The lowered, bumpy tail is a consequence of reduced path options: the frequency distribution of paths taken from PV to CV became somewhat clustered because options were reduced.

The experimental outflow profiles in Figures 3.6B and 3.6D show that two different mechanistic changes can cause similarly altered outflow profiles. Consequently, differences between two outflow profiles, as discussed above, can have multiple, equally valid, detailed explanations. I speculate that the same is true in situ and in vivo.

In Methods, I stated that BINDERS are INTRACELLULAR components that collectively represent transporters, enzymes, organelles, and other cellular components that bind or sequester drug molecules. DILTIAZEM "sees" more BINDERS than do the other three DRUGS. As mentioned in Results, the probability of a bound DILTIAZEM being METABOLIZED within a simulation cycle is very small: 0.02. That value may initially seem inconsistent with reports that diltiazem has a large, relative, intrinsic hepatic clearance. However, by positing that a large portion of DILTIAZEM binding within the ISL maps to intracellular ion-trapping and microsomal binding, I see that the ISL mechanisms and event parameterizations for DILTIAZEM, although less specific and more abstract, are completely consistent with mechanisms that motivated traditional PK models (1, 22). I could replace BINDERS with two (or more) new INTRACELLULAR object classes, one representing intracellular ion-trapping and microsomal binding, and another

representing drug metabolism. However, because current model use did not require that detail, I elected not to do so.

3.4.4 Designed ISL Capabilities

I have demonstrated and validated an ISL that provides new methods for achieving the PBPK vision. It proved essential that the ISL have the ten capabilities listed in Supplementary Material. ISL observables have been designed to be consistent with in situ observables. This design has enabled clear mappings between in vitro and in silico components and mechanisms that can be harnessed for prediction. As in situ, the ISL behaviors that have emerged during simulated perfusion experiments are the consequences of local mechanisms—local component interactions. Simulation details as they unfold are visualizable, measurable, and comparable to those of in situ perfused livers (see Supplementary Material). The ISL has been designed and constructed so that it can become the liver component in a larger whole organism model. To enable that adaptability and use the same ISL for several different compounds, I found it essential that components be autonomous and easily reconfigured. By making components and system dynamics discrete, it was relatively simple to increase or decrease detail as needed to simulate outflow profiles of additional COMPOUNDS, without requiring significant ISL reengineering. Because identical PK profiles can be a consequence of different mechanisms, I have made it easy to join, disconnect, and replace ISL components to explore alternate explanations.

4 Predictions of Hepatic Disposition Properties Using a Mechanistically Realistic, Physiologically Based Model[1]

4.1 Introduction

The In Silico Liver (ISL) (18) (33) (Figure 4.1) is a first generation example of synthetic, discrete, physiologically based, analogues that are intended for refining, exploring, and testing hypotheses about the details of hepatic drug disposition. Autonomous components represent spatial aspects of hepatic organization and function. Different drugs can be represented and studied simultaneously or separately. Each ISL component can interact uniquely with any drug-representing object that enters its local environment. The consequences of simulated systemic and local interactions can be measured and studied simultaneously, analogous to how wet-lab experiments are conducted.

[1] This chapter has been accepted for publication in Drug Metabolism and Disposition

In (33), the simulated hepatic disposition of atenolol, antipyrine, labetalol, and diltiazem, along with co-administered sucrose, used only one parameterized, ISL structure for all compounds. A subset of components interacted differently with the particular compounds. Monte Carlo ISL variants simulated compound-specific outflow profiles that matched referent profiles. The results supported two hypotheses. 1) The mappings in Figure 4.2 between ISL components and corresponding liver components were sufficiently realistic for the stated model use. 2) The simulated drug-ISL-component interaction events mapped to corresponding hepatic disposition events. A goal for this project has been to discover and verify ISL counterparts to the relationships between drug physicochemical properties (PCPs) and pharmacokinetic (PK) mechanisms, and use them to make predictions.

A vision motivating research on this class of models is identical to one that has motivated development of traditional physiologically based (PB) PK models: by "accounting for the causal basis of the observed data, … the possibility exists for efficient use of *limited drug-specific data* (italics added) in order to make reasonably accurate predictions as to the pharmacokinetics of specific compounds, both within and between species, as well as under a variety of conditions" (3). PCP-sensitive, PBPK model parameters necessarily conflate features and properties of the biology (aspects of histology, etc.) with drug PCPs. In doing so, there is a risk that "the causal basis" becomes obscured due to the conflated biological features that were especially influential in causing some property of the data. Interconnections between sinusoids might be such a feature. In the ISL, because I have built a collection of mechanisms from finer-grained components, I have precise control over conflation, yet the causal basis is

still there in the drug-component interaction logic (axioms, rules). My expectation has been that at some level of granularity, the complexity will be sufficiently unraveled so that the logic for a given ISL component-drug interaction (the "causal basis") will rely heavily on only a small subset of easily specified, drug and biological attributes. In some cases, information about the molecular attributes may be represented adequately by just one or a few PCPs (or, more specifically, knowledge of the drug-environment interaction for which a PCP is a measure). Once I have validated ISL models that exhibit some of these characteristics, I will have moved closer to the above vision of being able to accurately anticipate the hepatic disposition properties of a new compound, given only its molecular formula and structure (and/or validated PCPs).

Are the mechanisms in the current ISL sufficiently realistic to enable such predictions using only limited data? I hypothesized that the differences in the tuned values of the ten PCP-sensitive ISL parameters were due primarily to differences in PCPs. If that is the case, then quantitative PCP-to-ISL mappings exist and, when discovered, should enable us to predict the disposition of a new compound given its PCPs. Failure to provide better than ballpark predictions could be taken as evidence invalidating that hypothesis and possibly aspects of the ISL mechanisms. The goal of this project has been to suggest quantitative PCP-to-ISL mappings and then use them to predict the disposition properties of prazosin and propranolol given only their PCPs.

Three different methods were used to implement quantitative PCP-to-ISL mappings: linear regression, an artificial neural network, and a well-established fuzzy clustering algorithm. Each method predicted the PCP-sensitive, ISL parameter values for prazosin and propranolol. Those predicted values were then combined with the already-validated,

drug-insensitive ISL parameter values. Simulation of the resulting ISLs gave three different, independent versions of expected hepatic disposition details for prazosin and propranolol along with expected liver perfusion outflow profiles. Those profiles were surprisingly good matches to the observed profiles, strengthening the ongoing validation of ISL mechanisms and supporting the existence of the mapping relationships in Figure 4.2. Together they represent an important advance in our ability to predict PK properties for drugs given only their structure.

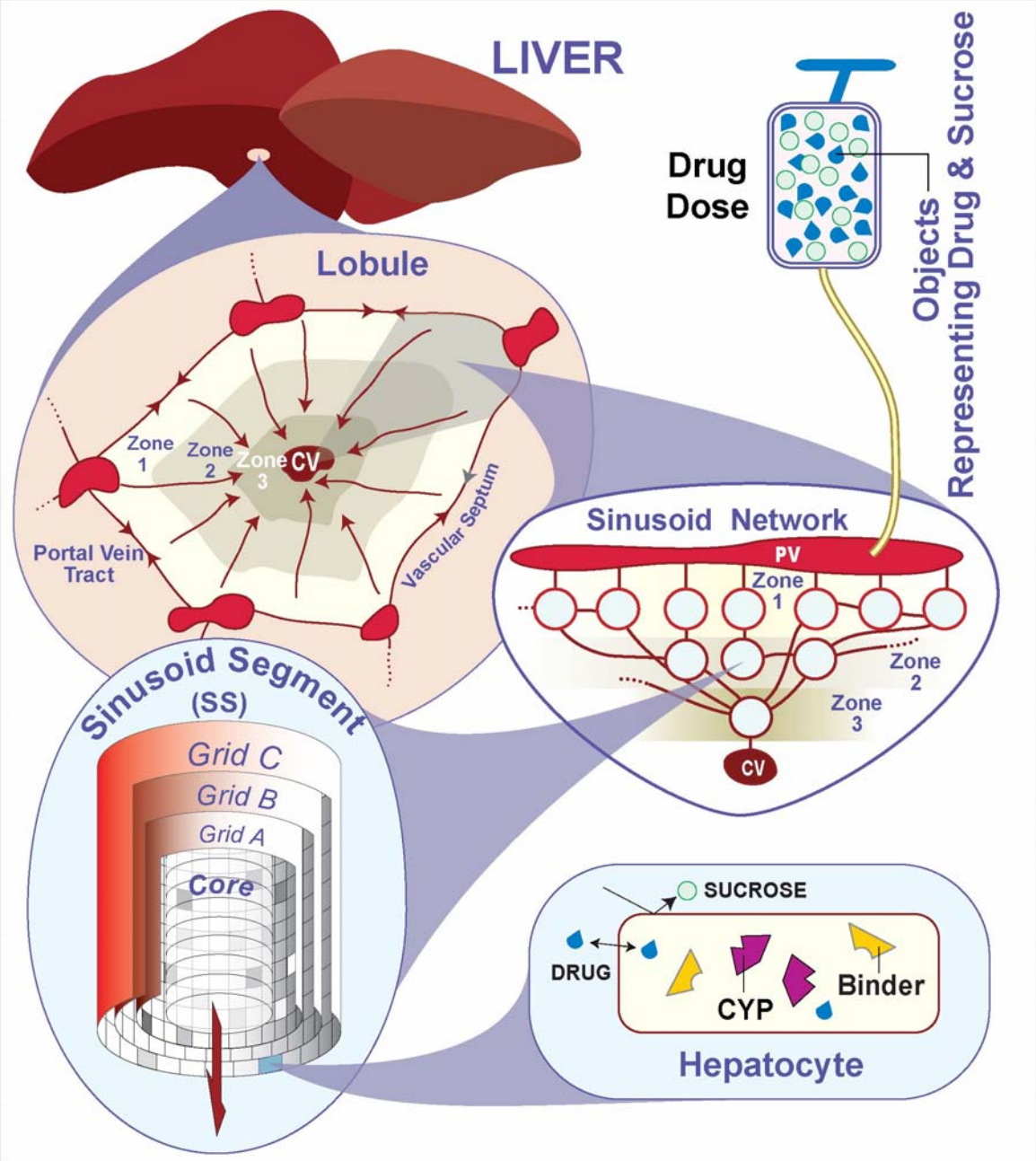

Figure 4.1 Illustrations of ISL model components. The six levels of ISL organization, from largest to smallest, are: 1) the ISL experiment system, 2) SINUSOID network, 3) SINUSOID segment, 4) SINUSOID spaces, 5) CELLS and EXTRACELLULAR domains, 6) ENZYMES and binders. Upper left: a schematic of a cross-section of a hepatic lobule. Blood flow enters lobules via terminal portal vein tracts, goes through sinusoids in three concentric zones, and drains into branches of a common central vein (CV). The three different zones have quantitative differences in structural characteristics and DRUG

interaction capabilities. Upper right: objects representing drug and sucrose are injected as a bolus into a simulated catheter that feeds into the portal vein tracts. Middle-right: a portion of the SINUSOID network is represented by an interconnected, three zone, directed graph. Objects called SINUSOID segments (SSs) are placed at each graph node: they simulate sinusoid functions. Lower-left: a schematic of SS component organization: Grids represent spaces and can contain objects representing the functions associated with a portion of SINUSOID space. Objects representing DRUG and SUCROSE enter and exit a SS via the Core and Grid A. From Grid A, a COMPOUND can access the other spaces. Grid locations have properties that govern their interaction with mobile COMPOUNDS. The potential for heterogeneous properties within different grids is illustrated by different shadings of grid locations. Objects functioning as containers (for other objects) are used to represent cells, and can be assigned to any grid location. The Core represents blood flow; Grid A represents the sinusoid rim; Grid B represents endothelial cells and fenestra; Grid C represents all other spaces, including the Space of Dissé and hepatocytes. Lower-right: A HEPATOCYTE container: objects representing all needed intracellular components can be placed within. All enzymes that metabolize drug are represented by ENZYME objects (e.g., CYP); cell components that simply bind, sequester, and release drugs are represented by BINDERS. Additional binders representing transporters and other capabilities can be easily added as needed. Cell containers can distinguish between DRUG and SUCROSE; they do not allow SUCROSE to enter. Bile was not needed for these simulations, but it can be represented easily as an extension, as described in (24).

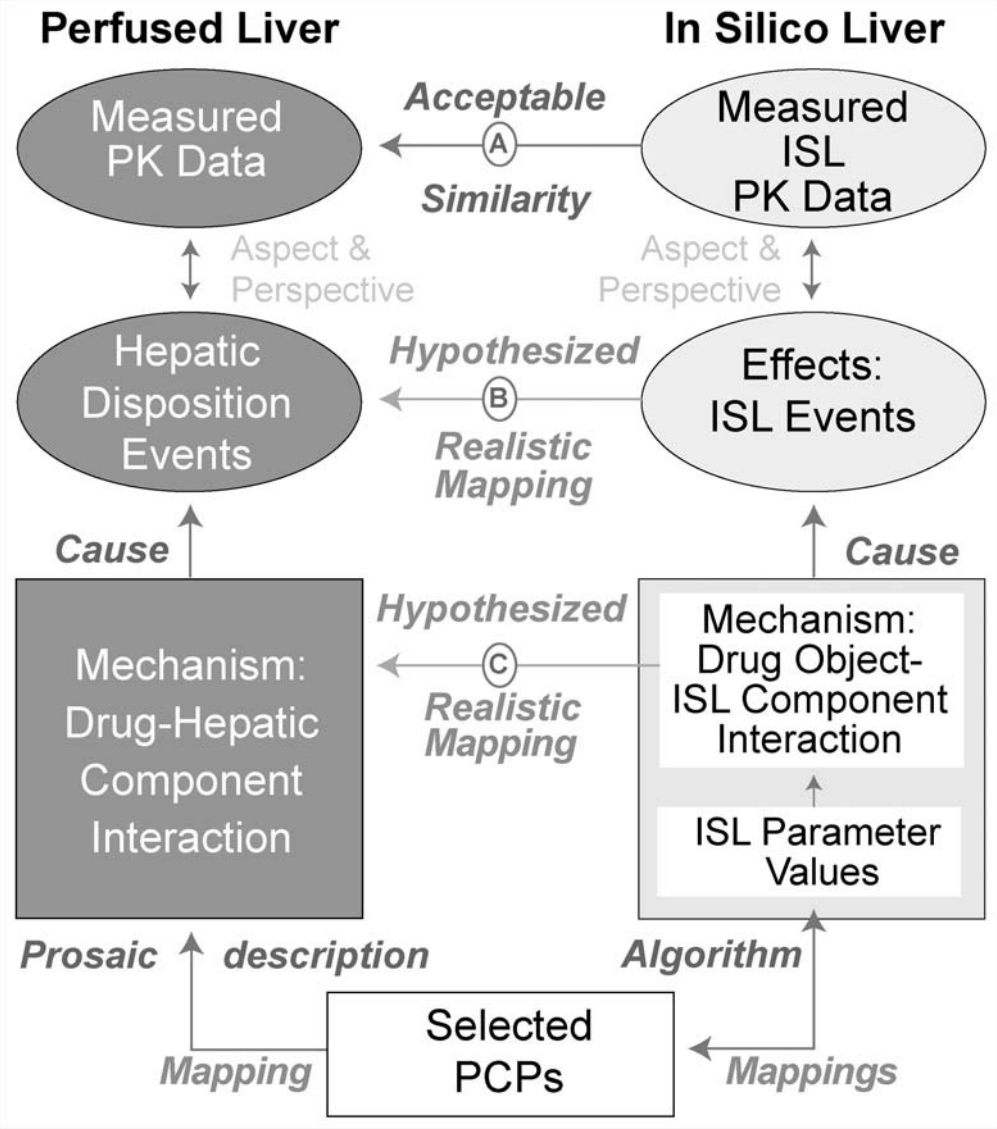

Figure 4.2 Illustration of relationships between ISL mechanisms and its components, along with relationships between the ISL and referent, perfused liver counterparts. To the extent that mappings A, B and C are realistic, the relationship between a drug's physicochemical properties (PCPs) and specific hepatic disposition events will have an ISL counterpart. Discovering plausible methods to make a quantitative mapping between a drug's PCP and PCP-sensitive ISL parameter values will allow prediction of DRUG–specific, PCP-sensitive ISL parameter values, given only a new drug's PCPs. ISL simulations using those predicted parameter values stand as a prediction of the expected disposition behavior of that new drug.

4.2 Methods

To distinguish clearly in silico compounds and processes from corresponding hepatic structures and processes, I use SMALL CAPS when referring to the former.

4.2.1 Model Structure and Design

ISL structure is illustrated in Figure 4.1. Brief summaries of its novel structure and design follow for convenience. Consult (18) and (33) for additional details. An ISL mimics essential features of the in situ perfused rat liver systems used in (1) to study the hepatic disposition of the cationic drugs. I assumed that rat, hepatic, secondary unit function is similar throughout a normal liver, and that, on average, the function of different lobules would be indistinguishable.

I therefore represented the entire liver as a collection of lobules. Blood flow is represented as entering lobules via portal vein tracts and vascular septa, and draining into branches of a common central vein (CV). The ISL represents each lobule as a network of cell-lined sinusoids connected by solute flows. The network is a directed graph; each node (or vertex) represents a cell-lined sinusoid and a link (or edge) between two nodes represents directional flow. For each LOBULE simulation, the graph structure is determined stochastically within parameter-specified constraints. Objects representing sinusoidal spaces and functions, along with related spatial features, are called “Sinusoidal Segments” (SSs). One SS is placed at each graph node. The parameters that control LOBULE form and function for the experiments in this report are listed in (33) along with all other ISL features. For convenience, they are also listed in Table 7.1 of the Appendix. Only ten parameters were modified during the course of this study. Except as noted

below, all other parameter values were the same as those specified in Table 7.1. The meaning of each parameter is in its name. For example, parameter *C2BJumpProb* is the probability that a COMPOUND, when given the option, will move from Grid C to Grid B; parameter *BindersPerCellMax* specifies the maximum number of binders that can be contained in any one CELL.

A SS is a tube-like structure with a blood "Core" surrounded by three identical size spaces in the form of 2D grids. Grid A represents the sinusoid rim. Grid B, which represents endothelial cells and fenestra, is wrapped around Grid A. Grid C, which represents all other spaces, including the space of Disse and hepatocytes, is wrapped around Grid B. Different sinusoids can experience different flows and have different surface-to-volume ratios. Such heterogeneity is represented by specifying two SS classes: S_1 and S_2. Compared to a S_2, a S_1 has a shorter internal path length and a smaller surface-to-volume ratio.

I represented the lobule interior as being subdivided into three concentric zones in order to distinguish the quantitative difference in intralobular structural characteristics and, when needed, enzymatic and transporter gradients. The number of directed graph nodes per zone is always in the order of Zone 1 > Zone 2 > Zone 3. Each LOBULE zone contains at least one node. Because interconnections between sinusoids are frequent in the periportal region and are rare near the CV, there are more interconnections between Zone 1 nodes than between Zone 2 nodes. There are no interconnections between Zone 3 nodes. All of these features are controlled by parameters (Table 7.1), many of which are stochastic to simulate uncertainty and biological variability.

