Title
In Silico Pharmacokinetic System

Author
Lam, Tai Ning

Publication Date
2010

Peer reviewed|Thesis/dissertation

In Silico Pharmacokinetic System

by

Tai Ning Lam, Pharm.D.

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

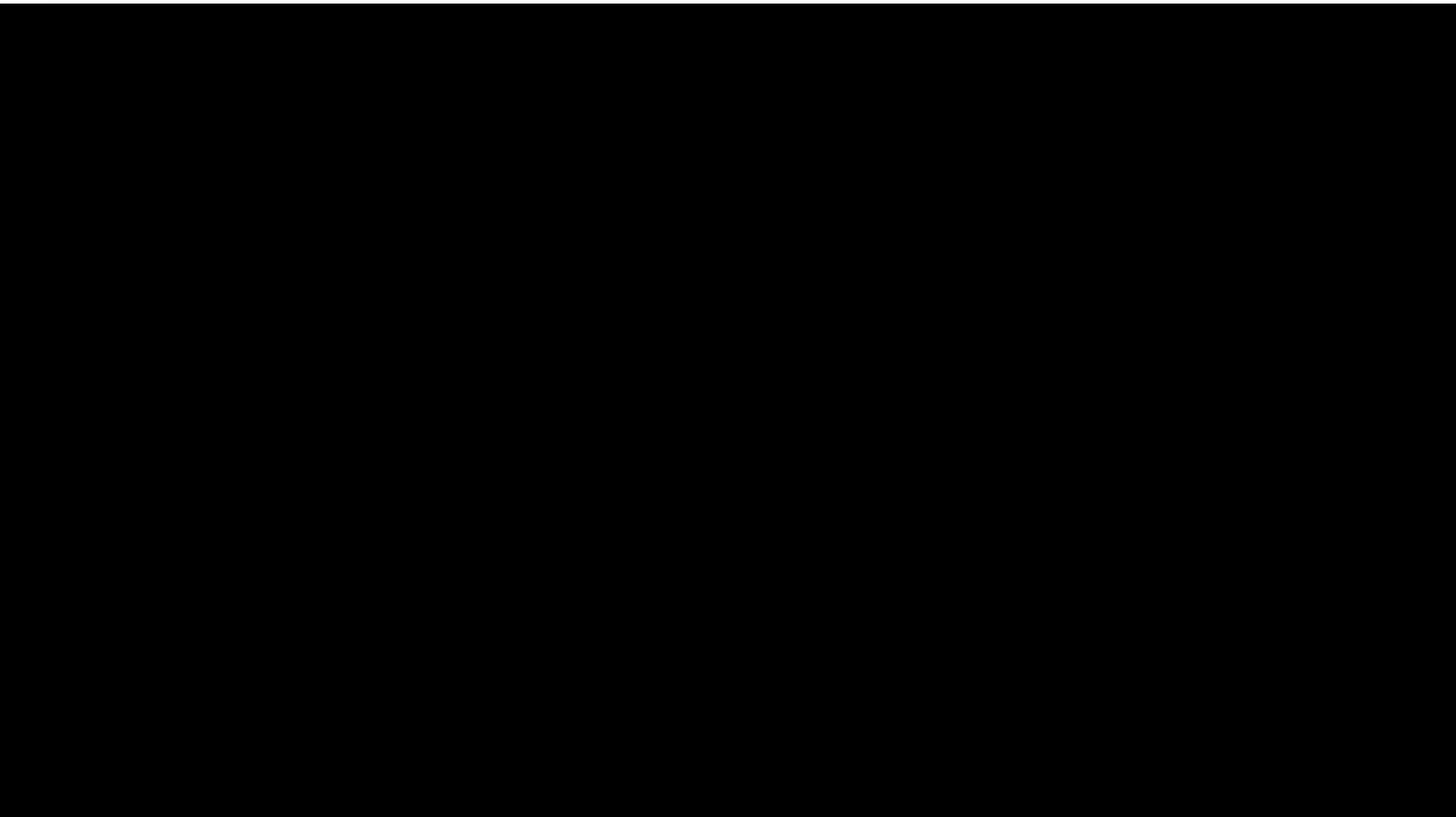
Pharmaceutical Sciences and Pharmacogenomics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO



In Silico Pharmacokinetic System

Athena, Ethan and Rosanna,

I love you to pieces;

to you I dedicate my thesis.

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my committee chair, Professor C Anthony Hunt, my mentor for 9 years. He has always been helpful, encouraging and holding the highest standard of academic rigor. He has provided clear advice and instructions all along my research yet fostered my own creativity and independence. Without him this dissertation would not be possible.

Also, a big thank you goes to my committee members Professor Leslie Z Benet and Professor Donna Hudson for providing numerous guidance and inspiration. And to Professor Betty-ann Hoener who got me immensely interested in Pharmacokinetics with her lively lectures (and jokes) in my first year Pharmacokinetic class, and who introduced me to Professor Hunt.

It was my pleasure and privilege to work with all the fascinating individuals from the Biosystems group - Glen Ropella, Dr Sunwoo Park, Dr Suman Ganguli, Song Li, Dr Yu Liu, Dr Yuanyuan Xiao, Dr Lana Garmire, Dr Li Yan, Dr Sean Kim, Dr Jon Tang, Jesse Engelberg, Dr Shahab Sheikh-bahaei, Anita Grover and our most missed administrative assistant, Pearl Johnson. All the group discussions, presentations and interactions have made up an integral part of my graduate training. I would also like to thank Dr Yvonne Lau, Dr Stephane Mouly and Professor Paul B Watkins for providing original data and helpful discussions.

I thank the CDH Computational and Systems Biology Predoctoral Fellowship and the University of California, San Francisco, School of Pharmacy Graduate Program in Pharmaceutical Sciences and Pharmacogenomics for financial support.

A portion of text of this dissertation is a reprint of the material as it appears in the journals Drug Metabolism and Disposition and Journal of Pharmacology and Experimental Therapeutics. The coauthor listed in the publications directed and supervised the research that forms the basis for the dissertation.

Last but not least, this acknowledgement would be incomplete without my utmost appreciation of my family: Rosanna, my parents and my brother, my little angel and demon Athena and Ethan, my two adorable cats, and my cute bearbears.

Abstract

In Silico Pharmacokinetic Systems

By

Tai Ning Lam

Models are frequently and ubiquitously used in all pharmacokinetic investigations. The familiar inductive equation-based pharmacokinetic models formulate hypotheses about data; they alone cannot provide mechanistic insights. We need models that have extant, working mechanisms that generate emergent properties analogous to how phenomena emerge during wet-laboratory experiments.