4.2.2 Simulation of Drug Disposition

ISL experiments follow the same protocol used for in situ perfused rat liver experiments. Specifically, in (1), a bolus containing similar amounts of the extracellular space marker sucrose and a cationic compound was injected into the portal vein through a catheter (because of this, drug does not enter the liver as a bolus or a short duration square wave). Consequently, the ISL dosage function is a modified gamma function (given in Table 7.1; see (Hunt et al., 2006) for details) rather than an impulse to simulate effects of catheters and large vessels. Perfusate was collected using a fraction collector. The fraction of the administered dose contained per collection fraction was measured. Hepatic compounds of interest in the ISL are represented as COMPOUNDS (also called DRUGS): they are objects that move through the LOBULE. Each COMPOUND represents a large number of drug molecules and can interact with each encountered SS feature. A COMPOUND's behavior is determined by the PCPs of its referent, along with the consequences of its interaction with the LOBULE and SS features encountered during its unique trek from portal vein tracts to CV.

After COMPOUNDS enter the LOBULE, they can enter a SS in Zone 1 at either the Core or Grid A. Thereafter, until each is collected at the CV, they have several stochastic options. The parameter *CoreFlowRate* biases COMPOUND movement in the Core in the CV direction, simulating blood flow. The parameter *SinusoidTurbo* biases COMPOUND movement in the three spaces in the CV direction, but much less so than *CoreFlowRate*; it simulates turbulent flow. In the Core or Grid A, a COMPOUND can move within that space, jump from one space to another, or exit the SS. When a COMPOUND jumps to Grid B, it can move within the space, jump back to the Grid A, or on to Grid C. All

COMPOUNDS can move within the extracellular portions of Grids B and C. When a COMPOUND encounters a CELL in Grid B or C, it can move into it if allowed to do so by the CELL (it is allowed when the parameter *isMembraneCrossing* is true). When a CELL sees SUCROSE, it also sees that its value of *isMembraneCrossing* is false; consequently, it will not allow SUCROSE to enter. All CELLS in Grid C represent HEPATOCYTES. After a COMPOUND exits a SS in Zone 3, it enters the CV. Its arrival is recorded, simulating being collected by a fraction collector. An animated visualization of these events during a simulation, from start to finish, at both individual ISL and SS levels, for SUCROSE and ANTYPYRINE administered together, is available through Supplemental Data of (33).

The only subcellular functions that were needed to simulate behaviors of the six drugs under study were binding and metabolism. As done in (16, 23, 31), transporters can be added when needed. At the start of each simulation, objects functioning as containers and representing CELLS are placed randomly at some fraction of the available locations in Grids B and C. All cellular components that bind or metabolize drugs, such as transporter, enzymes and other cellular materials, are conflated and represented by simple functional objects called BINDERS. Each CELL contains a randomly specified (within parameter-controlled limits) number of BINDERS in a well-stirred space. A BINDER within a Grid B CELL can only bind and later release a DRUG. A BINDER within a Grid C HEPATOCYTE, referred to as an ENZYME, can bind DRUG and either release or METABOLIZE it. In the later case, the DRUG vanishes; for this report, a METABOLITE replacement is not used.

Because there are 38 ISL parameters (Table 7.1) that specifically influence disposition events, a major simulation task in (33) was to locate a region of ISL parameter space

capable of providing parameter vectors that generate biologically realistic ISL outflow profiles for SUCROSE and each of the four DRUGS. To accomplish that, the LOBULE graph and SS structure was first tuned for SUCROSE; I held those values constant and then tuned the ten PCP-sensitive parameters until I matched outflow profiles for each of the four cationic drugs. For this study, the parameters in Table 4.1 were specified as being sensitive to differences in PCPs. All 28 other ISL parameter values (PVs) are unchanged: they control LOBULE and simulation features that were considered independent of administered COMPOUND and are identical to those specified in (33). Five of the ten parameters are used only for COMPOUNDS that can enter CELLS. I located a specific ISL parameter domain that could generate valid outflow profiles for all five compounds. The values of the PCP-sensitive parameters provided a base from which I could predict ISL PVs for two new simulated cationic drugs: PRAZOSIN and PROPRANOLOL.

Due to the stochastic nature of ISL simulations, each in silico experiment generates a slightly different outflow profile. In this report, the outflow profiles from 48 simulation trials are summed to represent one experiment.

Table 4.1 Tuned PCP- sensitive ISL parameter values for four cationic drugs

ISL attribute or event	ISL Parameters	Atenolol	Antipyrine	Labetalol	Diltiazem
A. Effective DRUG size	ISL2WetLabScaling[a]	7	7	6	1
B. Movement between spaces	A2BJumpProb	0.1	0.1	0.35	0.9
	B2AJumpProb	0.6	0.6	0.2	0.2[b]
	B2CJumpProb	0.3	0.35	0.5	0.5
	C2BJumpProb	0.6	0.65	0.5	0.2
C. Binding to LOBULAR components	BindersPerCellMin	5	5	5	10
	BindersPerCellMax	10	10	10	20
	SoluteBindingProb	0.35	0.5	0.6	0.35
	SoluteBindingCycle	25	25	25	20
D. METABOLISM	MetabolizeProb	0.35	0.4	0.3	0.02

[a] *ISL2WetLabScaling* = 7 for SUCROSE when administered alone or in combination with any drug.

[b] *B2AjumpProb* =0.35 for SUCROSE when co-administered with DILTIAZEM (compensates for difference *ISL2WetLabScaling* values).

4.2.3 Mapping PCPs to ISL Parameter Values

Traditional parameter estimation methods (34) are often infeasible for synthetic, discrete, agent-oriented, biomimetic models of the ISL type, because many of the mathematical assumptions underlying these traditional methods are not satisfied. Therefore, in addition to simply stating and describing the methods used, I also state considerations and provide the rationale.

It is widely held that a mapping exists between the hyperspace of drug PCPs and the hyperspace of measured PK properties (35). Correlation results, as in (1), frequently demonstrate relationships between a PCP and values of one or more fitted PK parameters, providing evidence of such mappings. Validation of the ISL for five compounds provided evidence for mappings A, B and C illustrated in Figure 4.2: the mapping between A) ISL outflow profiles and their in situ counterparts; B) ISL events and hepatic disposition events; and C) local ISL mechanisms and their drug-hepatic component counterparts. Because a mapping exists between PCPs and the causal hepatic mechanisms, I infer that a logical mapping can also be constructed between drug PCPs and the PCP-sensitive subset of tuned, ISL parameters of the four compounds in Table 4.1.

How can I establish specific, quantitative mappings so that ISL PVs for a new drug can be approximated directly based on its PCPs? How can I accomplish the task when I have "only limited drug-specific data"? An example of how this can be done for a synthetic, discrete event, agent-oriented model having design features similar to the ISL, was recently presented (24), and theirs is one of the methods that I use.

4.2.4 Methods for Predicting ISL Parameter Values

What is the best method and approach for predicting PCP-sensitive ISL PVs for a researcher-selected set of PCPs? I know from exploratory ISL experiments that the change in an outflow profile caused by changing the value of one parameter can be compensated by modest adjustments in the values of several other parameters: the relationships among ISL PVs and outflow profiles are nonlinear (18). From accumulated PK knowledge, I also know that mappings between PCPs and traditional PK model PVs can also be nonlinear. When attempting to identify and unravel complicated mappings, more data is always preferred. Unfortunately, the reality in the PK domain is that because of the high costs of wet-lab experiments and given the combinatorial size of all the factors involved in any PK mechanism, there will always be a shortage of specific data for any particular sub-system. In addition, most PCPs have their own uncertainty issues. Such considerations make predicting ISL parameters challenging. Before suggesting an answer to the lead question, I also need to take into account the following two issues.

1) A goal for this class of models is to incorporate the parameter prediction method(s) into the simulation system so that the ISL and its components can automatically adjust their DRUG-component interaction rules, "on the fly," using the PCP information carried by each COMPOUND. This will allow the ISL components to adjust their interactions with DRUGS even though the particular COMPOUND has not been previously encountered, and so no validated interaction rule is available.

2) The ISL is not a static model. It is expected to evolve iteratively with continued use. I can anticipate that components and interaction events will be invalidated and

changed. For example, a specific mechanistic detail may be invalidated when the set of targeted attributes is expanded by inclusion of a new attribute (based on wet-lab observations from other experiments). In order to validate against the expanded set of targeted attributes, a current mechanistic ISL feature may need to be replaced by a revised, possibly more detailed version. Consequently, parameter prediction methods need to be easily transferable to new components as the ISL evolves.

Research in quantitative (drug) structure-PK parameter relationships has considered many mapping methods ranging from simple linear regression to partial least squares regression (for examples see (36) and (37)) and nonlinear regression methods, including artificial neural networks (e.g., see (38)). Each has advantages and limitations. For results of the more sophisticated methods to exhibit statistical significance, larger data sets are required, often acquired considering the nature and capabilities of the prediction method. I can use any of these methods as long as I am cognizant of the limitations, especially when working with sparse data: over-fitting is an important limitation, but it does not preclude cautious use of a method. Another is the sensitivity of the model's behavior (e.g., the outflow profile) to changes in predicted PVs. The impact of such limitations depends on the intended use to which the predicted outflow profile will be put. Precise, verifiable predictions are beyond the scope of ISL class models (see discussions in (18) and (33)). At best, an ISL can be used to anticipate the behavior of a new compound in the context of considerable uncertainty, given its assumptions. An example of the latter: the hypothesized mappings B and C in Figure 4.2 are acceptably realistic for the intended use and can be extended to include the new COMPOUND. It

becomes clear that ISL simulations made using predicted PCPs is a rough approximation of what might be expected if all assumptions are valid. Further, this is not a stand-alone prediction: it should be considered only relative to validated ISL behaviors.

With the full weight of the above caveats in mind, a logical strategy would be to use several parameter prediction methods. Simulations made using each of the predicted parameter sets for the new drug can then be compared to in situ outflow profile of that same drug. I can state the following. Consider the hepatic disposition properties of prazosin and propranolol determined from wet-lab experiments and analogous properties of the two in an in silico experiment, where the in silico and wet-lab experimental designs are effectively the same. The relative differences and similarities in simulated disposition properties between the two compounds would approximate those of the wet-lab experiment.

My expectation is that the ISL will be revised and reused to represent the hepatic disposition properties of an increasing number of drugs under a variety of experimental conditions. With improvement, the two hypothesized mappings, B and C in Figure 4.2, will become more realistic because validated, finer granularity versions of the simulated mechanisms will be available. As the simulated mechanisms become less abstract, I can expect a shrinking subset of PCPs (or molecular descriptors calculated from structure information) will become increasingly important in determining the outcome of the referent event. In that case, the differences between predicted, PCP-sensitive PVs, arrived at using different methods may decrease. If so, the scientific usefulness of the simulations can be expected to increase.

Many molecular descriptors and PCPs (measured and calculated) are available (for

examples, see (37)). However, motivated by the desire to keep new methods as simple as possible, while still achieving the objective, I elected to focus initially on just the nine in Table 4.2, the same set used in (33). I elected not to include pK_a values for this study set for the sake of simplicity: I wanted to postpone dealing with the potentially complicated issues of multiple pK_a values and the variety of drug-specific pK_a values that one can encounter in the literature.

Below, I describe how I used three fundamentally different parameter prediction methods.

Method 1: Correlation and Linear Regression. The simplest method for specifying a quantitative mapping is to use correlation and linear regression. I correlated each of four sets of PCP values in Table 4.2 against each of four sets of PCP-sensitive, ISL PVs in Table 4.1. Next, I selected the one correlated pair having the largest correlation coefficient (CC). I then used the corresponding, least-squares regression line to predict a value of that ISL parameter for both PRAZOSIN and PROPRANOLOL, given the value of the selected PCP.

Method 2: Fuzzy c-Mean Clustering. To handle uncertainty that arises during mappings of this type, I used the Fuzzy c-Mean algorithm (FCMA) introduced by Bezdek et al. (39, 40). It can be used to transform descriptive mappings into quantitative relationships. It starts with influence relationships between subsets of PCPs and subsets of PCP-sensitive, ISL PVs from validated ISLs.

Sheihk-Bahaei and Hunt (24) used the FCMA and "only limited drug-specific data" to predict parameter values for their analogue of sandwich-cultured hepatocytes. They made the estimation method available to the model to enable it to reasonably anticipate

(predict) the biliary transport and excretion properties of a new compound, enkephalin, based on the acceptable, already tuned PVs of salicylate, taurocholate, and methotrexate. By assuming the similarity among PCPs extends to model PVs, acceptable values of the latter for enkephalin (with PCPs that are very different from the three other drugs) were predicted that had good similarity measure values. They established a single mapping between thirteen selected PCPs and fourteen PCP-sensitive model PVs (the ISL has just ten). However, correlations between specific PCPs and specific PK parameter values (35) suggest that a small set of PCPs will have a dominant influence on specific drug-component interactions during disposition, an observation confirmed recently by Ng et al. (37). It thus seemed reasonable to expect that different subsets of PCPs will map to the hepatic counterpart of different ISL events, such as moving between Grids B and C and movement into CELLS. Further, the artificial neural network method discussed below provides a mapping between all nine PCPs and all ten ISL PCP-sensitive, PVs. Therefore, for this second prediction method, I elected to seek usable relationships between subsets of PCPs and the ISL parameters controlling specific ISL events.

Consider the six compounds in Table 4.2. Each drug's nine PCP values, combined into a vector, can be viewed as a single point in a nine-dimensional hyperspace. The six vectors or points (also referred to by the name of the compound that the PCPs describe) can be classified into one-to-six groups based in part on their closeness in PCP hyperspace. Any compound can be a member of more than one group: the groups or clusters overlap. The researcher can specify the number of groups, or use the FCMA to search for the optimal number. The FCMA assumes that each point has some degree of membership in each group. It then assigns a membership degree value to every vector

in each cluster. The FCMA inputs are: 1) the data set of n cases to be clustered (for my use, each case is a single drug, a vector of nine PCP values); 2) a parameter *m* known as the fuzzy exponent; and 3) the number of user specified clusters *c*. The first two of these inputs are discussed further in the following sections. As recommended in (40), I always set $m = 2$. The output of a FCMA, illustrated in Figure 4.3, is a *c*-by-*n* matrix, containing the membership values of the fuzzy clusters.

ISL Parameter Categories. There are three categories of ISL parameters. Those that control lobule graph and sinusoid structures are the same for all drugs. Those that control the PCP-sensitive parameters comprise the third category. I need only predict the values of PCP-sensitive parameters. I divided the ten PCP-sensitive ISL parameters in Table 4.1 into four subgroups: those that reflect the influence of a COMPOUND'S A) effective size, B) movement between spaces, C) binding to LOBULAR components, and D) METABOLISM.

I took the nine PCPs in Table 4.2 and divided them into five groups, as follows, according to their perceived similarity in influence: I) MW and volume; II) partition coefficient ($\log P_{app}$); III) unbound fraction (f_{uB}); IV) topological parameters: number of rotatable bonds (RB count) and topological polar surface area (TPSA). These two PCPs are known to be good descriptors for characterizing drug absorption, including intestinal absorption, bioavailability, Caco-2 permeability, and blood-brain barrier penetration (41, 42). And, V) hydrogen bond related parameters: hydrogen bond donor (HBD) count, hydrogen bond acceptor (HBA) count, and tautomer count. These three PCPs are known to be important to a compound's chemical interactions with other solutes and migration capabilities in aqueous materials (41, 43). Because the relationship between

PCPs and ISL parameters are believed to be non-linear, I conservatively assumed that more than one set of PCPs could influence more than one set of ISL parameters. I then made the following mapping assumptions (Figure 4.4) and took them as specifications for moving forward:

1) logP_{app} is the primary PCP reflecting a drug's partitioning ability and so is expected to map to a compound's ability to move between spaces. I also assumed that parameters in PCP subgroup IV, which reflect drug flexibility and topology, also contribute to a compound's ability to move between spaces: in the ISL, that ability is controlled by ISL parameter subgroup B. I specify that a quantitative mapping can be established between PCP subgroups II and IV, and ISL subgroup B. The mapping is illustrated in Figure 4.4.
2) In addition to f_{uB}, parameters in PCP subgroup V are expected to contribute to drug binding to (or sequestration in) various cellular components. I therefore specify that a quantitative mapping can be established between PCP subgroups III and V, and ISL subgroup C (see Figure 4.4).
3) DRUG METABOLISM (ISL parameter group D) simulates a complicated process and so it may map to a broader range PCPs. I included the following: logP_{app} (subgroup II) along with the PCPs in subgroups IV and V.
4) Both MW and volume are known to influence a compound's effective size in biological systems (43). I specify that a quantitative mapping can be established between them, and *ISL2WetLabScaling*.

Fuzzy Cluster Prediction Strategy. In this section, I explain how I used FCMA results to construct quantitative mappings between the PCP and ISL parameter spaces,

and how I used those mappings to predict the PCP-sensitive, ISL PVs for PRAZOSIN and PROPRANOLOL. For clarity, I first describe the prediction method using an example.

The properties in PCP subgroups II and IV are posited to map to the PCP-sensitive subgroup B parameters (Figure 4.4). I illustrate a method to calculate a value of *B2CJumpProb* for PRAZOSIN. The subgroup II and IV values are listed in Table 4.3. I apply the FCMA to these five, three-value, parameter vectors, specifying three fuzzy groups: I get the degree of membership in each cluster listed in Table 4.3. I see that prazosin is primarily a member of a group that contains at least one of the four already-studied drugs. In this case, it is within a group to which labetalol and diltiazem also primary members. Atenolol and antipyrine are primarily in different groups. I assume that the position of prazosin in ISL parameter space relative to the other four drugs is approximately the same as its position in PCP space relative to the other four drugs. As a first approximation, I specify that it is the case. Given that assumption, I can calculate a value of *B2CJumpProb* using PRAZOSIN'S PCP degree membership values in each of the three clusters along with the average values of *B2CJumpProb* for all the members in each of the three clusters. LABETALOL and DILTIAZEM are members of cluster 1. The average *B2CJumpProb* value is 0.5; the value for the only primary member of cluster 2 (antipyrine) is 0.35; the value for the only primary member of cluster 3 is 0.3. Given that data, for PRAZOSIN, I have

$$B2CJumpProb = (0.506)(0.5) + (0.179)(0.35) + (0.316)(0.3) = 0.41$$

Additional, detailed calculations are provided in the Appendix.

The same approach can be used to calculate the remaining three members of ISL subgroup B. The method can also be extended to calculate expected values for the other

ISL parameters. The generalized statement of the method follows.

Parameter Estimation via Analysis of Fuzzy Clusters. In this section, I present an algorithm that uses fuzzy clustering to estimate simulation parameters for a new *situation* given the tuned parameters of several, previously validated *situations*. For additional detail, see (44). In this work, the new situation is the ISL encountering two new types of objects, one representing prazosin and the other representing propranolol. The algorithm starts with the highest possible number of clusters and iteratively decreases the number of clusters until the new compound is placed in a cluster with at least one of the already-tuned compounds. The compounds in that cluster represent those "most similar" to the new compound.