In this dissertation, I report a new class of synthetic, agent-based, discrete events models and simulations, with the objective to provide mechanistic insights. Validated, biomimetic software components are plugged together to form in silico analogues of the referent experimental systems. Each synthetic analogue is a mechanistic hypothesis: execution produces an observable phenomenon.

The recirculating in silico livers (RISLs) are in silico analogues of isolated perfused rat livers during an experiment in which digoxin is administered, alone or in combination with either an uptake or efflux inhibitor. A RISL that comprised four time-variant mechanisms and new enzyme and transporter components achieved the most stringent similarity measure: simulated digoxin and metabolite perfusate levels were experimentally indistinguishable from the referent data. The mechanisms simulated

unanticipated loss of hepatic viability during the original experiments: erosion of hepatic accessibility and of enzyme and transporter activities.

In silico experimental Caco-2 (cell monolayer) cultures (ISECC) are analogues of the confluent, asymmetric cell monolayer used in vectorial transport studies. To seek an explanation for the observed paradoxical saquinavir transport data, I followed an iterative refinement protocol that enabled discovery of plausible, new mechanistic details. The ISECC surviving the most stringent similarity challenge produced transport data statistically indistinguishable from referent observations. It required heterogeneous intracellular spaces; a biased distribution of metabolizing enzymes; and restrictions on intracellular drug movement.

Experimenting on synthetic analogues, such as RISLs and ISECCs, provides a formerly unavailable means of discovering and testing new mechanistic hypotheses. It is a powerful expansion of the scientific method: an independent, scientific means to challenge, explore, and improve any inductive mechanism. Validated, biomimetic analogues are concrete instances of hypothetical yet plausible mechanisms, and would replace vague, unverified concepts. The collection of mechanisms, rules, assemblies, and interactions of components can be subjected to testing and falsification, and in the absence of other competing theories, stands as the current best mechanistic hypothesis for the phenomena.

Table of Contents

List of Figures

List of Tables

1. Models in Pharmacokinetic Investigations

Pharmacokinetics (from ancient Greek *pharmakon* "drug" and *kinetikos* "to do with motion") is the study of what the body does to the drug – the processes of absorption, distribution, metabolism and elimination. Pharmacokinetics includes the study of the mechanisms of absorption and distribution of an administered drug, the rate at which a drug action begins and the duration of the effect, the chemical change of the drug in the body by, for example, metabolizing enzymes, and the effects and routes of excretion and/or elimination of the parent drug and its metabolites.

1.1. The use of models

I first introduce the use of models in pharmacokinetic investigations. All scientific investigations begin with a hypothesis, or a question (quantity of interest) about a specific aspect of a system of interest – the primary system, or the *referent* system ("referent" hereafter) – as well as knowledge and assumptions about that system. The knowledge is a mental model about the system of interest. The goal of scientific investigation is improved understanding of the referent, by confirming the hypothesis (or more frequently in statistical hypothesis testing, falsifying the

null-hypothesis) about the referent system. To test the hypothesis, a secondary system of the referent, frequently in the form of an extant *model*, has to be constructed. The secondary system can take many forms, some of which are described below.

While most scientific investigations involve one primary system (referent) and one secondary system (model), since the earliest era of pharmacokinetics, scientists have been using *two* classes of secondary systems (models) to study pharmacokinetic phenomena. The first one is an experimental model, consisting of living biological components, to which drugs are administered and measurements are made. The second is an induced mathematical model to describe the observed pharmacokinetic data, usually in the form of an equation-based model, whose structure is influenced by prior knowledge about the biological mechanisms.

The following aims to dissect the ordinary pharmacokinetic study to highlight the frequent, ubiquitous use of models in them. The goal is to shed light on the use of models at every step of common pharmacokinetic investigations.

Guided by theory and knowledge about the referent, scientists design experiments such that an intervention would allow measuring the quantity of interest, or expose or disturb biological mechanisms, the consequence of which is observed and measured. Most wet-laboratory experiments involve biological systems that are reductions from the referent and therefore less complicated, for good reasons: ease of manipulation, lower costs, higher efficiencies, ethical considerations, etc. Select man-made experimental apparatus are connected to the extracted biological parts, which are supported by specialized, engineered medium. An intervention, designed to disturb the biology, is given under abnormal physiological conditions. The experimental biological system is nontrivially distorted in many ways, and is, in short, radically different to the referent biological system. Hence, experiments themselves are models – *experimental models*. Despite the above limitations, experimental models may still be the closest alternative to the referent biological system. During this selection and construction of

experimental models, assumptions, whether necessary or for convenience, explicitly stated or implicit in the mind of the experimenter, become an inseparable part of the experimental model. Because most wet-lab experiments assemble parts to form a whole experimental system (model), they are constructive, or *synthetic*, in nature.

Next, observations are made from the experimental model. The biological system under experimentation, although disturbed, presents many aspects for observation. The concentrations of the administered drug and its metabolite in the medium are one aspect. The gross anatomical appearance of the tissue under experimentation is another aspect. Among them, only a few are observed and recorded. The experimental protocol (part of the experimental model) calls for even fewer to be measured. In most pharmacokinetic studies, samples of the medium from the experiment are taken and later assayed with validated bioanalytical methods. At the end, measurements usually consist of the time-course profile of an administered drug and its metabolite, expressed in concentration terms, and/or the amount of the drug and/or metabolite(s) in the biological tissue (cell lysate) or fluid (urine) at some point during or after the experiment. The observed data are a model of an aspect of the experimental system from the perspective of the experimenter's interaction with it. Outside of these data, some observations, especially unexpected, anomalous ones, are ignored. The unobserved aspects are lost forever as the experimenter concludes the experiment and cleans up the apparatus.

The measurements from experiments form a data set, to be analyzed, using an additional model – an *inductive* model. In most pharmacokinetic experiments, the expected phenomenon is not observable by direct visualization, but rather by changes in (trends in) measured quantities. Usually, a pharmacokinetic model is used to relate these measured quantities to the expected phenomena. With the exception of a question calling for a quantity to be directly measured, in pharmacokinetics studies, two types of questions are generally asked. Firstly, does an intervention alter the pharmacokinetics of the compound of interest? Secondly, what

pharmacokinetic model best describes the data, and what are the best parameter estimates for the selected pharmacokinetic model? In both type of questions, a model structure is assumed, and observed data are fitted to the selected model structure. More precisely, for most pharmacokinetic models involving differential equations, a measured state of the integrated equations resulting from a presumed, parameterized model is compared to the observed data. In the former type of question, the steps follow common statistical hypothesis testing procedures. More often, the investigator derives quantities using additional (pharmacokinetic) models, and uses the derived quantities for hypothesis testing. A test statistic is then calculated and the hypothesis is either rejected or confirmed.