In general, for a set *S*, containing *n data points* $S = \{c_1, c_2, \ldots c_n\}$ (one *data point* corresponds to one *situation*—one drug, in my case, with its associated PCP values), the following algorithm estimates the simulation parameters of a new *situation*, c_{n+1}:

Step 1. Let $q = n$, and $S_{new} = \{c_1, c_2 \ldots c_n, c_{n+1}\}$.

Step 2. If $q = 1$ go to step 4. Else, classify S_{new} into q clusters using the FCMA.

Step 3. If c_{n+1} is not in the same group with *at least* another member, then decrease q to $q - 1$. Repeat steps 2 and 3.

Else, let *G-value* be the number of group-mates of c_{n+1}

Step 4. Call the q groups $G_1, G_2, \ldots G_q$, where $c_{n+1} \in G_1$. Let μ_k be the membership degree of c_{n+1} in G_k. Estimate the needed ISL simulation parameters of c_{n+1} as:

$$\hat{\vec{P}}_X = \sum_{k=1}^{q} \mu_k \cdot \vec{P}_{Gk} \qquad (1)$$

where P_{Gk} is the weighted average parameter vector of all the members of group k:

$$\hat{\vec{P}}_{Gk} = \left(\sum_{j=1}^{m} \mu_j \cdot \vec{P}_j\right) / \sum \mu_j \qquad (2)$$

The accuracy or usefulness of the resulting estimates depends, of course, on how many *data points* similar to c_{n+1} exist in the data set, i.e. the higher the *G-value*, the better the accuracy. Because I am starting with just four compounds, four data points, my expectations must necessarily be modest.

Method 3: Artificial Neural Network (ANN). To establish a traditional nonlinear mapping, I used a simple feed-forward ANN having a single hidden node (45). The neuron transfer function is $y(x) = b/(1 + e^{-ax})$; I chose a = b = 1. I recognized that the small training set would lead to an over-fitted prediction model. Nevertheless, I used it because the method is known to be good at finding patterns and establishing quantitative relationships between data sets based on those patterns. I also used it to provide an indication of what might be expected from an ANN-based, predictive mapping, recognizing that such mappings have often done better than one of the many available multivariate linear mapping options, when the relationships between the data sets are known to be complicated, as is the case here. I used a backpropagation training algorithm that used normalized values of the four sets of nine PCPs in Table 4.2 as inputs; values of the four sets of ten ISL PVs in Table 4.1 were obtained as outputs. I then used the trained ANN to separately predict a full set of ten PCP-sensitive, ISL PVs for PRAZOSIN and PROPRANOLOL, given their full set of nine PCPs as inputs.

Values were normalized as follows: divide each point by its vector Norm. E.g., if drug (point) X has PCPs *A*, *B* and *C*, then it was divided by its length *N*, i.e., $N^2 = A^2 + B^2 + C^2$ and $X_{normalized} = (A/N, B/N, C/N)$. The output values were normalized in the same way.

Consequently, the predictions too were normalized; to be used, the predictions had to be scaled: multiplied by the norm of the training vector.

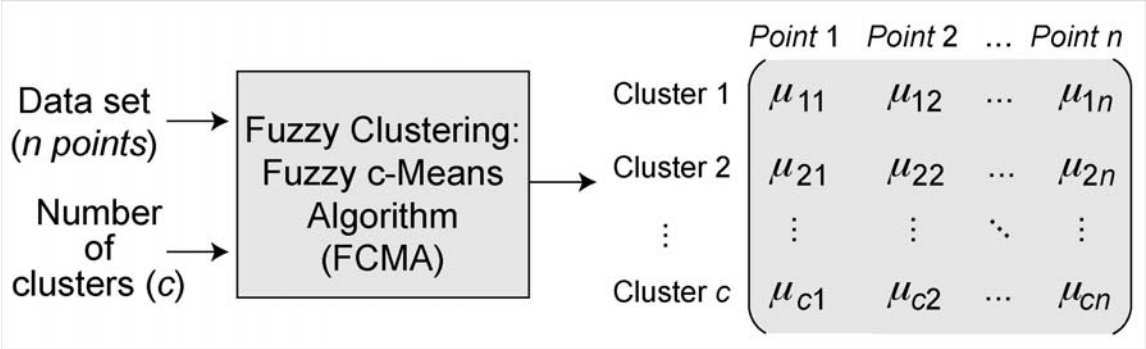

Figure 4.3 Inputs and output of the Fuzzy c-Means algorithm (FCMA). The FCMA locates *c* clusters within the set of *n* points (corresponding to *n* drugs) in PCP hyperspace and assigns a membership degree, μ_{ij}, for each point *i* to each cluster *j*.

Table 4.2 Physicochemical properties of sucrose and six cationic drugs

Category	Parameters	Sucrose	Atenolol	Antipyrine	Prazosin	Labetalol	Propranolol	Diltiazem
I: Size and shape	MW	342.3	266.34	188.23	383.41	328.41	259.35	414.53
	Volume	283.58	260.9	178.29	336.0	314.78	257.82	377.75
II: Partitioning	logPapp[a]	-3.7	0.14	0.33	1.88	2.69	3.10	3.53
III: Binding	fuB[b]	-	0.74	0.60	0.54	0.52	0.45	0.28
IV: Topology related	RB count[c]	5	8	1	4	8	6	7
	TPSA[d]	189.53	84.58	26.94	106.96	95.58	41.49	59.09
V: Hydrogen bound related	HBD count[e]	8	3	0	1	4	2	0
	HBA count[f]	11	4	2	8	4	3	5
	Tautomer count[g]	1	2	0	3	7	0	0

[a] logP_{app}: log of the octanol-water partition coefficient at pH 7.4, 37ºC
[b] f_{uB}: Unbound fraction of each drugs in the liver perfusion medium as determined in (1)
[c] RB: number of rotatable bonds, a measure of molecular flexibility; amine C-N bonds are not considered because of their high rotational energy barrier
[d] TPSA: topological polar surface area is a sum of fragment contributions; O- and N-centered polar fragments are considered (46)
[e] HBD: hydrogen bond donor
[f] HBA: hydrogen bond acceptor
[g] Tautomers are organic compounds that are interconvertible by tautomerization.

Table 4.3 Values used in the text example for predicting a value of B2CJumpProb for prazosin

Top: values comprising PCP subgroups II and IV. Middle: the results of applying the FCMA to the five sets of three PCP values. Bottom: tuned *B2CJumpProb* values for the four already-studied drugs along with the predicted value for prazosin.

PCP Subgroups II & IV	Atenolol	Antipyrine	Labetalol	Diltiazem	Prazosin
RB count	8	1	8	7	4
TPSA	84.58	26.94	95.58	59.09	106.96
logP_{app}	0.14	0.33	2.69	3.53	1.88
Cluster					
C 1	0.007	0.001	**0.930**	**0.895**	**0.506**
C 2	0.006	**0.998**	0.020	0.045	0.179
C 3	**0.987**	0.001	0.050	0.060	0.316
ISL Parameters					
B2CJumpProb	0.3	0.35	0.5	0.5	(0.41)

4.2.5 Similarity Measure

The Similarity Measure (SM) used to compare in silico with the lag-time adjusted, referent outflow profiles is detailed and discussed in (18) and (33). Briefly, I assumed that the coefficients of variation of identically timed observations between t = 1 and t = 100 seconds from repeat experiments (referent) are constant. For each referent outflow profile measure P, I create two curves, $P^l = P(1 - d)$ and $P^u = P(1 + d)$. They form the lower and upper bounds of a band around P. The SM value is the fraction of evenly spaced, outflow profile measures falling within the band. Here, d is the standard deviation (SD) of the relative differences between each of six repeated in situ experiments and the mean value for a given collection interval, pooled over all collection intervals. In (33), where the ISL outflow profile was iteratively tuned to give a good match to the referent profile, I used d = 1 (SD), and an ISL outflow profile was deemed similar to the referent if 80% or more of ISL outflow values were within the band (SM ≥ 0.8). Here, using predicted PVs without any post-prediction refinement, I speculated that having a predicted profile fall within a factor of two of the referent profile could be considered a good result. Nevertheless, I calculated SM values using two reference bands: d = 1 and 1.5, which corresponds to factors of 0.45 and 0.675.

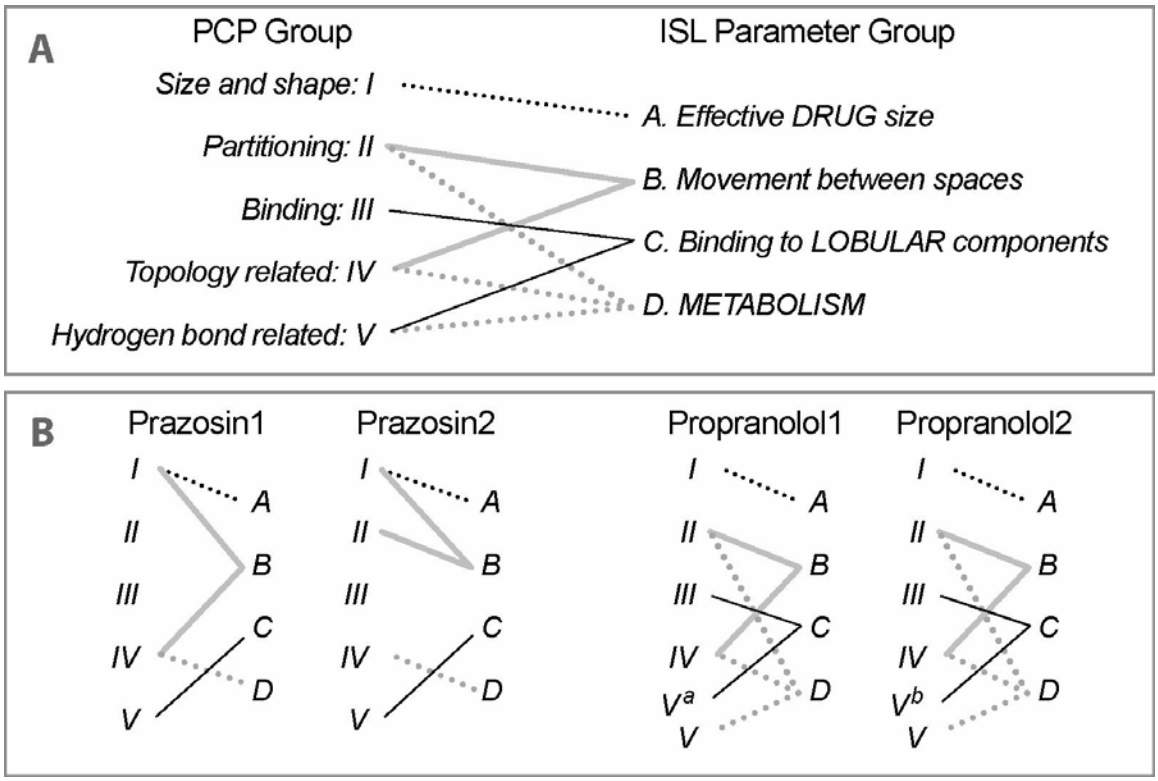

Figure 4.4 Mappings between PCP subgroups and ISL parameter subgroups. The four ISL subgroups are specified in Table 4.1; the five PCP subgroups are specified in Table 4.2. Each line represents a hypothesized strong influence. E.g., the partitioning and topology related PCPs are thought to have a strong influence on the wet-lab counterpart of the four PCP-sensitive ISL parameters values controlling movement between spaces within SSs. The identified PCP subgroups are the ones used by the FCMA to provide the membership degree values used in Eq. 1 to predict ISL PVs for PRAZOSIN and PROPRANOLOL in Table 4.5. The FCMA treats the identified PCPs as having equal influences. The influence of the other PCPs is neglected.

Table 4.4 Correlation coefficients between each PCP-sensitive, ISL parameter and each PCP for atenolol, antipyrine, labetalol, and diltiazem

The least squares regression line corresponding to the largest correlation coefficient (bold) was used to predict that ISL parameter value for PRAZOSIN and PROPRANOLOL, given their value of that PCP. Those ten, predicted parameters values are listed in Table 4.5 and were used to generate the outflow profiles in Figure 4.5.

	ISL2WetLabScaling	*A2BJumpProb*	*B2AJumpProb*	*B2CJumpProb*	*C2BJumpProb*
MW	0.87	0.92	0.87	0.79	**0.94**
Volume	0.82	0.88	0.87	0.77	0.91
logP_{app}	0.83	0.91	**0.98**	**0.97**	0.89
TPSA	0.06	0.04	0.41	0.25	0.10
RB	0.28	0.34	0.50	0.34	0.42
f_{uB}	**0.93**	**0.95**	0.81	0.85	0.92
HBD	0.46	0.36	0.14	0.02	0.31
HBA	0.71	0.75	0.69	0.55	0.80
Tautomer	0.31	0.16	0.44	0.38	0.16

	BindersPerCellMin	*BindersPerCellMax*	*BindingProb*	*BindingCycle*	*MetabolizeProb*
MW	**0.80**	0.80	0.27	0.80	**0.92**
Volume	0.75	0.75	0.25	0.75	0.88
logP_{app}	0.73	0.73	0.08	0.73	0.84
TPSA	0.16	0.16	0.12	0.16	0.06
RB	0.20	0.20	0.20	0.20	0.40
fuB	0.88	0.88	0.08	0.88	0.90
HBD	0.57	0.57	0.40	0.57	0.36
HBA	0.66	0.66	0.43	0.66	0.80
Tautomer	0.45	0.45	0.70	0.45	0.23

4.2.6 ISL Execution

Experiments were executed on an eight node OSCAR cluster (oscar.openclustergroup.org/) running RedHat's Fedora 5. The distribution of the runs uses MPICH 1.2.7 (www-unix.mcs.anl.gov/mpi/mpich1/). The ISL was compiled using GCC 4.1.1 against the Swarm 2.2.3 Objective-C libraries (swarm.org/wiki/Main_Page). The initial PRN seed was extracted from the machine's clock. Each completed experiment was archived and logged with the date and time using a Makefile target. Examination and processing of data from simulations used a combination of Matlab (7.0.0) and R (2.4.0) (www.r-project.org/) scripts, and Microsoft Excel. As done previously (18) (33), I smoothed results from each simulation experiment using the Smodulus_smoothing function from the Rwave (version 1.22) package of R. For the results presented, a wavelet window of three observations at 0.5-second intervals was used.

4.3 Results

4.3.1 Predicted PCP-sensitive ISL Parameter Values

Ninety correlations were performed: four values of each of ten ISL PVs (from having already tuned the ISL for four drugs) with the four values of each of nine PCPs corresponding to those four drugs. Table 4.4 lists the CC values. The regression line for the ten pairs having the largest CCs (bold) was used to predict the ISL parameter value for both PRAZOSIN and PROPRANOLOL listed in Table 4.5.

The PCP-sensitive ISL PVs for PRAZOSIN and PROPRANOLOL predicted using Eqs. 1

and 2 in combination with the mappings specified in Figure 4. 4 are also listed in Table 4.5, along with the ISL PVs for the two drugs predicted using the ANN. Because the referent outflow profiles for prazosin and propranolol included co-administered sucrose, equations 1 and 2 were also used to calculate new PCP-sensitive ISL PVs for SUCROSE. Those values are also listed in Table 4.5. Several examples of using the FCMA method to predict specific ISL parameter values are provided in the Appendix along with the predicted sucrose outflow profiles (Figure 7.6) when SUCROSE was co-administered with PRAZOSIN and PROPRANOLOL.

4.3.2 Hepatic Disposition Predictions

I used the predicted ISL PVs in Table 4.5 along with the previously specified, PCP-independent, ISL parameter values (Table 7.1) to parameterize new ISLs, three for PRAZOSIN and three for PROPRANOLOL. Simulations using each of these ISLs generated the expected outflow profiles presented in Figures 4.5-4.7. One simulated outflow profile is the result of pooling 48 separate Monte-Carlo trials. One trial represents a single lobule and 48 trials represent the whole liver.

As previously discussed (33), the structural and microarchitectural details of the ISL used for each simulation trial were nondeterministic. Even though the number of nodes per zone and the number of intra- and interzone edges were specified, the connectivity pattern was determined randomly at the start of each simulation. SS structures for a given set of parameter values are also stochastic: the actual structures of the 72 SSs used for each ISL run were highly constrained yet different. Because of theses differences, combined with the many forms of discretization, there was variability between ISL

instantiations and, consequently, the raw outflow profiles were noisy. That variability reflects, and to some extent simulates, interindividual hepatic differences along with the uncertainty in my knowledge about the exact details of the generative mechanisms. Increasing the number of averaged simulation runs makes outflow profiles smoother. The additional wavelet smoothing was described under Methods.

As described under Similarity Measure, when matching the simulated outflow profile with lag-time adjusted, *in situ* outflow profiles, I used a band of ±1 SD flanking the original wet-lab outflow profile values as a measure of the similarity of ISL simulations to those wet-lab results. For the outflow profiles generated from the predicted parameter values, two reference bands were used: ±1 SD and ±1.5 SD, and SM values were calculated for each. For these data, a band of ± 1 SD corresponds to a constant, 45% coefficient of variation; ± 1.5 SD corresponds to a constant, 67.5% coefficient of variation. The ±1 SD SM values, in the order PRAZOSIN and PROPRANOLOL, shown in Table 4.5, were as follows. Method 1 (simple linear regression): 0.36 and 0.68; Method 2 (the FCMA): 0.50 and 0.58; and Method 3 (ANN): 0.79 and 0.97. The corresponding ±1.5 SD SM values: Method 1: 0.55 and 0.82; Method 2: 0.77 and 0.72; and Method 3: 0.97 and 0.98.

Table 4.5 Predicted ISL parameter values for sucrose, prazosin, and propranolol

For each of the three methods, the listed parameter values were combined with the PCP-insensitive parameter values. The resulting ISLs were used to generate the expected outflow profiles in Figures. 4.5–4.7

ISL parameter values	Method 1: Simple Linear Regression		Method 2: FCMA			Method 3: ANN	
	PRAZOSIN	'ROPRANOLOL	SUCROSE[A]	PRAZOSIN	PROPRANOLOL	PRAZOSIN	PROPRANOLOL
ISL2WetLabScaling	5.3	4.1	6.0	3.0	7.0	4.6	2.8
A2BJumpProb	0.35	0.52	0.35	0.37	0.86	0.62	0.75
B2AJumpProb	0.37	0.21*	0.22	0.40	0.21*	0.29	0.21*
B2CJumpProb	0.42§	0.50*	0.49	0.41§	0.50*	0.50§	0.50*
C2BJumpProb	0.56§	0.31	0.50	0.48§	0.22	0.42§	0.32
MetabolizeProb	0.34	0.13*	-	0.25	0.12*	0.19	0.09*
BindersPerCellMin	6	12	-	7	6	8	8
BindersPerCellMax	7	14	-	14	13	15	17
SoluteBindingProb	0.47§	0.39*	-	0.41§	0.42*	0.43§	0.42*
SoluteBindingCycles	24§	23*	-	23§	24*	24§	23*
SM values (±1 SD)	0.36	0.68		0.50	0.58	0.79	0.97
SM values (±1.5 SD)	0.55	0.82		0.77	0.72	0.97	0.98

[a]The predicted outflow profiles for SUCROSE co-administered with PRAZOSIN and PROPRANOLOL, obtained using these FCMA-predicted parameter values, are provided in Figure 7.6 in the Appendix along with the predicted sucrose outflow profiles (Figure 7.6) when SUCROSE was co-administered with PRAZOSIN and PROPRANOLOL.

§ The three predicted values of these four prazosin parameters are within 15% of each other.

* The three predicted values of these five propranolol parameters are within 15% of each other.

4.3.3 Drug-specific Prediction Strategies

Although the predicted ISL parameter values from the FCMA method generated good predicted PRAZOSIN and PROPRANOLOL outflow profiles, they were the result of applying simultaneously a common PCP-to-ISL mapping strategy to both drugs. I recognize that prazosin and propranolol have significantly different PCPs. What if I had sought parameter predictions for each drug separately? Would individualized mappings yield improved outflow profile predictions? The mapping between each compound's PCPs and the matched ISL parameter values are not expected to be identical. I explored the consequences of individualizing the mappings. Several dozen mappings were explored. Starting with the strategy in Figure 4.4, I made one mapping change, then predicted parameter values, and simulated outflow profiles. I then repeated the process for a second, third, etc. mapping change. For both drugs, I kept the number of PCP-ISL mappings constant at four, for four subgroups of ISL parameters, as in Figure 4.4. I removed or added single PCP values or single PCP groups. Most of the mapping options resulted in expected drug outflow profiles having SM values for both PRAZOSIN and PROPRANOLOL that were smaller than those in Table 4.5 for the FCMA method. In several cases, the SM values showed little change. Only a few showed improved SM values, but they were modest. The individualized mappings for two of the latter cases (Figure 7.7), along with the corresponding simulated outflow profiles (Figure 7.8) are provided in the Appendix.

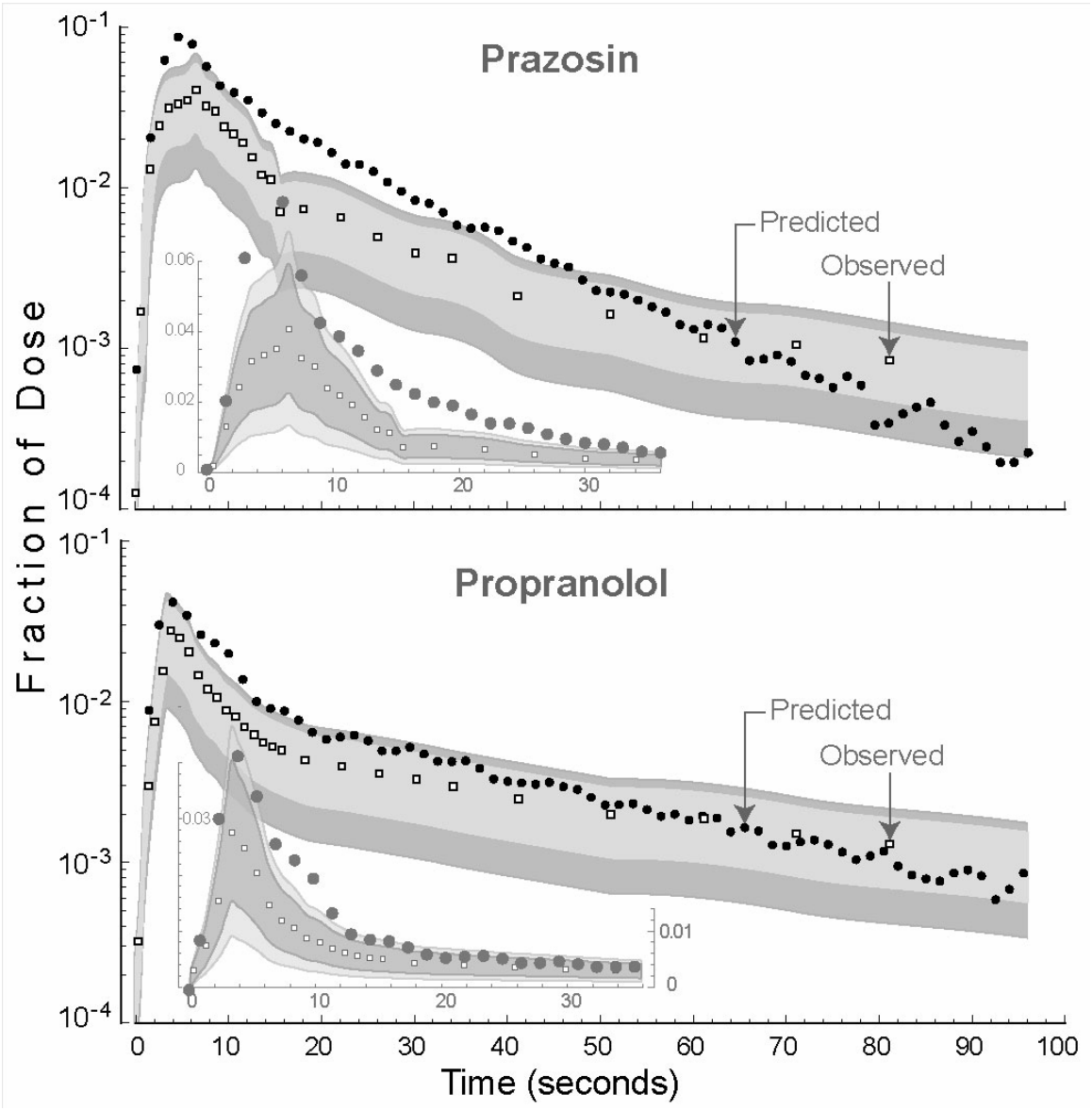

Figure 4.5 Expected outflow profiles (black circles) for prazosin and propranolol using predicted parameter values. These semilog plots show results of simulation experiments for which the PCP-sensitive PVs were predicted using the Simple Linear Regression Method values listed in Table 4.5. All other ISL parameter values were identical to those reported in (33). Open squares: observed, referent wet-lab data. The data are the fraction of dose per outflow unit (per ml for the referent) as a function of time (1 simulation time unit = 1 second) after dosing with DRUG. The light gray band spans the range for the mean of the referent data ± one standard deviation (SD) as specified in the text. The dark gray band spans the range for the mean of the referent data ± 1.5 SD. Each ISL datum is the smoothed (window size: 3 values) mean value of 48 independent

ISL runs.

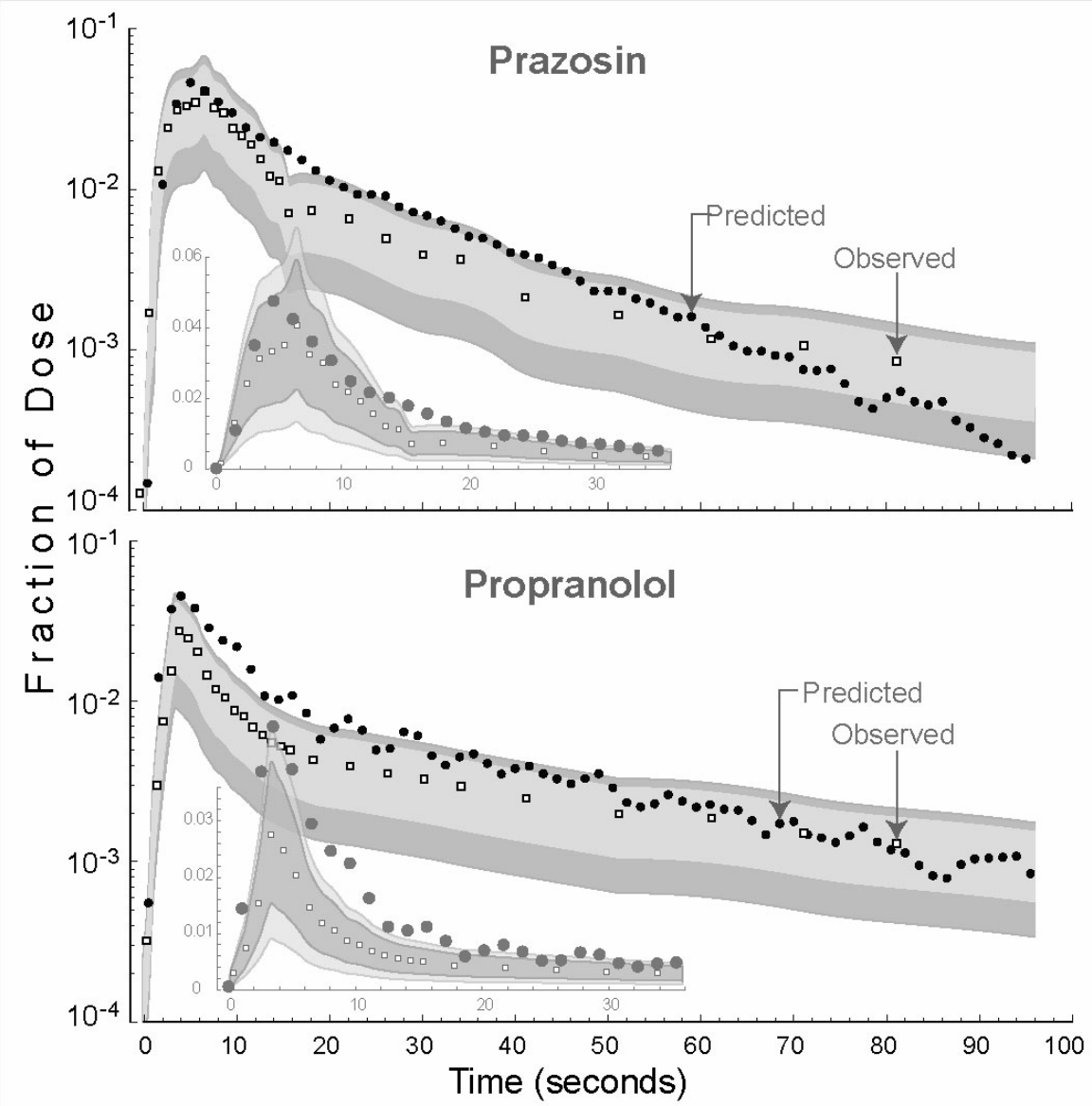

Figure 4.6 Expected outflow profiles (black circles) for prazosin and propranolol using predicted parameter values. These semilog plots show results of simulation experiments for which the PCP-sensitive PVs were predicted using the FCMA values listed in Table 4.5. Otherwise, the graph components are same as specified in Figure 4.5.

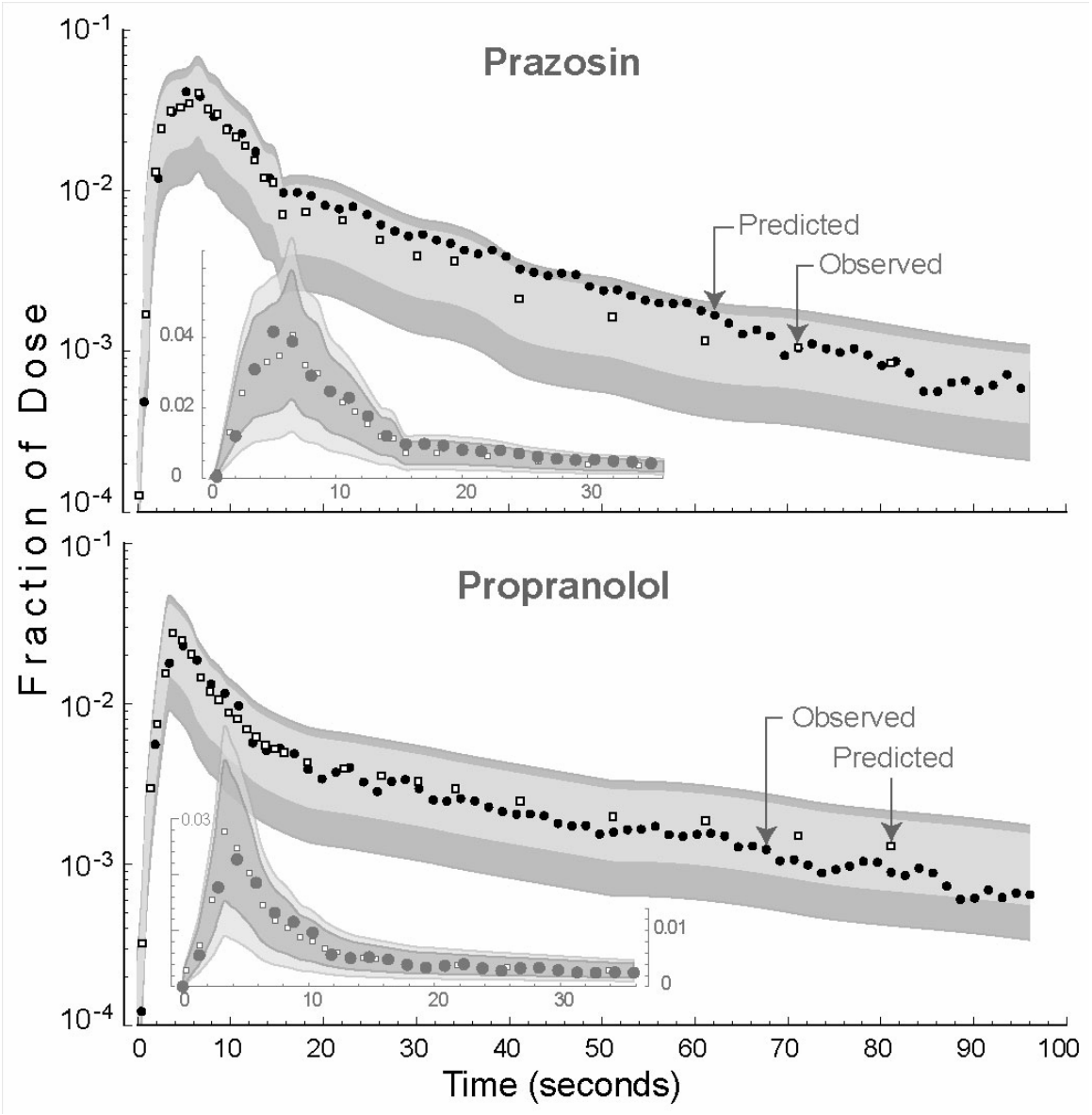

Figure 4.7 Expected outflow profiles (black circles) for prazosin and propranolol using predicted parameter values. These semilog plots show results of simulation experiments for which the PCP-sensitive PVs were predicted using the ANN Method values listed in Table 4.5. Otherwise, the graph components are same as specified in Figure 4.5.

4.4 Discussion

Prediction of PK parameter values plays an increasingly important role in decision making concerning drug candidates in development. The methods and issues, which have been thoroughly reviewed, fall into two categories: empirical, data-based approaches (47-50) and structure-based approaches (37, 48). Data-based approaches can be subdivided into linear and non-linear methods, and each approach may involve clustering. Examples of the former include multiple linear regression, partial least squares analysis, and linear discriminant analysis. Examples of the latter include ANNs, genetic algorithms, support vector machines, and k-nearest neighbor algorithms. Popular clustering techniques include partitional algorithms (including *k*-Means Clustering), nearest neighbor clustering, ANNs, evolutionary approaches, search-based approaches (including simulated annealing), and fuzzy clustering. I employed clustering, specifically fuzzy clustering, as one of the three methods used because it is difficult to classify PCPs and the drug's they characterize into crisp, disjoint clusters. Conventional clustering produce partitions: each pattern fits in one and only one cluster. Fuzzy clustering assigns each pattern within each cluster a membership value. I speculated that this feature may make it an effective tool for clustering PCPs for mapping to ISL PVs, especially when data are sparse.

The relative merits of linear and non-linear methods when dealing with traditional PK parameter estimation are now reasonably well understood. However, the ISL is representative of a new class of models. It contains more mechanistic detail than traditional PK models, and only a sub-set of its parameters is PCP-sensitive. Consequently, I had no a priori basis for selecting one of the preceding methods over

another. I chose three quite different methods to begin building insight into their relative merits for parameter estimation within the ISL context, given the multiple levels of uncertainty.

I started with a simple version of each method. I treated the predictions arising from their use as semi-independent "votes" for how to parameterize an ISL, given a set of PCPs of prazosin and propranolol, along with ISL PVs that are independent of PCPs from having validated ISLs for sucrose, altenolol, antipyrine, labetalol, and diltiazem. Simulations arising from use of the three different sets of predicted ISL parameter values gave us different hepatic disposition and outflow predictions. Because referent data were sparse, I was not trying to select a best prediction method. I took each outflow prediction in Figures 4.5–4.7 as a plausible estimate what one might expect following administration of prazosin and propranolol to a perfused liver. In the case of the ANN (Figure 4.7), I recognized the risks associated with over-parameterization (45). All three methods gave rise to predicted profiles that were within a factor of two of the actual profiles. I found it noteworthy that the simple, single PCP correlation method (Figure 4.5) lead to ISL outflow profiles that were close to those of FCMA and ANN.

A single point in discretized PCP space represents an unique compound. The space is not continuous. Each point can map to a relatively unique hepatic outflow profile, which is a consequence of many entangled local, intralobular mechanisms. A mapping from PCP space to a disposition outflow profile (an attribute of the drug-liver phenotype) is complex. A small shift in location in PCP space is expected to correspond to distinguishable outflow profile changes. Table 4.2 shows that for this work, PCP space has nine dimensions. It is relatively simple to calculate dozens of additional

physicochemical descriptors, ranging from simple geometric to sophisticated electrostatic and thermodynamic. On the other hand, with larger data sets, one can reduce the number of dimensions using a linear dimensionality reduction method such as Principal Component Analysis. However, the goal of this work is not finding the optimal size of PCP space. Rather, it is to explore the usefulness of the new methods in this context.

The ISL is an assembly of componentized mechanisms: purposefully separated and abstracted aspects of hepatic form, space, and organization interacting with COMPOUNDS. Each component mechanism has been unraveled from the complex whole of the hepatic-drug phenotype; it has its own unique phenotype, but that phenotype is much simpler than that of the entire lobule mechanism. Hence, a mapping from component behavior space to PCP space is also simpler and fundamentally different from the PCP-outflow profile mapping described above. For each ISL component, a small location change in PCP space may correspond to a negligible or modest change in the properties of that component mechanism, as well as to the value of the parameter controlling its behavior. I see evidence of this in Tables 4.1 and 4.5: there are cases where different compounds have similar ISL PVs. Note also that for both PRAZOSIN and PROPRANOLOL, there are cases where each of the three methods predicted similar PVs for the same drug. Therefore, considered in retrospect, it is not surprising that even an overly simple estimation method can lead to an acceptable parameterization. Interestingly, the current level of mechanistic granularity is sufficiently fine-grained so that the subtle patterns in PCPs and in the PCP-sensitive, ISL PVs that are needed to achieve larger SM values (Figure 4.7; Table 4.5) are just enough for the simple ANN parameter estimation technique, the more complicated of the three, to find them.