> Example: suppose the investigator hypothesizes a drug-drug interaction between drugs A and B. Specifically, co-administration of drug B would reduce the rate of elimination (a natural phenomenon) of drug A. There are multiple ways to test that hypothesis. Assuming drug A's metabolism is inhibited by drug B, one way is to measure the amount of metabolite of drug A formed at the end of experiment. The hypothesis can be (re)stated: metabolite formation of drug A is reduced with co-administration of drug B. This restatement does not need an additional pharmacokinetic model. However, the prior assumption may be unrealistic. Assuming only that elimination is linearly related to drug concentration, an alternative is to quantify the rate of elimination by measuring the concentration-time profile of drug A, with and without co-administration of drug B, and selecting a pharmacokinetic model to describe the time-course profiles. Suppose drug A is known to follow one-compartment kinetics, and so the respective pharmacokinetic model is used. In the selected pharmacokinetic model, *clearance* (a pharmacokinetic parameter) represents the rate of drug elimination. Hence, the hypothesis can be (re)stated: clearance of A is decreased when co-administered with drug B. Either way, the investigator, based on his knowledge about pharmacokinetic properties of drugs A and B, designs an experiment where one group of subject animals are first given drug A alone (control) and another group given drugs A and B together (treatment). In the former way, the amount of metabolite is collected and measured at the end of experiment, summarized, and compared with a

Student's t-test. In the latter way, a time series concentration of drug A is taken, clearance values of drug A from both control and treatment are calculated and compared. Under the null hypothesis, where there is no interaction, a statistical model is assumed: the difference in the amount of metabolite, or in the clearance values between both control and treatment groups would follow a t-distribution with a mean equal to zero. A test statistic – the ratio of difference in mean values to its respective variance – evaluates the probability that the measured amounts of metabolite or calculated clearance values are consistent to the selected statistical model (i.e. the t-distribution.). A decision can be made to reject the null-hypothesis, and to conclude there is drug-drug interaction, if the probability (p-value) is sufficiently small. Depending on the drugs' properties, different or additional pharmacokinetic models may be needed to calculate clearance values. Note that rejecting the null-hypothesis when the amount of metabolite is measured leads to knowledge about the mechanism: that metabolite formation of drug A is inhibited by drug B. However, decrease in clearance value with co-administration of drug B does not tell anything about why or how the rate of elimination of drug A is reduced – there is no mechanistic knowledge generated.

In the latter type of questions, where the estimation of pharmacokinetic parameters is of interest, investigator would select a pharmacokinetic model based on the pattern in observed data, as well as knowledge and assumption about the experimental system. The observed data are then "fitted" to the selected model, and the model parameters are estimated. The fitting steps are usually done with computational software, using an optimization algorithm. An optimized parameter vector is calculated (estimated), which together with selected model structure, best describes the observed the pharmacokinetic data. In essence, the information from observed data – a large collection of observations of low information content – is abstracted and simplified to form a representative model, which is a simpler description of higher information content, by using a selected model structure and estimated model parameter values. A statistic of goodness-of-fit criterion can be defined and evaluated, and a model with sufficiently high goodness-of-fit value is deemed a "good", "validated" or "adequate" description of the data.

Example: suppose a whole animal pharmacokinetic experiment was performed. The subject animal was given an extravenous dose of the study drug, and blood samples of the animal were taken at specific time points, and concentrations of the study drug in the blood sample were assayed. The concentration-time data were plotted and revealed an upward curve, followed by a peak and a two phase decline. Based on the observed pattern and prior knowledge, the investigator concluded – induced – a two-compartment model with first order input would adequately describe the data. In other words, the investigator decided the pharmacokinetics of the test drug was consistent with the mechanisms implied in a two-compartment model with first order input. Specifically, the investigator inferred that absorption into central compartment was proportional to the amount of the drug present, and that the distribution of the drug inside the body can be simplified to distribution between a central, rapidly equilibrating pool and a peripheral pool. A corresponding equation in the form of a sum of exponentials was derived, and entered into a computer program whose algorithm was to "fit the data to the curve". More specifically, given the model structure (set of equations) and input data (observed concentration-time profile) the computer algorithm found an optimized set of parameters that would give a minimized prediction error (or maximized likelihood.) The model prediction – measured states of the optimized model, in the form of a set of parameterized equations – was compared to the observed concentration-time profile, and was deemed adequate. The model structure and its parameters were reported and the modeling process was then complete. The pharmacokinetic quantities of interest may then be calculated from model parameters. Knowledge about the pharmacokinetics of the study drug was limited to conceptual requirement of the validated (or falsified) pharmacokinetic model: that the concentration-time profile appeared to be consistent (or inconsistent) with the selected model. Whether those requirements were mechanistically realistic, or what mechanisms led to the inconsistency is unknown.

In both types of questions, the selected model is induced from the observed data – they are *inductive* in nature. Interpreting observed data creates understanding about the referent system. Either a quantity is measured, or a hypothesis is confirmed or rejected. Describing the referent system using a pharmacokinetic model creates predictive power, as well as the ability to leverage

established knowledge from other pharmacokinetic systems that are adequately described by the same pharmacokinetic model. Parameters from pharmacokinetic models, although not physiologic parameters, are usually useful in predicting drug disposition, above and beyond describing them.

It is important to draw the distinction between pharmacokinetic properties and pharmacokinetic parameters. Pharmacokinetic properties are observations of the intrinsic properties of a unique system: a drug interacting with a particular biological system. A pharmacokinetic property is an aspect of that system (e.g. whether the drug is excreted unchanged in urine, or the extent it is distributed in adipose tissue under a particular dosing condition). The properties are not fully separable into those of the drug and those of the biology. When we nevertheless do so, e.g., to compare the "pharmacokinetic properties" of two drugs, we apply a separation model that can be characterized by the assumptions made. For example, one can compare the extent being excreted unchanged in urine for two drugs, *assuming* the renal function is the same and “normal”. Pharmacokinetic properties are independent of any pharmacokinetic models. Pharmacokinetic parameters are numerical values of pharmacokinetic models’ parameter (e.g. renal clearance, volume of distribution.) Conflating them is the result of failure to recognize the use of inductive models in pharmacokinetic studies. Pharmacokinetic models are selected such that their parameters map to, or estimate, important pharmacokinetic properties. This is why pharmacokinetic models are useful and valuable. Pharmacokinetic parameters do not inform how pharmacokinetic properties come about. Hence, understanding pharmacokinetic properties, or in other words, the generating mechanisms of the observed (and measured) phenomena, is what makes pharmacokinetic studies interesting and worthwhile, and that requires additional methods beyond fitting pharmacokinetic models.

1.2. The gap

The observation in the preceding section about the two classes of models can be organized in the following figure, which depicts models used in biomedical and pharmaceutical research.