The results suggest that when using the synthetic method of assembling separated aspects of mechanism, a parameter estimation method, which reasonably quantifies the relative differences between compound-specific behaviors at the level of detail represented by those mechanisms, will provide satisficing (the minimum satisfactory outcome), ballpark estimates. That bodes well for using the ISL class of models for predicting PK properties, given only molecular structure information. However, the six drugs in this study are all weak bases. As such, they occupy a common sub-region of PCP space. More drug cases, spanning additional regions, will be needed to build confidence that the above, apparently favorable parameter prediction situation is a consequence of the synthetic method and its instantiation for the ISL (rather than fortuity).

Because it is a synthetic model, the ISL successfully shrinks the PCP-sensitive phenomenal manifold relative to that of tightly coupled mathematical PK models. Traditional equation-based, PK models reduce the phenomenal manifold to a set of real-valued parameters interacting within and assuming a relational continuum. Although one can reduce the number of parameters being considered with these models, in order to maintain accurately quantified relationships, the set of reduced parameters must still accurately capture the aggregated relationships of the larger parameter set. This implies that, for tightly coupled PK models, the phenomenal manifold remains just as complex even though there are fewer parameters in the PK equations. The parsimony achieved by such reduction is, in a sense, illusory. The fine-grained relationships that traditional PK models conflate into mathematical parameters are deeply embedded in the coarse-grained relationships of these parameters. In contrast, the ISL can be actually

simplified because it is an assemblage of coupled, abstract mechanisms, such as transfer between spaces and entry into cells, where the fine-grained relationships can be completely ignored.

Of course, a negative consequence of the ISL's componentized simplification is that the model can sometimes yield dramatically inaccurate and abiotic simulations. Here are three examples: 1) a parameterization makes too much space available to DRUGS causing a prolonged, abiotic lag-time to occur before a slowly rising wave of DRUG emerges. 2) When the diversity in travel paths becomes too limited, travel paths can form distinct clusters, rather than being smoothly distributed. That discontinuity in path options can cause abiotic, oscillations to occur in outflow profiles. Evidence of oscillations can be seen in the PROPRANOLOL profile prediction in Figure 4.6. 3) Inaccurate and abiotic crowding can occur when too many DRUGS converge too quickly on a too small space; this can occur when too many edges connect to one SS. However, *all models are inaccurate*. A central purpose to the synthetic method is to clearly identify and denote both the accurate and inaccurate assumptions and behaviors that characterize a model. Similar phenomena occur with all wet-lab models. Aspects of *in vitro* phenotype are expected to map to corresponding attributes in animal species or patients. However, combinations of experimental specifications can cause *in vitro* behaviors and phenomena for which there is no counterpart in the referent.

As long as I maintain awareness of the above negative consequence, such effects can be minimized. For research that adopts the motivating vision of PBPK, it seems important to use both the synthetic method for chronicling modeling assumptions (heuristic value) and the inductive method to achieve high tolerance predictive value. I

posit that, by having simplified, yet fine-grained, heuristic models, like the ISL, which are commensurate with relatively simple parameter estimation methods, the biological mechanisms become more understandable and accessible. Increased accessibility can provide the methodological leverage needed to build synthetic analogues that will enable us to explore, digest, categorize, understand, and use the massive amounts of data being generated by "omic" technologies.

5 Summary and Perspectives

5.1 Summary

The ISL is a representative of an entirely new class of physiologically based models. By designing components and composing them in a synthetic model capable of exhibiting ten targeted capabilities (discussed below) I have created a vehicle for continual clarification at multiple levels of what occurs within the liver during passage of compounds of interest.

5.1.1 Ten Capabilities Achieved and Demonstrated

1. *Intrahepatic events:* The fixed and dynamic relationships within the ISL (Figures 3.1 and 3.2) are relativistic and realistic at the current level of resolution. For example, for simple task of representing sucrose and antipyrine outflow profiles, components that are more accurate and detailed than those used proved not to be needed. Instead, it was important that component interactions in silico reasonably matched the corresponding referent interactions, such as the fraction of time ANTIPYRINE objects spend in different environments.

2. *Physiological mapping:* By conceptually extending the ISL model and the in silico experimental procedure beyond the liver to include key components associated with the actual wet-lab system, catheters and fraction collection, for example, it proved easier to make the observables from each consistent. It has been widely recognized that such system factors, outside of the biological component, must be factored into hepatic clearance models (51). The dosing function can be easily changed to represent different experimental systems without having to change any LOBULE property. Two classes of observables influenced physiological mappings: histological and dynamic. I used the former to set constraints: acceptable in silico components and relationships needed to fall within to be consistent. Where they fell, such as the number of graph nodes, their interconnections and SS properties, and number of ENZYMES per HEPATOCYTE, were controlled by the dynamic observable, of which I had one type: outflow profiles. By increasing the number of compounds that an ISL can represent (discussed below), I will be iteratively improving the physiological mapping along with the accuracy of the detail within.
3. *Turing test:* Because the many Monte Carlo controlled options at each cycle, each trek of dose through the same lobule is unique and no two lobules are identical. By pooling results from several simulations, I generated a unique outflow profile. In these ways, the ISL has represented uncertainty in observations as well as in the biology and the outflow profiles. Using an acceptably parameterized ISL, I demonstrated that once the SM criteria were satisfied, then from data alone, the results from wet-lab and in silico experiments

were experimentally indistinguishable.

4. *Transparency:* Synthetic models transparency helps bring conceptual clarity to the opacity of biological systems. I can record where SUCROSE and ANTIPYRINE go and what they do. I can then visualize such detail as the simulation progresses (Figures 7.6 & 7.7). I can, for example observe how any SS is connected to others or how many HEPATOCYTES an ANTIPYRINE visits before being METABOLIZED. From a modeling perspective, such capabilities provide a means of identifying potential flaws in the software, its assembly, and operation in ways that can feed back and influence the design of future experiments and how I think about the referent. Having transparency makes it possible to contrast visual and abstract features of the ISL and to observe how processes give rise to characteristic features. I can trace cause-effect relationships to specific subcellular components.
5. *Articulation:* To make model evolution facile, the PBPK modeler needs the ability to easily explore an alternate sinusoidal topography, for example, or to alter the spatial arrangement of TRANSPORTERS and ENZYMES (16, 23). I have demonstrated that it is easily to change number of SS nodes in three different zones, the inter-zone connections, and the intra-zone connections (Figures 3.4, 3.5, 7.1 and 7.4). The same can be done within the SS functional unit: a grid or cells can be removed and modified. Cells can be added to or removed. The functions and interrelations of the other components remain unaltered.
6. *Flexibility:* With traditional PBPK models, changing assumptions as I did for Figure 7.6, can require a different set of differential equations where the

mapping between the new and the original parameters is not straightforward. This reality often hampers exploring the consequences of alternative sets of assumptions. Consequently, some PBPK models can acquire a degree of inertia. Researchers find it easier or more cost effective to adjust their requirements and model use to available models, thereby avoiding dealing with such challenges and forgoing the benefits they might have obtained. Elimination of such situations is the motivation for capability eight.

7. *Reusability:* I have demonstrated a powerful, scientifically useful characteristic of the ISL. An ISL parameterized for one COMPOUND, for example, ANTIPYRINE, can be used to represent the outflow profile of new compounds, such as ATENOLOL, LABETALOL, and DILTIAZEM, without compromising the model's ability to interact as before with the previous compound. I need only change those parameters that are influenced differently by the new compound's physicochemical properties. Should one or two new features be needed, they can be added without compromising existing ISL capabilities. The ISL exists and is capable of functioning whether or not it is dosed with SUCROSE and/or DILTIAZEM. Each COMPOUND has properties that can be recognized by the ISL components. They can even be given recognition algorithms so that after acquiring a mobile object's properties, the ISL components automatically adjust their responses according to a separate or learned algorithm and the objects they encounter.
8. *Discrete interactions:* I demonstrated that adopting graphs (networks) as a fundamental modeling tool allows us to achieve the above nine capabilities. By

restricting the model to discrete interactions, it ensures that graphs can be used throughout the ISL, which ensures the above nine capabilities and makes the resulting device more explorable.

9. *Adaptability:* Having the ISL become a component in a whole organism model is a special case of the *articulation, reconfigurability, and flexibility* capabilities. Attaining it is essential to achieve the full PBPK vision and strengthen the case being made for this class of models. Aspects of an ISL can be separately validated against *in vitro* data, for example, prior to inclusion within the whole organism model.
10. *Prediction:* I have demonstrated a promising application of ISL in the pharmaceutical industry. The ISL can be used to predict new compounds' hepatic disposition based on new compounds' PCPs and the similarity of the new compound's PCPs to those of known compounds.

5.1.2 Detail, Accuracy, and Realism

How much detail does an ISL need; how accurate and realistic does it need to be? The answers require having other information, such as a clear statement of why the model is needed. The PBPK vision is such a statement. I also need to state the uses to which the model will be put, such as to mimic specific aspects of hepatic functionality along with PCs of the referent system that I deem important for the analogue to possess. These PCs were a significant determinant of the detail reflected in Figures 3.1 and 3.2. Depending on how PCs are specified, they will circumscribe a space of allowable ISLs that can range from relatively simple to complex. The referent PCs are expected to have

direct counterparts in the ISLs. For example, if at least 50% of the sinusoids near the PV in the referent lobules are (or are assumed to be) interconnected, and this property is among the listed PCs, then at least 50% of the SSs near the PV in the ISL should be interconnected.

Most important are the measurable observables of the referent. They are used for model validation. They circumscribe the behavior space of each ISL. The similarity measures used to compare referent and ISL observables provide an acceptance criterion for the accuracy; and so, indirectly they are part of the ISL specification. They influence the ISL's detail and fidelity. For example, if I had decided to be satisfied with outflow profiles that fit within a ± 2SD band about the target outflow data, and I had decided to tolerate a few modest bumps in the simulated outflow profile, then I could have accepted lobule graph structures having significantly less detail.

Several ISL features can be cited as being realistic; others can be criticized as being unrealistic. I have dealt with the former. Relative to the latter, here are three examples. 1) In a lobule, only adjacent sinusoids are interconnected, whereas in the ISL, connections between nodes are randomly assigned without consideration of relative location. 2) Because the S_A and S_B are randomly assigned, the range of path lengths in terms of total grid length can be unrealistically broad. For example a one-node path could be as short as two grid spaces in length; whereas, a three-node path could be 50 grid spaces or longer. It seems unlikely that the range in actual lobule path lengths will be that broad. To make relative path lengths more realistic, I need to add a data-driven constraint to our list of targeted PCs. 3) The graph structures (Figures 3.3 and 7.3) are intended to represent plausible *paths*, not actual measurable lengths. They are not

intended to be scale models of actual sinusoidal arrangements.

Because of the random assignment of edges, a Zone 1 or 2 node can be connected to itself by a single edge that forces some compound to revisit the same region of functionality many times. I did not eliminate this feature because, even though it does not seem realistic, I did not have experimental evidence to rule out such path options occurring at a low frequency. Should the evidence become available (or a convincing argument be offered) that no such paths exist in livers, a constraint can be added to disallow them.

The ISL is complicated. When dealing with sucrose and six cationic drugs, I only address a few intracellular events encountered by drugs, such as intrahepatic binding and metabolization. To achieve the PBPK vision will require a combination of sophisticated knowledge, skills, and tools. The problems can have complex origins. Complicated problem solving and decision-making skills require complicated technology (or experienced experts or both). As the overlap between in silico and referent phenotypes increases, the complexity of ISLs and similar devices will approach that of their referents.

5.2 Perspectives

There is increasing impetus for the inclusion of in silico PBPK modeling in defining compound PK properties and their suitability for clinical use (52). To achieve these goals it is becoming necessary to provide increasingly more useful predictions and detailed explanations for observed PK behaviors and properties. Achieving the required detail is limited by the abstractive nature and inherent heuristic limitations of current inductive PBPK models. The synthetic ISL is an example of a class of models

specifically intended to help provide the required details and improved insights, and deliver useful predictions. The results represent an important advance in the science and methods of PBPK modeling and simulation.

The greater the similarity between the properties and characteristics of the ISL and the liver attributes to which they map, the more useful the ISL is expected to become as a PBPK tool. Progresses can be made in achieving the larger PBPK vision by systematically increasing liver and ISL similarity at four levels:

- Outflow or PK profiles at both normal and diseased conditions
- Measures of related phenotypic attributes
- Generative components
- Component interactions and their organization

Internal and external factors can lead to important differences between livers. The PBPK vision anticipates being able to explore the significance and consequences of such differences. The ISL components can be easily reconfigured to represent different histological, physiological, or experimental conditions. Using directed graphs to represent connectivity between units of function, one of several possible approaches, enables us to manipulate the representation of both histology and physiology. I have shown that it is easy to alter these representations and explore the consequences. It is similarly easy to alter the experimental context as well. The ISL represents an isolated, perfused liver. By altering the lobule input functions, I can change the system to represent the liver in a living rat or within a different perfusion system. I can unplug

HEPATOCYTES from the ISL and study their properties in simulated *in vitro* experimental conditions (31, 53). Given appropriate *in vitro* data I can also refine, reparameterize and validate the HEPATOCYTES, and return them to the ISL to observe the consequences within the whole-liver context.

In addition to predict drug's hepatic disposition in normal livers, another direction to expand ISL is to use it to represent diseased livers. Liver diseases alter hepatic microcirculation (54), impair the disposition kinetics, and decrease the metabolism of many drugs (55, 56). Fenyves *et al*. (57) hypothesized that the disposition of propranolol in cirrhosis is partially dependent on capillarization and intrahepatic shunts, oxygen delivery, and possibly acinar heterogeneity. When studying drug pharmacokinetics in carbon tetrachloride-induced cirrhosis rat and alcohol-induced chronic hepatocellular injury rat, Yung *et al*. (20) found that the disposition of cationic drugs may be directly related to the extent of fibrosis caused by the liver disease and the physicochemical properties of the drugs. Alternations in sinusoidal morphology, hepatic microcirculation, and decreased enzymatic activity in the liver caused reduction in PS, CL_{int}, and perfusate flow rate. In addition, hepatic distribution is reduced in cirrhosis due to reduced tissue binding as a result of a low synthesis of intrahepatocellular proteins, loss of ion-trapping effect, and change in organelle volumes. Induced pharmacokinetic changes due to liver disease depend on drug properties as well as on the severity of liver disease (58). As mentioned above, one of the advantages of ISL is that it can be easily reconfigured to represent different histological and physiological conditions. Therefore I can change SS network connections to represent alternations in sinusoidal morphology and hepatic microcirculation. Girds that represents CELLS and HEPATOCYTES and objects

representing intracellular components can be made heterogeneous to represent inter and intracellular heterogeneity in diseased condition. Based on in situ (referent) cationic drug hepatic disposition data obtained from diseased liver (20), I am able to find the corresponding ISL parameterizations to represent injured liver and use this expanded ISL as platform to study the causality mechanism between altered hepatic disposition with drug's properties and changed liver environments, and predict hepatic dispositions in various injured conditions. The resulted finding could help us better understand and estimate how changes in hepatic diseases qualitatively and quantitatively affect hepatic drug behaviors, which is very helpful in facilitating drug preclinical development in pharmaceutical industry.

6 References

1. Hung DY, Chang P, Weiss M, Roberts MS. Structure-hepatic disposition relationships for cationic drugs in isolated perfused rat livers: transmembrane exchange and cytoplasmic binding process. J Pharmacol Exp Ther 2001;297(2):780-9.
2. Rowland M. Physiologic pharmacokinetic models: relevance, experience, and future trends. Drug Metab Rev 1984;15(1-2):55-74.
3. Rowland M, Balant L, Peck C. Physiologically based pharmacokinetics in drug development and regulatory science: a workshop report (Georgetown University, Washington, DC, May 29-30, 2002). AAPS PharmSci 2004;6(1):E6.
4. Andersen ME. Toxicokinetic modeling and its applications in chemical risk assessment. Toxicol Lett 2003;138(1-2):9-27.
5. Leahy DE. Progress in simulation modelling for pharmacokinetics. Curr Top Med Chem 2003;3(11):1257-68.
6. Corley RA, Mast TJ, Carney EW, Rogers JM, Daston GP. Evaluation of physiologically based models of pregnancy and lactation for their application in children's health risk assessments. Crit Rev Toxicol 2003;33(2):137-211.
7. Roberts MS, Anissimov YG. Modeling of hepatic elimination and organ distribution kinetics with the extended convection-dispersion model. J Pharmacokinet Biopharm 1999;27(4):343-82.
8. Roberts MS, Magnusson BM, Burczynski FJ, Weiss M. Enterohepatic circulation: physiological, pharmacokinetic and clinical implications. Clin Pharmacokinet 2002;41(10):751-90.
9. Czarnecki K, Eisenecker U. In: Generative Programming: Methods, Tools, and Applications. New York: Addison-Wesley; 2000. p. 10, 251-254.
10. Ropella GE, Hunt CA, Nag DA, editors. Using heuristic models to bridge the gap between analytic and experimental models in biology. San Diego, CA.: SCS Press;

2005.

11. Ropella GE, Hunt CA, Sheikh-Bahaei S. Methodological Considerations of Heuristic Modeling of Biological Systems. In: Proc 9th World Multi-Conf Systemics, Cybernetics and Informatics. Orlando, Florida; 2005.
12. Steels L, Brooks R. In: The Artificial Life Route to Artificial Intelligence. New Jersey: Lawrence Earlbaum Associates, Inc.; 1995. p. 83-121.
13. Zeigler BP, Praehofer H, Kim TG. In: Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems. California: Academic Press; 2000. p. 3-36, 75-85, 99-104, 137-147.
14. Andersen ME, Thomas RS, Gaido KW, Conolly RB. Dose-response modeling in reproductive toxicology in the systems biology era. Reprod Toxicol 2005;19(3):327-37.
15. Rosen R. Role of similarity principles in data extrapolation. Am J Physiol 1983;244(5):R591-9.
16. Liu Y, Hunt CA. Mechanistic study of the cellular interplay of transport and metabolism using the synthetic modeling method. Pharm Res 2006;23(3):493-505.
17. Thompson M, Ellison SLR, Wood R. Harmonized guidelines for single-laboratory validation of methods of analysis. Pure Appl. Chem. 2002;74(5):835-855.
18. Hunt CA, Ropella GE, Yan L, Hung DY, Roberts MS. Physiologically based synthetic models of hepatic disposition. J Pharmacokinet Pharmacodyn 2006;33(6):737-72.
19. Anissimov YG, Roberts MS. A compartmental model of hepatic disposition kinetics: 1. Model development and application to linear kinetics. J Pharmacokinet Pharmacodyn 2002;29(2):131-56.
20. Hung DY, Chang P, Cheung K, McWhinney B, Masci PP, Weiss M, et al. Cationic drug pharmacokinetics in diseased livers determined by fibrosis index, hepatic protein content, microsomal activity, and nature of drug. J Pharmacol Exp Ther 2002;301(3):1079-87.
21. Liu L, Pang KS. An integrated approach to model hepatic drug clearance. Eur J Pharm Sci 2006;29(3-4):215-30.
22. Siebert GA, Hung DY, Chang P, Roberts MS. Ion-trapping, microsomal binding, and unbound drug distribution in the hepatic retention of basic drugs. J Pharmacol Exp Ther 2004;308(1):228-35.
23. Liu Y, Hunt CA. Studies of intestinal drug transport using an in silico epithelio-mimetic device. Biosystems 2005;82(2):154-67.
24. Sheikh-Bahaei S, Hunt CA, editors. Prediction of In Vitro Hepatic Biliary Excretion Using Sto-chastic Agent-Based Modeling and Fuzzy Clustering. Monterey, CA; 2006.
25. Steen H, Oosting R, Meijer DK. Mechanisms for the uptake of cationic drugs by

the liver: a study with tributylmethylammonium (TBuMA). J Pharmacol Exp Ther 1991;258(2):537-43.