Fig. 1.1 depicts the space of major model classes used today in pharmaceutical and

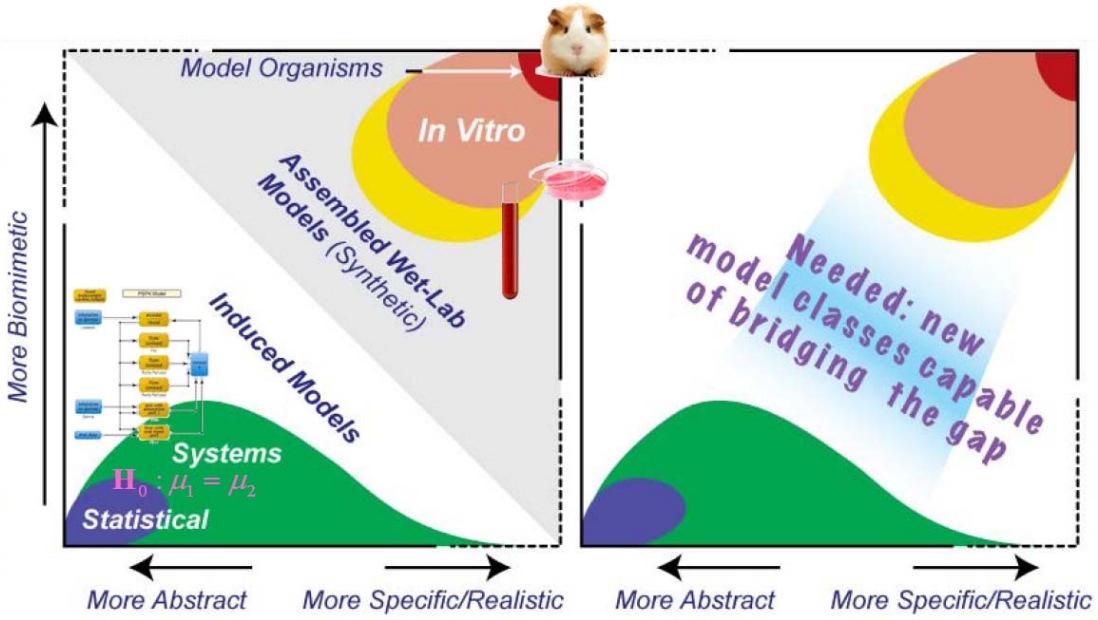


Figure 1.1 Major classes of models used in pharmaceutical biomedical research. Left: For purposes of illustration, model types, and the analytic and explanatory methods that use them, are arranged according to abstraction level versus biological character; in reality they are not independent. The arrangement of model types is discussed in the text. More abstract indicates a greater capability for simple and focused representation. More realistic indicates a greater capability for aggregating collections of facts. The biological axis (biomimetic) indicates the degree to which a model resembles and behaves, at some level of detail, like its wet-lab referent. An inductively defined, equation-based model, for example, can mimic time-course measures of an aspect of a biological system very well (high, aspect-specific biomimesis), but, as a complex algorithm implemented atop a numerical integrator, it is not at all realistic (yet the conceptual model to which it is tied may include some realistic features). An unvalidated agent-based model can implement detailed representations of almost any physiological process and yet be incapable of behaving like the referent in any particular context; hence, it exhibits high realism but little biomimicry. **Right**: Illustrated is the gap that exists between inductive, mathematical models and the wet-lab models used in biomedical research. Models that can bridge the gap will be biomimetic computational analogues of their wet-lab counterparts.

biotechnology research. Model types on both sides have different uses. The ultimate referents for both are specified subsets of patients. Not represented are the conceptual mental models on which all scientists rely and the prosaic models (often supported by sketches of idealized mechanistic events) that describe these mental models. Moving to the right in Fig. 1.1, model aspects (the perspective taken when a system is observed) become more realistic, relative to their referents. Moving up, similarities between model and referent attributes increase: the models become more biomimetic. The diagram excludes patients. Model organisms are in the upper right. *In vitro* cell and tissue models are next. Below them are cell-derived systems. Statistical and correlative models are to the lower left. Above them are the familiar, induced mathematical computer models. Included within the latter are network, pathway, pharmacokinetic, pharmacodynamic, and physiologically-based models; they are induced from data.

In Fig 1.1, the most biological model is a model organism. A specific species of organism is carefully selected or even engineered for experimentation to provide the needed biological parts and mechanisms to answer the scientific question. Often that requires substantial similarity between the model organism and the referent (human pharmacokinetics). Frequently used model organisms in the pharmacokinetic context are mice, rats, dogs, monkeys, and their engineered, or genetically modified, counterparts.

In vitro systems usually consist of living biological components taken from model organisms. Only the biological parts thought to be relevant to the research question are extracted and studied. The extracted, yet living biological parts are placed in a well-controlled, engineered environment in which the experiment is performed. Because of its lower cost, ease of use and lesser variability and complexity compared to *in vivo* systems, *in vitro* systems are often preferred by experimentalists for deducing biological mechanisms of action. However, the controlled conditions present in the *in vitro* system differ significantly from those *in vivo*, and may give

misleading results. Therefore, *in vitro* studies are usually confirmed by *in vivo* studies. An example of an *in vitro* system is cell cultures grown in Petri dishes.

An even less biological model can be made with an assembly of non-living biochemical molecules. An example is an enzymatic study in a test-tube.

The familiar statistical models, being numbers on paper completely removed from their biological context, are the most abstract and least biomimetic models; they are at the lower left corner. In essence, the statistical models assert that observations made could be adequately described by some statistical distributions. Falsification of that assertion based on statistical probability leads to a more probable alternative hypothesis. Because the alternative hypothesis is usually formulated with conceptual mapping to the referent, rejecting the null and accepting the alternative hypothesis leads to an improved conceptual understanding of the referent; however, no causal basis can be asserted from statistical tests.

Above and to the right of statistical models are mathematical models used to describe the relationships between quantities measured from biological experiments, which in turn reflect the biological processes in a quantitative way. Within this class of models, the model could take many forms, ranging from simple mathematical relationships to complex systems of differential equations. These mathematical relationships are more biologically relevant in a number of ways: 1) the quantities in the equations often map to related quantities in the biological process, 2) predictions from these models are usually verified by performing the biological experiment and measuring the relevant biological quantities; 3) the form and structure of the mathematical models are induced by observations made from biological experiments; 4) selection of one model versus another is often informed by prior knowledge about biological mechanisms and as such the selected model usually conceptually maps to biological mechanisms in some ways. Examples include the familiar Michaelis-Menton enzymes kinetics, pharmacokinetic compartmental model, and physiologically-based pharmacokinetic models.