26. Cheung K, Hickman PE, Potter JM, Walker NI, Jericho M, Haslam R, et al. An optimized model for rat liver perfusion studies. J Surg Res 1996;66(1):81-9.

27. Teutsch HF, Schuerfeld D, Groezinger E. Three-dimensional reconstruction of parenchymal units in the liver of the rat. Hepatology 1999;29(2):494-505.

28. Miller DL, Zanolli CS, Gumucio JJ. Quantitative morphology of the sinusoids of the hepatic acinus. Quantimet analysis of rat liver. Gastroenterology 1979;76(5 Pt 1):965-9.

29. Gumucio JJ, Miller DL. Zonal hepatic function: solute-hepatocyte interactions within the liver acinus. Prog Liver Dis 1982;7:17-30.

30. Garmire LX, Garmire DG, Hunt CA. An in silico transwell device for the study of drug transport and drug-drug interactions. Pharmaceutical Research 2007:Published online.

31. Sheikh-Bahaei S, Ropella GEP, Hunt CA, editors. In Silico Hepatocyte: Agent-based Modeling of the Biliary Excretion of Drugs In Vitro. San Diego, CA: SCS Press; 2006.

32. Sheikh-Bahaei S, Ropella GEP, Hunt CA. In Silico Hepatocyte: Agent-Based Modeling of the Biliary Excretion of Drugs. In: 2006 Spring Simulation Multiconference, The Society for Modeling and Simulation International; 2006 April 2-6; Huntsville, AL; 2006.

33. Yan L, Ropella GEP, Park S, Roberts MS, Hunt CA. Modeling and Simulation of Hepatic Drug Disposition Using a Physiologically Based, Multi-Agent In Silico Liver. Pharm Res 2007:Published online: 28 November 2007 (DOI: 10.1007/s11095-007-9494-y).

34. Ljung L. System Identification: Theory for the User (2nd Edition). 2nd Edition ed: Prentice Hall PTR; 1998.

35. Mager DE. Quantitative structure-pharmacokinetic/pharmacodynamic relationships. Adv Drug Deliv Rev 2006;58(12-13):1326-56.

36. Hou T, Wang J, Zhang W, Wang W, Xu X. Recent advances in computational prediction of drug absorption and permeability in drug discovery. Curr Med Chem 2006;13(22):2653-67.

37. Ng C, Xiao Y, Putnam W, Lum B, Tropsha A. Quantitative structure-pharmacokinetic parameters relationships (QSPKR) analysis of antimicrobial agents in humans using simulated annealing k-nearest-neighbor and partial least-square analysis methods. J Pharm Sci 2004;93(10):2535-44.

38. Nestorov I, Aarons L, Rowland M. Quantitative structure-pharmacokinetics relationships: II. A mechanistically based model to evaluate the relationship between tissue distribution parameters and compound lipophilicity. J Pharmacokinet Biopharm 1998;26(5):521-45.

39. Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. Computers and Geoscience 1984;10:191-203.

40. Pal N, Bezdek J. On cluster validity for the fuzzy c-means model. IEEE Trans Fuzzy Syst 1995;3(3):370-379.

41. Ekins S, Waller CL, Swaan PW, Cruciani G, Wrighton SA, Wikel JH. Progress in predicting human ADME parameters in silico. J Pharmacol Toxicol Methods 2000;44(1):251-72.

42. van de Waterbeemd H, Gifford E. ADMET in silico modelling: towards prediction paradise? Nat Rev Drug Discov 2003;2(3):192-204.

43. Hansch C, Leo A, Mekapati SB, Kurup A. QSAR and ADME. Bioorg Med Chem 2004;12(12):3391-400.

44. Sheikh-Bahaei S, Hunt CA. Parameter Estimation via Analysis of Fuzzy Clusters (PEAF): An Algorithm to Estimate Parameters of Agent-Based Models. In: Proceedings of Computer Applications in Industry and Engineering (CAINE-2007), 20th Int'l. Conference; 2007 November 7-9, 2007; San Francisco, CA, USA; 2007.

45. Hudson DL, Cohen ME. Neural Networks and Artificial Intelligence for Biomedical Engineering (IEEE Press Series on Biomedical Engineering): Wiley-IEEE Press; 1999.

46. Ertl P, Rohde B, Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. J Med Chem 2000;43(20):3714-7.

47. Beresford AP, Segall M, Tarbit MH. In silico prediction of ADME properties: are we making progress? Curr Opin Drug Discov Devel 2004;7(1):36-42.

48. Li H, Yap CW, Ung CY, Xue Y, Li ZR, Han LY, et al. Machine learning approaches for predicting compounds that interact with therapeutic and ADMET related proteins. J Pharm Sci 2007;96(11):2838-2860.

49. Stouch TR, Kenyon JR, Johnson SR, Chen XQ, Doweyko A, Li Y. In silico ADME/Tox: why models fail. J Comput Aided Mol Des 2003;17(2-4):83-92.

50. Yamashita F, Hashida M. In silico approaches for predicting ADME properties of drugs. Drug Metab Pharmacokinet 2004;19(5):327-38.

51. Anissimov YG, Bracken AG, Roberts MS. Catheter effects in organ perfusion experiments. J Theor Biol 2002;214(2):263-73.

52. Food and Drug Administration. Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products. 2004 [cited; Available from: http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html

53. Sheikh-Bahaei S, Ropella GEP, Hunt CA. Agent-based simulation of in vitro hepatic drug metabolism: in silico hepatic intrinsic clearance. In: Spring Simulation Multiconference, The Society for Modeling and Simulation International. San Diego, CA; 2005.

54. Varin F, Huet PM. Hepatic microcirculation in the perfused cirrhotic rat liver. J Clin Invest 1985;76(5):1904-12.

55. Gariepy L, Fenyves D, Kassissia I, Villeneuve JP. Clearance by the liver in cirrhosis. II. Characterization of propranolol uptake with the multiple-indicator dilution technique. Hepatology 1993;18(4):823-31.

56. Reichen J, Arts B, Schafroth U, Zimmermann A, Zeltner B, Zysset T. Aminopyrine N-demethylation by rats with liver cirrhosis. Evidence for the intact cell hypothesis. A morphometric-functional study. Gastroenterology 1987;93(4):719-26.

57. Fenyves D, Gariepy L, Villeneuve JP. Clearance by the liver in cirrhosis. I. Relationship between propranolol metabolism in vitro and its extraction by the perfused liver in the rat. Hepatology 1993;17(2):301-6.

58. Rodighiero V. Effects of liver disease on pharmacokinetics. An update. Clin Pharmacokinet 1999;37(5):399-431.

7 Appendix

7.1 Description of ISL Parameters and Their Values

Table 7.1 Description of ISL parameters and their values for the simulations in text Figures 3.4, 3.5, 3.6, 4.3 and 4.4, and Appendix figures 7.1, 7.2, 7.4 and 7.5, except where noted otherwise

Parameter[a]	Description	Values[b]	Ranges[c]
Device Framework Parameters			
cycleLimit	(B) [d] Provides the simulation with an artificial stopping criterion so that it will not run forever.	200	200
monteCarloRuns	(I) [d] The number of runs in a Monte-Carlo set. Aggregate measures for the whole system are calculated at the end of all Monte-Carlo runs.	48	[2, 50]
similarityMeasure	Specifies which similarity measure to use.	global_sd	global_sd
nominalProfile	(S) [d] Specifies which model to use as the nominal for use in calculating similarity between the experimental and nominal models.		
experimentalProfile	(S) Specifies which model to use as the experimental for use in calculating similarity between the experimental and nominal models.		
In Silico Liver Parameters			
StepsPerCycle	(I) Parameterizes the granularity of time for the Research model (the ISL). The data has a "sampling rate" that provides a default time scale for it. Finer granularity requires interpolation between the data points. The Reference (PK) model [a] is a time-reversible, continuous function, which allows sampling at any frequency.	2	2
GraphInputFile	(S) Specifies the file to read if the SS graph is to be specified by an explicit graph (file format is GML)	null	
GraphSpecFile	(S) Provides the lobule graphical specification, and specifies the base file name (extension .ls) to be used if the graph is to be specified according to the Lobule Specification		

	Nodes in Zone I	45	[20, 50]
	Nodes in Zone II	21	[10, 35]
	Nodes in Zone III	6	[3, 15]
	Intra-Zone I edges	18	[10, 30]
	Intra-Zone II edges	10	[5, 25]
	Intra-Zone III edges	0	0
	Zone I → Zone II edges	15	[2, 45]
	Zone I → Zone III edges	0	[0, 20]
	Zone II → Zone III edges	8	[5, 25]
GraphSpecIterates	(I) Tells the framework to modify LOBULE specification and run a Monte-Carlo set (consisting of *N* runs) for each different LOBULE specification. Set to 1, it runs 1 set and provides 1 set of outputs. Set to 5, the first run uses the current contents of lobule.ls; it then runs 4 more sets, slightly modifying the LOBULE specification each time, resulting in 5 sets.	1	[1, 10]
DirSinRatio	(F) [d] Specifies the percentage of SS that are type $\boldsymbol{S}_A$ ("direct")	0.75	[0, 1]
TortSinRatio	(F) Specifies the percentage of SS that are $\boldsymbol{S}_B$ ("tortuous")	0.25	[0, 1]
DirSinCircMin	(I) Sets the upper and lower bounds for a pseudo-RNG (random number generator) using the uniform distribution to choose a circumference for each SS	50	[1, 50]
DirSinCircMax		50	[1, 100]
TortSinCircMin		4	[1, 8]
TortSinCircMax		4	[1, 20]
DirSinLenAlpha	(F) Sets the parameters for a pseudo-RNG using a modified form of the Gamma distribution; The modification consists of a left-right shift of the distribution, allowing the user to clip off the front.	3.0	[0.5, 5.0]
DirSinLenBeta		0.215	[0.1, 0.325]
DirSinLenShift		0.0	0.0
TortSinLenAlpha		10.0	[4, 18]
TortSinLenBeta		0.125	[0.03, 0.15]
TortSinLenShift		-40	[40.0, 40.0]
SinusoidTurbo	(F) The complement of the amount of turbulence allowed in any given SS. Lower turbo means more tendency of any one COMPOUND to wander sideways or backwards. Higher Turbo means a stronger flow from the input to the output of the SS.	0.3	[0.05, 0.95]
ECDensity	(F) Specifies the relative ENDOTHELIAL CELL density, given that the number of grid points in Grid B is set by the geometry parameters of the SS; it specifies the percentage of Grid B points that index an endothelial cell.	0.8	[0.05, 1.0]
HepDensity	(F) Specifies the relative HEPATOCYTES density within Grid C, given that the number of grid points in Grid C is set by the geometry parameters of the SS, it specifies the percentage of Grid C points that index a hepatocyte.	0.6	[0.05, 1]
CoreFlowRate	(I) The number of slots solute in the SS core moves forward during each step	2	[1, 8]
A2BJumpProb	(F) Specifies the probability that, when given the option, SOLUTE will jump from the sinusoid rim space to the endothelial space	0.1	[0.05, 1.0]
B2AJumpProb	(F) Specifies the probability that, when given the option, SOLUTE will jump from the endothelial	0.6	[0.05, 1.0]

	space to the sinusoid rim space		
B2CJumpProb	(F) Specifies the probability that, when given the option, SOLUTE will jump from the endothelial space to the space of dissé.	0.35	[0.05, 1.0]
C2BJumpProb	(F) Specifies the probability that, when given the option, SOLUTE will jump from the space of dissé to the endothelial space	0.65	[0.05, 1.0]
BindersPerCellMin	(I) Max and min for a uniform pseudo-random draw for the number of binding agents inside each CELL. Simple binders for ECs and ENZYMES for HEPATOCYTES.	5	[1, 30]
BindersPerCellMax		10	[10, 100]
MetabolizeProb	(F) Probability that an ENZYME will metabolize a SOLUTE before it releases it.	0.4	[0, 1]
SoluteBindingProb	(F) Probability that, when a BINDER and SOLUTE make contact, the SOLUTE will be bound.	0.5	[0, 1]
SoluteBindingCycles	(I) Number of cycles a binder will bind a SOLUTE.	25	[5, 100]
ISL2WetLabScaling	(F) Specifies the number of molecules being represented by one SOLUTE. Because everything in the model is normalized, this amounts to a magnitude factor on the output fraction.	7.0	[1.0, 8.0]
BolusStartTime	(I) Dictates when to let SOLUTES flow into the LOBULE	5	[2, 9]
DosageParamA[e]	(F) Parameter (A) of the dosage function: $D(t) = A(B^C t^{(C-1)} e^{-Bt})/(C-1)!$, where t = current cycle Parameter (A) simply raises the amplitude of the function	2000	[1000, 7000]
DosageParamB	(F) Parameter (B) of the dosage function	1	[1, 2]
DosageParamC	(F) Parameter (C) of the dosage function	2	[2, 3]
	Actual dose for Figure 3.4, 3.6, 4.4, and 4.5	3,682	

[a] There are four classes of parameters: Device Framework, Research Model, Reference Model, and Data Model. Here, the Research Model is the ISL. Only a subset of the Device Framework parameters is listed here. The Data Model includes all the data against which the SM is being applied, and uses a parameter that specifies whether to interpolate between observations of the in silico data. The Reference Model is a traditional PK model previously fit to the in silico experimental data; it is run concurrently with the Research Model.

[b] Parameter values when SUCROSE and ANTIPYRINE were dosed in combination.

[c] Ranges from which values were drawn during searches of model and parameter space.

[d] B: Boolean; F: floating; I: integer; S: string

[d] The total dose is the area under the dosage function curve.

7.2 Additional Outflow Profiles for Chapter 3

7.2.1 Outflow Profiles for Four Cationic Drugs plus Sucrose Using Four Different Graph Structure

During the parameter refinement process, from among those parameterizations tested I chose a best LOBULE parameterization for each of four drugs. Keeping that LOBULE parameterization constant, I adjusted the remaining parameter values to obtain best SM value for that drug. Those results are presented in Figures 7.1 and 7.2, and each drug's corresponding lobule graph structure parameter and similarity values are list in Table 7.2. Notice that for Figure 3.4 and Figures 7.1 and 7.2 I reused for the other three drugs the lobule graph structure that worked best for ANTIPYRINE (Figure 3.3), whereas for Figures 7.4 and 7.5 below, I reused for the other three drugs the lobule graph structure that worked best for LABETALOL (Figure 7.3).

Table 7.2 Parameter values for four different lobule graph structures

Lobule Graph Structure Parameter	Values			
	Atenolol	Antipyrine	Labetalol	Diltiazem
Node in Zone I	45	45	30	50
Node in Zone 2	25	21	15	35
Node in Zone 3	9	6	6	5
Intra-Zone I edges	22	18	15	18
Intra-Zone II edges	10	10	7	10
Intra-Zone III edges	0	0	0	0
Zone I → Zone II edges	21	15	10	21
Zone I → Zone III edges	0	0	0	0
Zone II → Zone III edges	1	8	6	12
SM Values	0.88	0.81	0.91	0.99

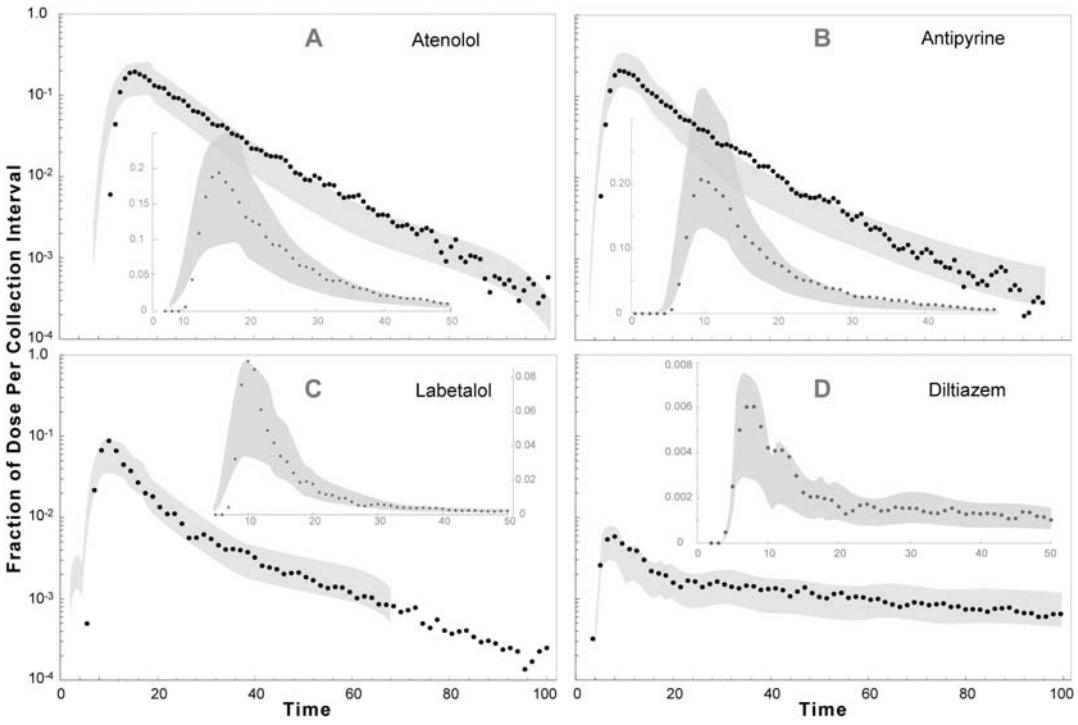

Figure 7.1 Semilog and scatter plots of acceptable outflow profile matches for four drugs using the lobule graph structures in Table 7.2. (A) ATENOLOL, (B) ANTIPYRINE, (C) LABETALOL, and (D) DILTIAZEM were co-administrated with SUCROSE. For all four DRUGS, the SS structure parameter values along with those influenced by drug PCPs are same as in Table 3.2, except for ATENOLOL, which has: *C2BjumpProb* = 0.7 and *MetabolizeProb* = 0.4. The width of the gray SM band assumes the same coefficient of variation as in Figure 3.4. Each datum is a smoothed mean value of 48 Monte carlo simulations of the same ISL.