The preceding analysis show the gap between the two classes of models – inductive, mathematical models on the lower left, and experimental, biological models on the upper right. Models that span the gap will be biomimetic analogues of their wet-lab counterparts; they can be expected to be used for evaluating explicit mechanistic hypotheses in the context of many aspects of the referent.

1.3.Hypotheses about the mechanisms

When a hypothesis is confirmed by the familiar statistical test, or by rejecting a simpler pharmacokinetic model and accepting a more complex one with different parameter values, how much can we say about what spatiotemporal mechanisms play roles in the emergence of a pharmacological response or a pharmacokinetic phenomenon? Not much: hypotheses confirmed or rejected are about the *data* observed from the experimental model, not about the *mechanisms* of the referent, as illustrated in Fig 1.2. In the familiar inductive pharmacokinetic modeling and simulation, the determination of pharmacokinetic parameters limits the use of the model (and its parameters) to the current system of interest where the current pharmacokinetic model applies; it is model-dependent. It is fragile to changes in context. The same model can be used to represent any number of individuals and even different mammals, but only under similar conditions. Even in the specific cases where pharmacokinetic models are informed by prior knowledge on mechanisms, the gained knowledge from these models remains conceptual in nature. Concepts require subjective interpretations by scientists. Only an understanding of the pharmacokinetic properties and the mechanism underlying the properties would provide concrete answers to questions about how biological phenomena emerge, and to translate the results to systems in a different context or to a different level of biological complexity. The inductive procedures alone described above lack the possibility of providing concrete mechanistic insight; it cannot generate nor test hypotheses about mechanisms, or in other words, the generator of the phenomena.

The exploration of an inverse map from phenomena to generators requires the researcher to hypothesize generators that could result in the observable phenomena. The question posed is this: given phenomena, what hypothetical generators (and measures) might generate them? The goal is to find a hypothetical generator in the form of mechanisms and interactions constructed from its components, i.e., a simplified, synthetic wet-lab experimental apparatus and corresponding measurements, or an *in silico synthetic analogue* (a collection of biomimetic mechanisms constructed using object-oriented software tools). The constructed experiment, or computational analogue, when executed, is expected to produce the target phenomena. When one is found, either wet-lab or in silico, we are still ignorant of the referent system as we were before, but we now have a concrete instance of a plausible mechanism where before we had only vague,

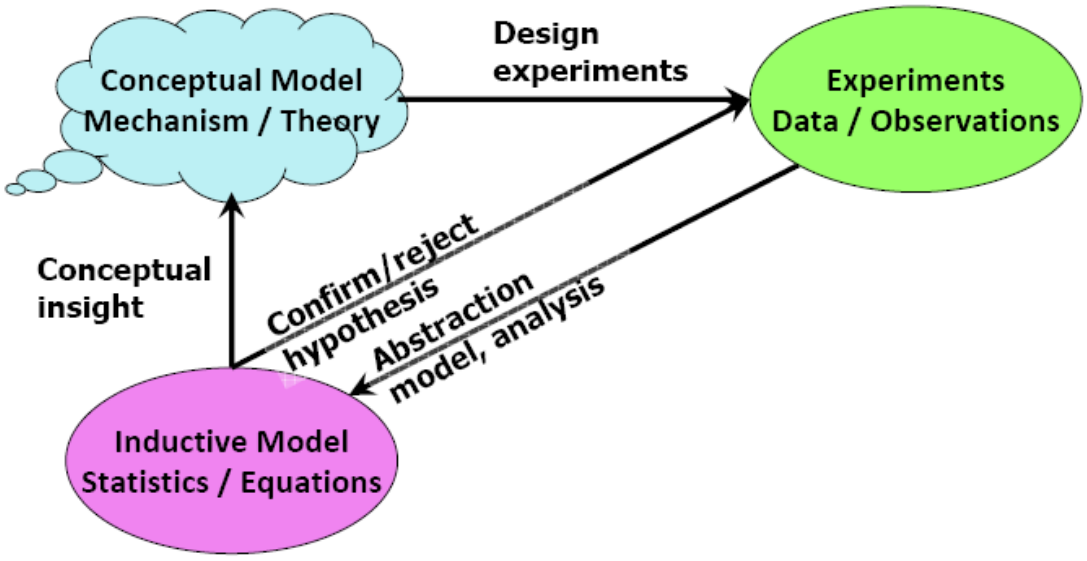


Figure 1.2 Hypothesis about data. Guided by prior knowledge about the biological system of interests, a scientist designs experiments and protocols. Measurements are made from the experiments and are subsequently abstracted and analyzed (modeled) with the familiar inductive methods. Hypothesis confirmed or rejected is about the data generated from the experiment. The question being asked is, whether the observed data are consistent with the model? In itself, it does not inform the underlying mechanism. If the inductive model is specified by mechanistic knowledge, then conceptual insight may be possible: rejecting such a hypothesis (model) may mean the hypothesized mechanism is inadequate. It remains unknown how that rejected mechanism is inadequate.

unchallenged concepts. In the absence of other concrete, competing theories, this plausible mechanism can stand as the current best explanation for the phenomena. A full study of any inverse map requires abduction (several hypotheses), deduction (simulation, or experimentation), and induction (falsification) and that can only be done with concretizable hypotheses: synthetic wet-lab experiments, or synthetic computational analogues. It cannot be done with hypotheses that remain conceptual.

1.4. Synthetic, mechanistic analogues

In order to demonstrate that we have achieved deeper insight into how molecular details interface with and exert influence at higher levels and emerge as features of observations made from experiments, and to understand biological responses and their plausible generative mechanisms, we need models and methods that can bridge the gap as presented in Fig 1.1. We need models that are made increasingly more similar to their referents—models that have extant (actually existing, observable) mechanisms that generate emergent properties analogous to how phenomena emerge during wet-lab experiments. We need to build these extant, working mechanisms that exhibit some of those same phenomena. We need analogue mechanisms that are transparent such that investigators can trace how temporally and hierarchically networked generators cause their respective phenomena. These extant, plausible, mechanistic analogues are fundamentally different from the traditional pharmacokinetic models that "fit the data." Such analogues are instantiations (represented with concrete instances) of (abstract) mechanistic hypotheses – hypotheses about the generation of a phenomenon – that can be subjected to testing and falsification.

While we cannot yet build such analogues from biochemicals, we can build extant biomimetic mechanisms using object-oriented software tools. Those models will be synthetic, as are wet-lab models. When object-oriented, software engineering methods are used to implement a mechanism as described in the following chapters, the resultant software implementation is an

abstract, but extant mechanistic hypothesis: the components (objects), their assembly and their interactions collectively will produce a mechanism upon execution. By so doing, we have instantiated a mechanism in silico, as depicted in Fig 1.3. A consequence of mechanism execution will be the emergence of phenomena that are similar (or not) to pre-specified phenomena, such as a response following exposure to a xenobiotic. An important use for such models will be testing hypotheses about mechanisms (rather than about patterns in data). Execution produces a simulation with features that we can measure; those measurements enable testing the hypothesis.

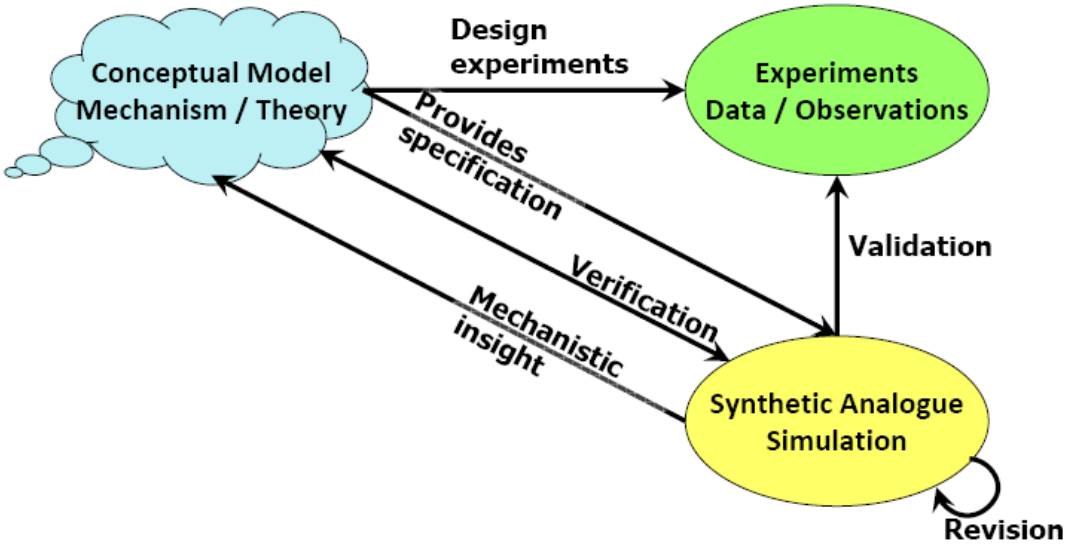


Figure 1.3 Hypothesis about mechanisms. Alternative to using familiar inductive models, constructing a synthetic analogue begins when a scientist organizes his or her knowledge, assumptions, and ignorance about mechanisms to form an abstract mechanistic model. The prior knowledge provides specification as to how the analogue should be constructed and what components may be needed. Next, the scientist instantiates the abstract mechanistic model using software algorithms and tools. Implementation of the algorithm gives an executable program. Execution of the program (simulation) provides observations, and measurements are made and compared to those from its referent experiments. The scientist may verify if the simulated results are consistent with prior knowledge and assumptions. The analogue may undergo many cycles of revision, during which hypothesized in silico mechanisms may be added and modified, before it is validated by experimental data. Each of the analogues is therefore a mechanistic hypothesis. When a degree of similarity between simulation results and experimental observations is established, one can assert (or hypothesize) that the in silico dynamics and mechanisms, consistent with their level of resolution, may have biological counterparts. This provides mechanistic insight into the referent.

If phenomena similarities meet some pre-specified criteria, then the simulation stands as a challengeable yet tested theory about abstract yet plausible mechanistic events that may have occurred during the wet-lab experiments.

My thesis is thus: experiments with synthetic, mechanistic analogues of biological system provide mechanistic insights to the underlying biology. The analogues serve as informative experimental devices to explore potential consequences of mechanistic interventions and to facilitate hypothesis generation, falsification and selection. In the remainder of this dissertation, I present a novel method for the study of pharmacokinetics using computational, agent-based modeling and simulation: the use of synthetic, mechanistic analogues to provide mechanistic insights.

This dissertation is organized as follows. Chapter 2 provides an overview of models used in pharmaceutical research, in pharmacokinetics in particular. I begin by illustrating the two modeling approaches used in today's pharmaceutical research: the wet-laboratory, synthetic, experimental systems using biological components in a well-engineered experimental apparatus and the computational, inductive, data-driven mathematical and statistical models consisting of equations and systems of ordinary differential equations. For biological models, I focus on the *in vitro*, Caco-2 cell monolayer cultures in a Transwell system and an *ex vivo*, isolated perfused rat liver system. For mathematical models, I briefly describe the widely-used analysis and modeling methods including compartmental analysis, noncompartmental analysis, and physiological-based pharmacokinetic modeling. Then, I summarize the key features and principles for agent-based models used in developing and instantiating a synthetic, mechanistic model of a referent pharmacokinetic system.

In chapter 3 I describe a few basic features of the In Silico Pharmacokinetic System (ISPKS) that is common in the subsequent chapters. They provide foundations on which the

Recirculating In Silico Liver (RISL) and the In Silico Experimental Caco-2 Cell Monolayer Cultures (ISECC) are built.

In chapter 4 I detail the effort of extending the ISPKS to model a recirculating isolated rat liver, the result of which is the Recirculating In Silico Liver (RISL). From those experiments, I assert that there was deterioration in the function and viability of the isolated liver during the referent experiment, resulting in time-dependent changes in pharmacokinetic mechanisms.

In chapter 5 I describe the instantiation of ISPKS to represent a Caco-2 monolayer culture in a vectorial transport experiment. I was able to provide a plausible explanation for the unresolved paradoxical observation from the investigators: that heterogeneous intracellular microenvironment may have played a role in saquinavir transport and metabolism.

I conclude with my perspective of how synthetic modeling and simulation may be made scientific and how it will impact today's pharmaceutical research and care in chapter 6.

2. Pharmacokinetic Model Approaches

2.1. Introduction

The need to achieve deeper mechanistic insight into biological systems has been the motivation for the recent development and application of a new class of models referred to by Fisher and Henzinger (2007) as executable biology. Hunt et al. (2010) have referred to the method used as the synthetic modeling and simulation method. The method is distinct from the traditional inductive pharmacokinetic (PK) and pharmacodynamic modeling method. A computational model from this new class represents a hypothesis about the mechanism underlying the data of interest, whereas a traditional PK model formulates hypotheses about the data.

Aspects of the methods and approach used in developing and validating this new class of models are fundamentally different from those used in developing traditional pharmacokinetic and pharmacodynamic models. In this chapter, I present background information on the two classes of models and explore differences and similarities between them.