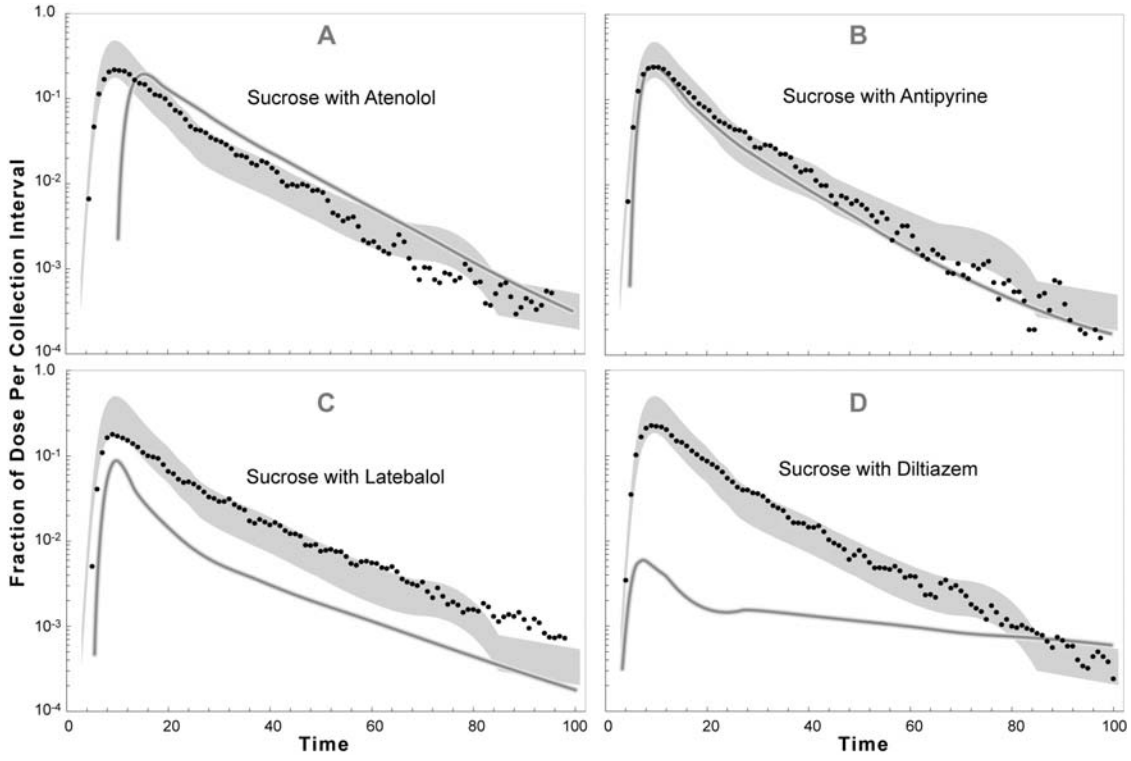

Figure 7.2 Semilog plots of outflow profiles for SUCROSE co-administrated with (A) ATENOLOL, (B) ANTIPYRINE, (C) LABETALOL, and (D) DILTIAZEM in Figure 7.1. The width of the gray SM band assumes the same coefficient of variation as in Figure 3.5. The gray curve is the trend line for the co-administrated DRUG (from Figure 7.1).

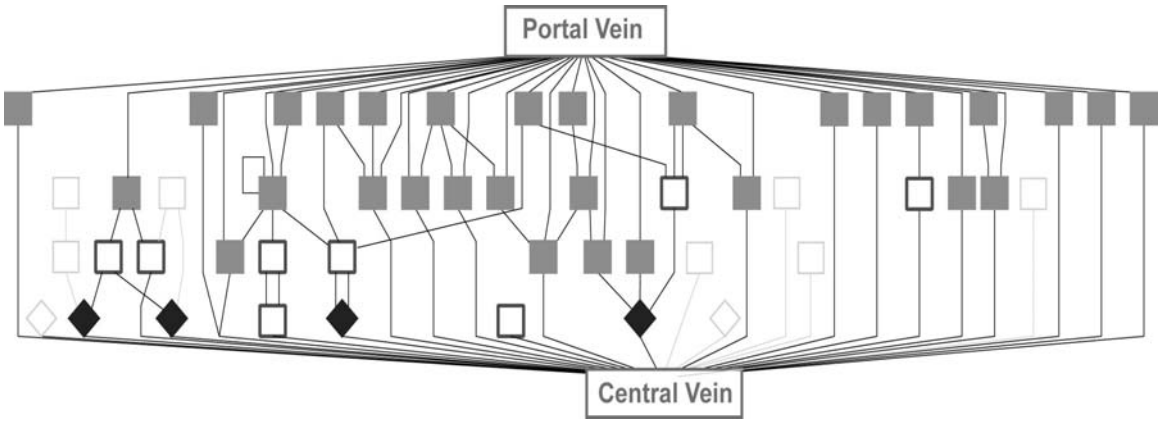

Figure 7.3 An example of the 48 Monte Carlo generated ISL lobule graph structure used in Figure 7.1C and in Figure 7.4 A–D. As in Figure 3.3, all Zone 1 nodes (gray squares) have one incoming edge from the PV. They have one outgoing edge; it can be connected to another SS or to the CV. Zone 2 nodes (unfilled squares) have an edge incoming from a Zone 1 SS; the outgoing edge can be connected to a SS in Zones 2 or 3. Zone 3 nodes (black diamonds) have one incoming edge from a Zone 2 SS; the outgoing edge is connected to the CV. The nine, light gray, unfilled squares and diamonds represent Zone 2 and 3 nodes that, by chance, were not assigned an incoming edge.

7.2.2 Outflow Profiles for Four Cationic Drugs Plus Sucrose Using Different Single Lobule Graph Structure

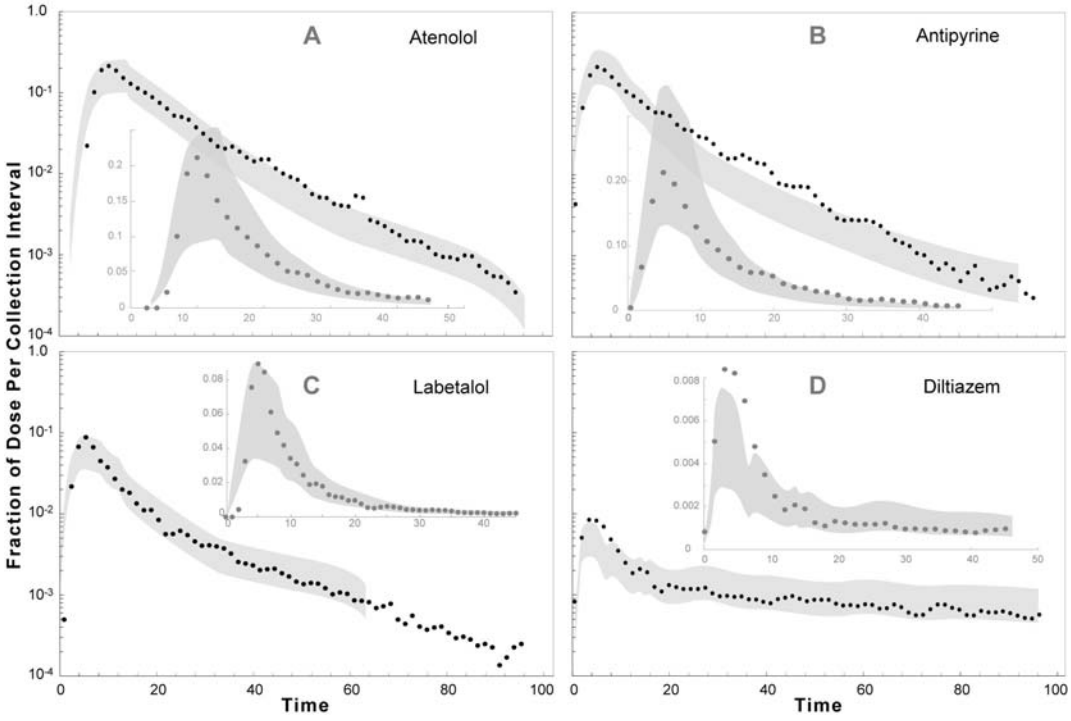

Figure 7.4 Semilog and scatter plots of acceptable outflow profile matches for four drugs using the lobule graph structure exemplified in Figure 7.3. (A) ATENOLOL, (B) ANTIPYRINE, (C) LABETALOL, and (D) DILTIAZEM were co-administrated with SUCROSE. Other experimental conditions were same as in Figure 3.4.

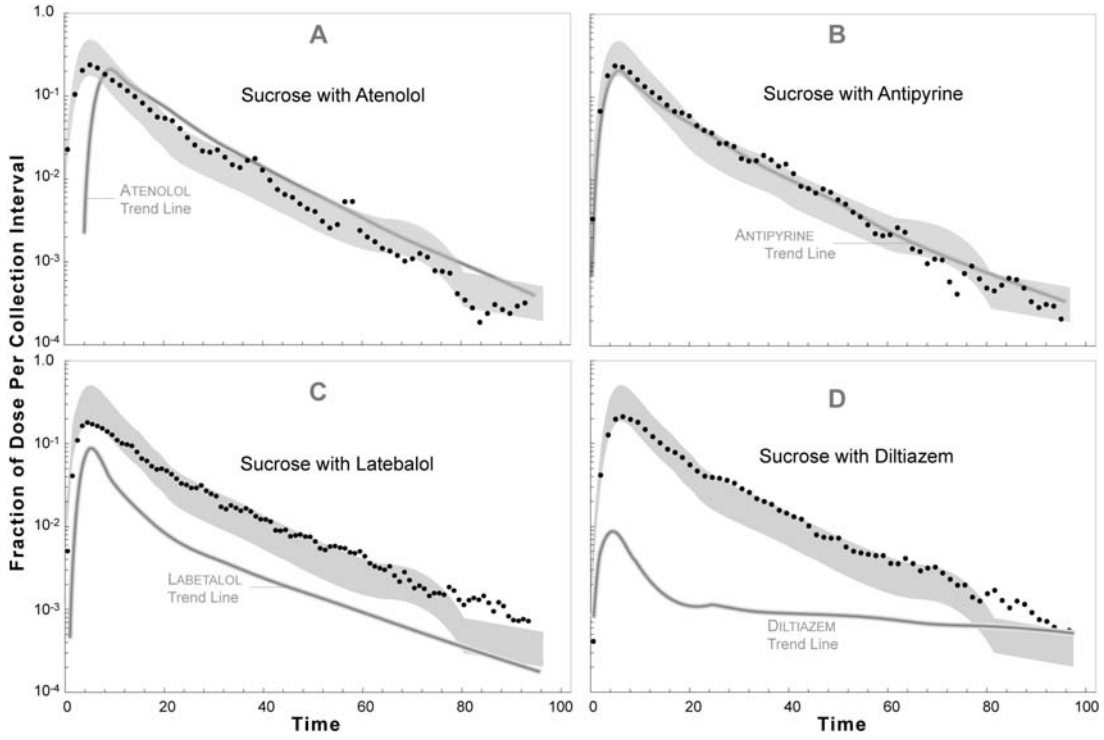

Figure 7.5 Semilog plots of outflow profile data for SUCROSE co-administrated with (A) ATENOLOL, (B) ANTIPYRINE, (C) LABETALOL, and (D) DILTIAZEM in Figure 7.4. The width of the gray SM band assumes the same coefficient of variation as in Figure 3.5. The gray curve is the trend line for the co-administrated DRUG (from Figure 7.4).

7.3 Predicting Sample ISL Parameters Using FCMA

Application of the FCMA was automated. Nevertheless, below we walk through examples of predicting two ISL parameter values for prazosin and propranolol.

7.3.1 Predict a Value of the Parameter *B2CJumpProb* for Prazosin and Propranolol

7.3.1.1 For prazosin

Step 0. Selection of PCPs and creating the data set.

The data set contains n=4 compounds: S={Atenolol, Antipyrine, Labetalol, Diltiazem} for which the simulation parameters are known:

ISL Parameters	Atenolol	Antipyrine	Labetalol	Diltiazem
B2CJumpProb	0.3	0.35	0.5	0.5

B2CJumpProb is one of the four parameters in the ISL subgroup B, moving between spaces (Table 4.1). All four are calculated in the same way. Figure 4.4 specifies that the three PCP properties in subgroups II (partitioning) and IV (topology related) are expected to be equally influential in determining the value of all ISL subgroup B parameters.

PCP Subgroups II & IV	Atenolol	Antipyrine	Labetalol	Diltiazem		Prazosin
RB count	8	1	8	7		4
TPSA	84.58	26.94	95.58	59.09		106.96
logP_{app}	0.14	0.33	2.69	3.53		1.88

Step 1. Let $q = 4$, and S_{new}= {Atenolol, Antipyrine, Labetalol, Diltiazem , Prazosin }.

Step 2. If q = 1 go to step 4. Else, classify S_{new} into q clusters using Fuzzy c-Means algorithm.

Step 3. If Prazosin is not in the same group with *at least* another member then decrease q to q-*1*. Repeat steps 2 and 3.

Else, let *G-value* be the number of groupmates of Prazosin, Go to step 4.

Step 4. The 5 compounds are clustered into 3 groups with the following membership degrees (The G-value is 2 since Prazosin has 2 other groupmates):

Cluster	Atenolol	Antipyrine	Labetalol	Diltiazem	Prazosin
C3	**0.987**	0.001	0.050	0.060	0.316
C2	0.006	**0.998**	0.020	0.045	0.179
C1	0.007	0.001	**0.930**	**0.895**	**0.506**

PRAZOSIN'S simulation parameter, *B2CJumpProb***,** is estimated as:

$B2CJumpProb_{\text{Prazosin}} = 0.506 \bullet P_{C1} + 0.179 \bullet P_{C2} + 0.316 \bullet P_{C3}$

where P_{Ci} is the average ISL parameter values of members of cluster i:

$P_{C1} = (B2CjumpProb_{\text{labetalol}} + B2CjumpProb_{\text{diltiazem}})/2 = 0.5$

$P_{C2} = B2CjumpProb_{\text{antipyrine}} = 0.35$

$P_{C3} = B2CjumpProb_{\text{atenolol}} = 0.3$

Consequently,

$B2CjumpProb_{\text{prazosin}} = 0.506 \times 0.5 + 0.179 \times 0.35 + 0.316 \times 0.3 = 0.4104$ (rounded to 0.41).

7.3.1.2 For Propranolol

PCP Subgroups II & IV	Atenolol	Antipyrine	Labetalol	Diltiazem		Propranolol
RB count	8	1	8	7		6
TPSA	84.58	26.94	95.58	59.09		41.49
logP_{app}	0.14	0.33	2.69	3.53		3.10

Steps 1-3. Let $q = 4$, and S_{new}= {atenolol, antipyrine, labetalol, diltiazem, propranolol}.

For q = 4 and the values above, propranolol is in the same group with *at least* one other member, and G = 1.

Step 4. The classification results are:

Cluster	Atenolol	Antipyrine	Labetalol	Diltiazem	Propranolol
C4	**1**	0	0	0.014	0.014
C3	0	**1**	0	0.010	0.014
C2	0	0	**1**	0.067	0.036
C1	0	0	0	**0.908**	**0.936**

$P_{C1} = B2CjumpProb_{diltiazem} = 0.5$

$P_{C2} = B2CjumpProb_{labetalol} = 0.5$

$P_{C3} = B2CjumpProb_{antipyrine} = 0.35$

$P_{C4} = B2CjumpProb_{atenolol} = 0.3$

Consequently,

$B2CJumpProb_{propranolol} = 0.936\times0.5 + 0.036\times0.535 + 0.014\times0.35 + 0.014\times0.3 =$ 0.4951 (rounded to 0.50)

7.3.2 Predict Parameter *SoluteBindingProb* for Prazosin and Propranolol

7.3.2.1 For Prazosin

Step 0. Selection of PCPs and creating the data set.

The data set contains n=4 compounds: S={Atenolol, Antipyrine, Labetalol, Diltiazem} for which the simulation parameters are known:

ISL Parameters	Atenolol	Antipyrine	Labetalol	Diltiazem
SoluteBindingProb	0.3	0.5	0.6	0.35

SoluteBindingProb is one of the four parameters in the ISL subgroup C, binding to LOBULAR components (Table 4.1). All four are calculated in the same way. Figure 4.4 specifies that fuB (unbound fraction) and the three PCP properties in subgroups V (hydrogen bound related) are expected to be equally influential in determining the value of all ISL subgroup C parameters.

PCP Subgroups III & V	Atenolol	Antipyrine	Labetalol	Diltiazem		Prazosin
fuB	0.74	0.6	0.52	0.28		0.54
HBD count	3	0	4	0		1
HBA count	4	2	4	5		8
Tautomer count	2	0	7	0		3

Steps 1-3. Let $q = 4$, and S_{new}= {atenolol, antipyrine, labetalol, diltiazem, prazosin}.

For q = 4 and the values above, prazosin is in the same group with *at least* one other member, and G = 1.

Step 4. The classification results are:

Cluster	Atenolol	Antipyrine	Labetalol	Diltiazem	Prazosin
C4	**0.980**	0.001	0.001	0.010	0.320
C3	0.007	**0.994**	0.000	0.037	0.191
C2	0.007	0.000	**0.999**	0.003	0.118
C1	0.006	0.004	0.000	**0.950**	**0.371**

P_{C1} = SoluteBindingProb $_{\text{Diltiazem}}$ = 0.35

P_{C2} = SoluteBindingProb $_{\text{Labetalol}}$ = 0.6

P_{C3} = SoluteBindingProb $_{\text{Antipyrine}}$ = 0.5

P_{C4} = SoluteBindingProb $_{\text{Atenolol}}$ = 0.35

Consequently,

$SoluteBindingProb_{Prazosin} = 0.371\times0.35 + 0.118\times0.6 + 0.191\times0.5 + 0.320\times0.35 =$ 0.4082 (rounded to 0.41)

7.3.2.2 For Propranolol:

PCP Subgroups III & V	Atenolol	Antipyrine	Labetalol	Diltiazem		Propranolol
fuB	0.74	0.6	0.52	0.28		2
HBD count	3	0	4	0		3
HBA count	4	2	4	5		0
Tautomer count	2	0	7	0		0.45

Step 1-3. Let $q = 4$, and S_{new}= {atenolol, antipyrine, labetalol, diltiazem, propranolol}.

For q = 4 and the values above, propranolol is in the same group with *at least* one other member, and G = 1.

Step 4. The classification results are:

Cluster	Atenolol	Antipyrine	Labetalol	Diltiazem	Propranolol
C4	**0.989**	0.006	0.000	0.002	0.325
C1	0.004	**0.977**	0.000	0.004	**0.376**
C2	0.003	0.002	**1**	0.000	0.045
C3	0.003	0.014	0.000	**0.993**	0.254

P_{C1} = $SoluteBindingProb_{Antipyrine}$ = 0.5

P_{C2} = $SoluteBindingProb_{Labetalol}$ = 0.6

P_{C3} = $SoluteBindingProb_{Diltiazem}$ = 0.35

P_{C4} = $SoluteBindingProb_{Atenolol}$ = 0.35

Consequently,

$SoluteBindingProb_{Propranolol} = 0.376\times0.5 + 0.045\times0.6 + 0.254\times0.35 + 0.325\times0.35 =$ 0.4176 (rounded to 0.42)

Table 7.3 Two sets of predicted ISL parameter values for prazosin and propranolol calculated using the two, different, FCMA-generated, drug-specific prediction strategies illustrated in Figure 7.7.