2.2. Wet-lab experiments

An experimental system used in pharmacokinetic experiments includes a test drug and its mode of administration, biological components (organ, cell, or enzymes) of interest, as well as everything else called for in the experimental protocols; for example, perfusate or growth media, the investigators and their surgical techniques, sample collection and storage technology, monitoring devices, and analytical equipment. Together, the experimental apparatus are the world within which the biological components in the experimental system live, and within which drugs are absorbed, distributed, metabolized, and eliminated. Consequently, they influence, and so are "part of," the biological mechanisms of interest. During a pharmacokinetic experiment, cellular components interact with transiting drug molecules. The mechanistic details, envisioned in the mind of the scientist, are represented at the middle in Fig. 2.1. They range from macromolecules binding the drug and cell organelles that are thought to sequester the drug, to microarchitectural features of intracellular organelles or extravascular space. Collectively, these events change the concentration-time profile of the drug. They cause systemic behaviors—phenotypic attributes that may be observed and measured in a variety of ways. Although several attributes may be measured during such experiments, the drug and metabolite levels in perfusate or the medium are most often the focus of attention. The raw, drug-level time series data are often transformed or manipulated for easier viewing or interpretation. For example, the final observable may be the fraction of administered dose per unit of perfusate plotted versus time or ratio of flux of dose to initial concentration divided by surface area (i.e. apparent permeability). Together, this is the experimental (apparatus) model. The time series data resulting from measurements of the observable, as just one set of possible measurements that can result from applying measures, are a concrete filtrate of events that occurred within the system during experiments, and so represent a perspective-dependent model of the experimental system. I present two particular experimental models in the following sections, the Caco-2 cell monolayer culture, and the isolated perfused rat liver.

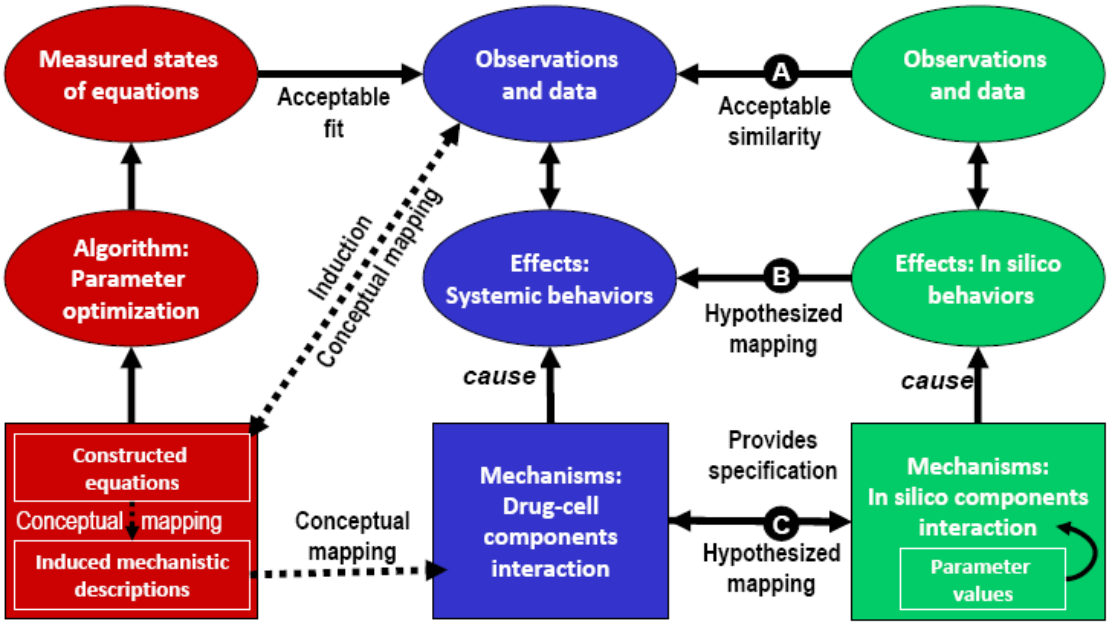


Figure 2.1 Inductive and synthetic modeling. Two different modeling approaches, inductive traditional pharmacokinetic modeling and synthetic modeling and simulation, are fundamentally different from each other. Here we have the two classes of model to the same referent system: perfused rat livers. Shown are relationships between the wet-lab models used for perfused liver experiments, traditional PK models, and our analogue Recirculating In Silico Liver (RISL). **Center**: Perfused rat livers in their experimental context are the referent systems. During experiments, liver components interact with transiting drug molecules causing changes in the compound's concentration-time profile. Systemic behaviors are reflected in the collected data. **Left**: The researcher identifies patterns in the wet-lab data and induces a mechanistic description of what is thought to have occurred. That mental model provides abstract, conceptual mappings from that description to hepatic mechanisms. A set of PK equations is selected that can generate most of the identified time-course patterns. A discretized, validated model of the equations in software is executed providing simulated output. Metrics specify the goodness of fit of the simulated output to the data. **Right**: The abstract mechanistic description is more knowledge-based. We start by coming up with a set of mechanisms that, when executed, produce simulated dynamics that are acceptably similar to experimental observations. Abstract, software components are designed, coded, verified, assembled and connected, guided by that mechanistic description. The product of the process is a collection of mechanisms rendered in software. A clear mapping—C—is intended to exist between components and how they plug together, and hepatic physiological and microanatomic details. Relative similarity is controlled in part by parameterizations. Compilation and source code execution gives rise to a working analogue. Its dynamics (mapping B) are intended to represent abstractly corresponding dynamics (believed to occur) within the liver during an experiment. Measures of simulated dynamics provide time series data that are intended to mimic counterpart measures of wet-lab perfusion experiments. Achieving measurable similarities enables mapping A to be made concrete.

2.2.1. Caco-2 cell monolayer culture to study vectorial transport

The Caco-2 monolayer is widely used across the pharmaceutical industry as an *in vitro* model of the human small intestinal mucosa to predict the absorption of orally administered drugs (Meunier et al., 1995). Caco-2 cells, when grown on polycarbonate filters, form confluent monolayers that can be used as an *in vitro* model of the intestinal mucosa. By measuring the permeability of a compound across these cell monolayers, one can estimate the extent of its permeation through the intestinal mucosa. As oral administration of drugs remains the most acceptable route of administration, models for intestinal permeation and *in vitro* permeability assays remain a valuable tool of screening scientists for lead compound optimization (Press and di Grandi, 2008).

The Caco-2 cells are derived from a colon carcinoma. When cultured under specific conditions, the cells become differentiated and polarized such that their phenotype, morphologically and functionally, resembles the enterocytes lining the small intestine (Hildago 1989). Caco-2 cells express tight junctions, microvilli, and a number of enzymes and transporters that are characteristic of such enterocytes, such as peptidases, esterases, P-glycoprotein, and uptake transporters for amino acids, bile acids and carboxylic acids. Caco-2 cells are most commonly used not as individual cells, but as a confluent monolayer on a cell culture insert filter (e.g., Transwell®). When cultured in this format, the cells differentiate to form a polarized epithelial cell monolayer that provides a physical and biochemical barrier to the passage of ions and small molecules (Gao et al., 2000).