The listed parameter values (as for Table 4.5) were combined with the PCP-insensitive parameter values. The resulting ISLs were used, as for Figures 4.5-4.7, to generate the expected outflow profiles in Figure 7.8

	FCMA Drug-specific Prediction Strategies			
ISL parameter values	Prazosin1	Prazosin2	Propranolol1	Propranolol2
ISL2WetLabScaling	3.0	3.0	7.0	7.0
A2BJumpProb	0.37	0.54	0.86	0.86
B2AJumpProb	0.21	0.32	0.21	0.21
B2CJumpProb	0.50	0.44	0.50	0.50
C2BJumpProb	0.49	0.39	0.22	0.22
MetabolizeProb	0.23	0.23	0.12	0.12
BindersPerCellMin	6	6	5	6
BindersPerCellMax	11	11	10	13
SoluteBindingProb	0.40	0.40	0.56	0.46
SoluteBindingCycles	24	24	24	24
SM values (±1 SD)	0.76	0.73	0.85	0.70
SM values (±1.5 SD)	0.97	0.97	0.88	0.85

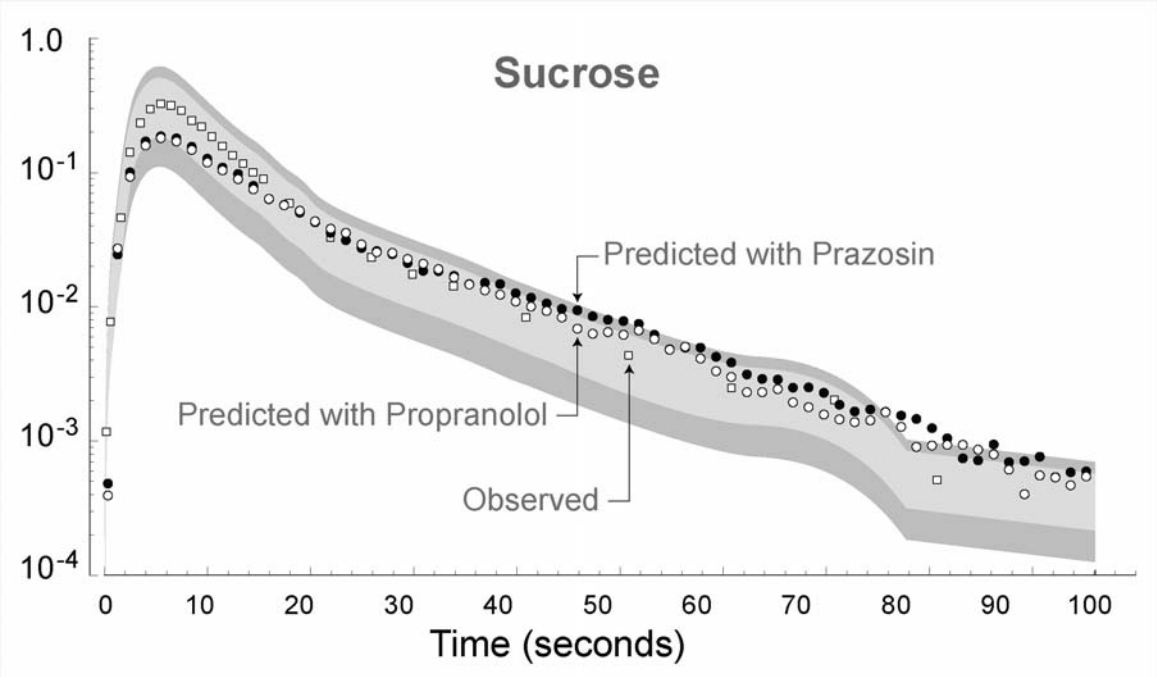

Figure 7.6 Expected outflow profiles for sucrose co-administrated with prazosin and propranolol using predicted parameter values. These semilog plots show results of simulation experiments for which the PCP-sensitive, PVs were predicted using the Method 2 values listed in Table 4.5. Black circles: SUCROSE co-administrated with PRAZOSIN; open circles: SUCROSE co-administrated with PROPRANOLOL. Otherwise, the graph components are same as for Figure 4.5.

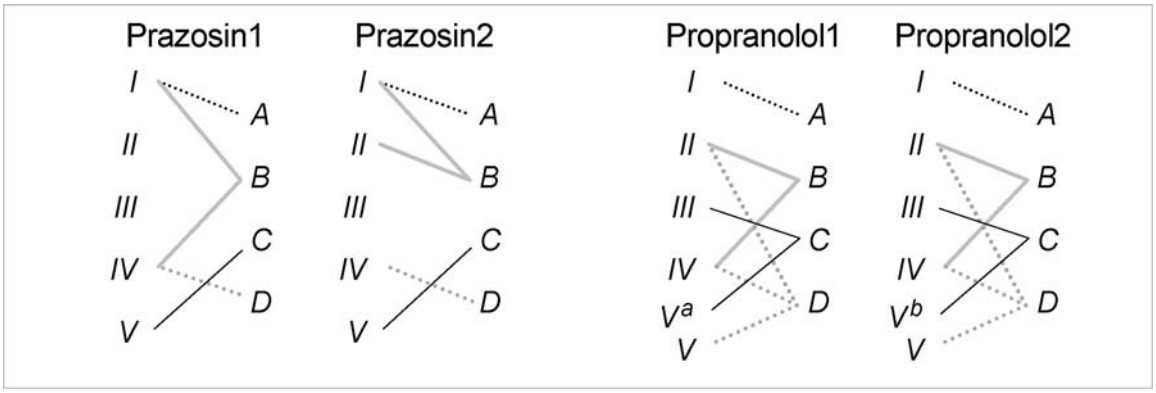

Figure 7.7 Mappings between PCP subgroups and ISL parameter subgroups (as in Figure 4.4). The four ISL subgroups are specified in Table 4.1; the five PCP subgroups are specified in Table 4.2. Each line represents a hypothesized strong influence. Except as follows, the identified PCP subgroups are the ones used by the FCMA to provide the membership degree values used in Eq. 1 to predict ISL parameter values for prazosin and propranolol in Table 7.4, exactly the same as done for Table 4.5. V^a: only HBA from group V was used; V^b: only Tautomer count from group V was used. The FCMA treats the identified PCPs as having equal influences. The influence of the other PCPs is neglected.

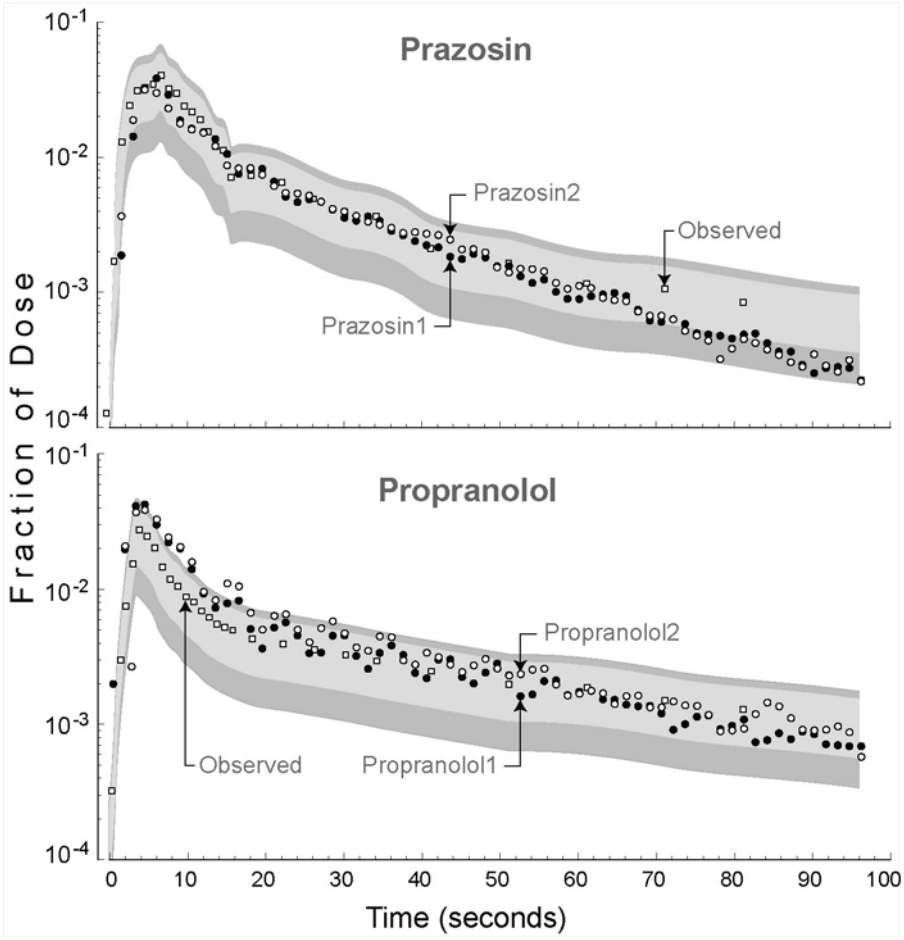

Figure 7.8 Expected outflow profiles for prazosin and propranolol using predicted parameter values. These semilog plots show results using the PCP-sensitive PVs in Table 7.3. Except as noted, the graph components are same as specified in Figure 4.5.

7.4 Sinusodial Segment Interactive 3D Visualization

The visualization (*http://furm.org/islvis/Structure9.html*) is implemented in macromedia's Director. The browser needs Shockwave plugin to view it.

7.4.1 Description

This movie is an interactive 3D visualization of two types of simulated compound, one that enters the *in silico cells* (modeling antipyrine) and one that does not (modeling sucrose), percolating through the *In Silico* Liver (ISL). The data driving the visualization is taken from the simulation through a combination of non-invasive instrumentation in the code as well as post-processing of simulation inputs and outputs. This visualization only shows the schematic structure of the ISL and how simulated solute travels through it. It does not show all the aspects of the model. In particular, The movie represents only a single Monte-Carlo trial through a single graph (modeling the lobule structure), whereas model results consist of many Monte-Carlo trials through several variant graphs. And it does not show the fine-grained spatial detail within each space. Metabolic events, for example, are not shown: the *antipyrine* simply vanishes.

The graph, shown in Screenshot 1 (Figure 7.6), is the primary component of the ISL. The nodes of the graph are called Sinusoidal Segments (SSes) and the edges are idealized connections between the SSes. The SSes are arranged at two levels: in 3 zones according to proximity to the input node (called the *portalVein*) and by the network's topology. The zonal assignment of the SSes is not shown here. Solute injected as a bolus (according to the Isolated Perfused Rat Liver experimental protocol) enters the *portalVein*, travels through the graph and exits the *hepaticVein*. An individual SS is shown in Screenshot 2

(Figure 7.7). The three grid spaces (pink: Grid A, yellow: Grid B, and blue: Grid C from Figure 2) are rendered as concentric cylinders. Two views are provided of these cylinders. One is a side view (upper right) and the other is a head-on view (lower left). At the lower right the raw grid data is directly represented. Each grid point:

- is empty, that is, no *in silico cell* object is there,
- contains an empty *cell* (rendered as mostly transparent white),
- contains *antipyrine* or *sucrose* objects outside of a *cell* (red or green),
- contains a *cell* with one or more free *antipyrine* objects within it, or
- contains a *cell* with one or more bound *antipyrine* objects within it. For example: "3 bound" means 3 bound *antipyrine* objects in a *cell*.

7.4.2 Control

In the graph view, click on a SS to see the SS view. Then click on "View network Graph" in the bottom left hand part of the SS view to return to the graph view.

In the SS view, each "smooth" control on the upper left is used to distinguish the spaces; it causes the grid data to be blurred or not. The blurring is being applied as a texture to a cylinder and could be viewed as representing the likelihood that a compound of that type is actually in that location.

The "Show" control selects the space (Grid A + Core, Grid B, and Grid C) to be observed (or not).

Selecting "Rotate" causes the cylinder to rotate.

The ">||" and "||<" controls causes in silico time to move one step forward or backward.

7.4.3 Notes

1. Compounds are injected into the In Silico Liver using a tight distribution function;
2. ANTIPYRINE, but not SUCROSE, can enter CELLS;
3. There are objects within “CELLS” that can bind ANTIPYRINE;
4. ANTIPYRINE bound within “HEPATOCYTES” in the Disse layer can be metabolized;
5. Metabolites of ANTIPYRINE are not visualized;
6. In this movie there is no more than one free solute per CELL.

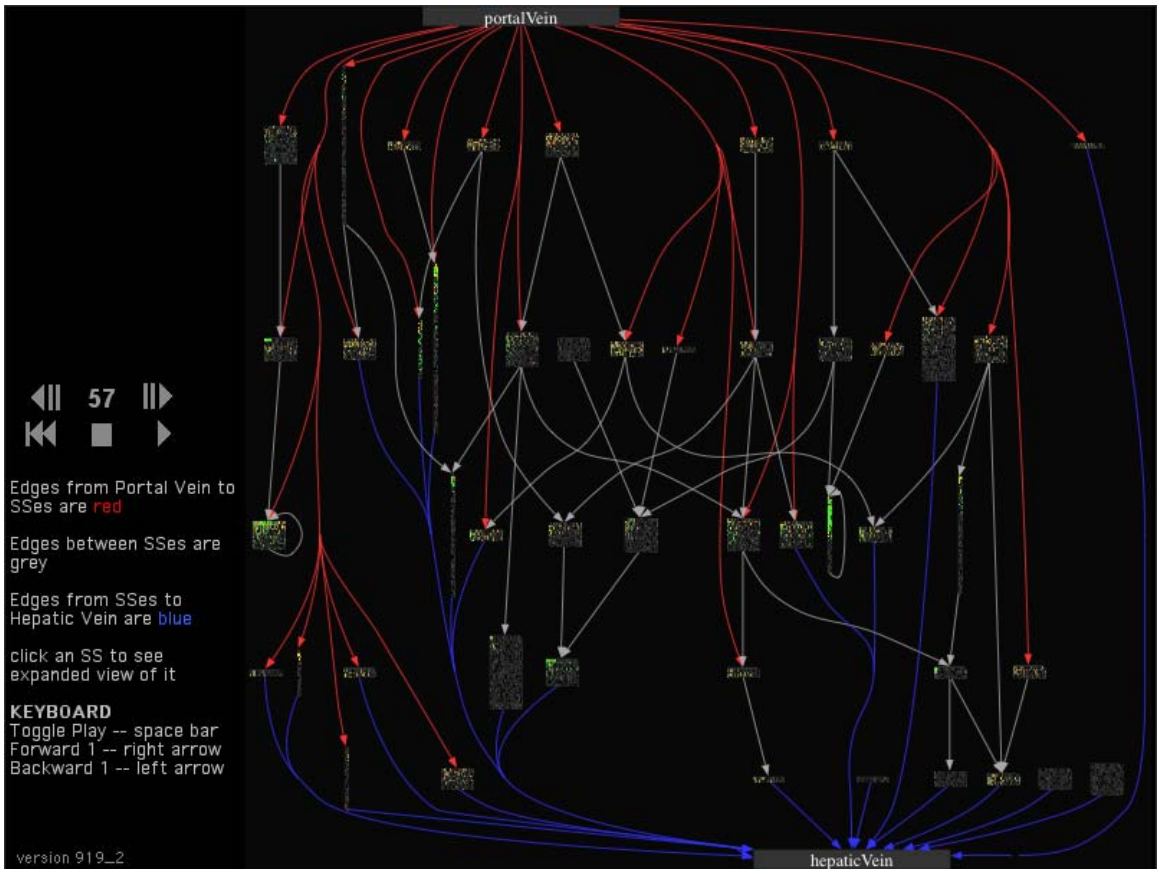

Figure 7.9 Screenshot 1: Lobule graph view

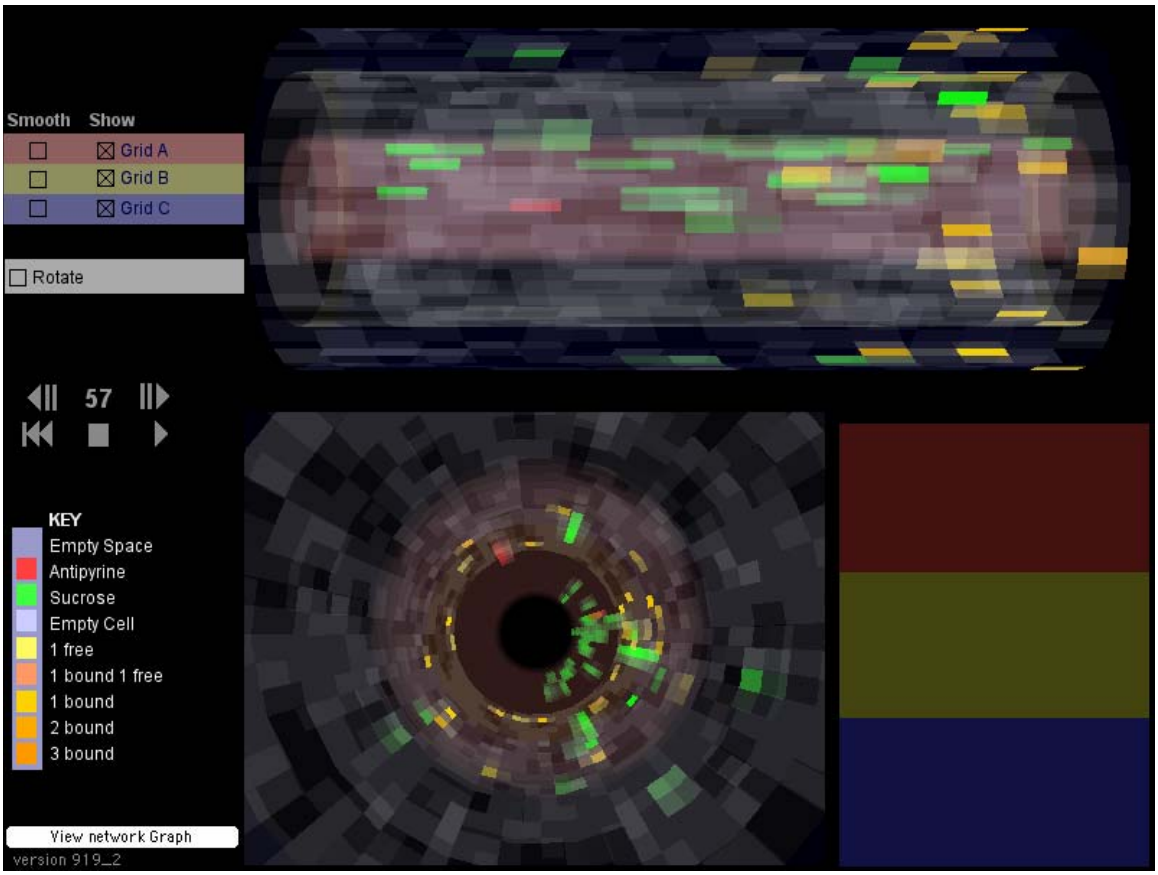

Figure 7.10 Screenshot 2: Sinusoidal segment (SS) view.