The procedure to set up a Caco-2 monolayer transport experiment can be summarized as follows and as depicted in Fig. 2.2. A filter is placed in a Petri dish and Caco-2 cells are seeded onto the filter and incubated for 3-4 weeks at 37 ºC until the cells are fully differentiated and polarized and the proteins are fully expressed. The cells are grown in media that provide nutrients;

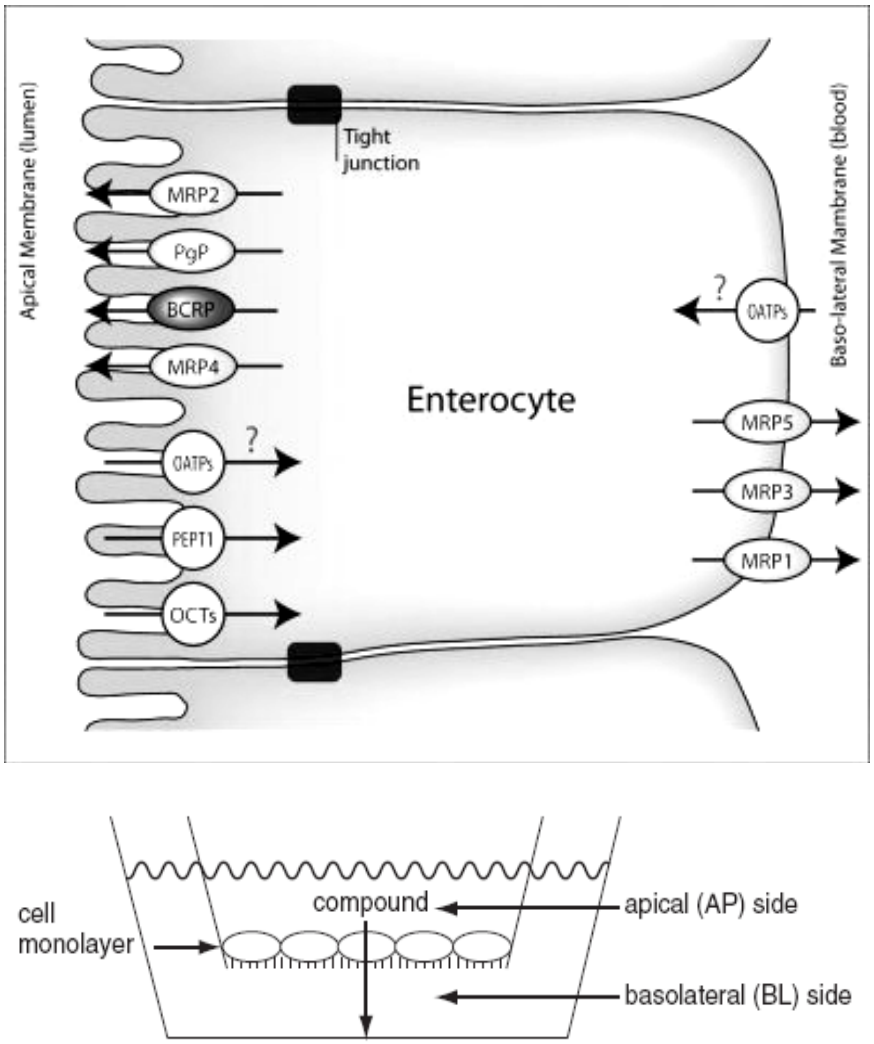


Figure 2.2 Top: Schematic of a Caco-2 cell. Caco-2 cells form tight junctions with each other and thus form a permeability barrier to drugs. They express various transporters. **Bottom:** Schematic of Caco-2 cell monolayer on a Transwell device. Caco-2 cells are grown to a confluent monolayer on a permeable filter. The layer of cell separates the medium into two compartments, the apical compartment and the basolateral compartment. Drug is dosed in one compartment and transport across the cell monolayer is assessed.

additives may be added to the medium to stimulate or modify expression of particular enzymes or transporters. For example, Caco-2 cells may be induced to express higher levels of CYP3A4 by treatment with vitamin D_3 (van Breemen and Li, 2005). When a confluent monolayer of Caco-2 cells is formed on the filter, the monolayer is ready for experiments. Typically, medium in the apical compartment is spiked with test compound, and the aliquots from the basolateral

compartment are taken to assay for the appearance of test compound over the duration of the experiments, which usually runs between 1 to 3 hours.

The correlation between the *in vitro* apparent permeability (P_{app}) across Caco-2 monolayers and the *in vivo* fraction absorbed is well established (Artursson and Karlsson, 1991; Hildago, 2001). The apparent permeability coefficient (P_{app}) of a test compound can be determined using the data generated from the transport experiment, and the magnitude of this P_{app} value is a reflection of how well the compound permeates cell monolayers.

$$P_{app} = \left(\frac{\Delta Q}{\Delta t} \right) \frac{1}{A \times C_0}$$

where P_{app} is the apparent permeability; ($\Delta Q/\Delta t$) is the linear appearance rate (flux) of the test compound; A is the surface area of the cell monolayer; and C_0 is the initial concentration of the test compound in the donor compartment.

Importantly, the Caco-2 monolayer Transwell system can be used to assess the effect of transporters. The test compound can either be dosed from the apical compartment and the appearance of the test compound in the basolateral compartment is monitored, or it can be dosed in the basolateral side and the appearance in apical side is monitored. If the permeability of a test compound is significantly higher for basal to apical transport than for apical to basal transport, the compound (for example, taxol) could be a substrate for an efflux transporter, for example, P-glycoprotein.

Other important applications of Caco-2 cells monolayer have been described in the literature. Examples included constructing structure-activity relationships of carrier mediated drug transport, assessing potential drug interactions, elucidating pathways for drug transport, metabolism, and interactions, and determining optimal characteristics for drug permeation and drug formulation (Gao et al, 2000; Sun et al., 2008).

The development of cell culture assays represents one of the most exciting advances in the pharmaceutical sciences (Borchardt, 1995; Artursson and Borchardt, 1997). The versatility of Caco-2 cells is demonstrated by the fact that, even to this day, they are serving as the basis for the creation of innovative new models that are contributing to our understanding of drug efflux transporters, such as P-glycoprotein (*ABCB1*) and breast cancer resistance protein (BCRP) (*ABCG2*). These types of systems, if properly used, can not only lead to an improved understanding of the biochemical basis of the barrier properties of the intestinal mucosa, but can also potentially expedite the process of drug discovery and development and thus improve the efficiency of the drug discovery/development process in the pharmaceutical industry. Meaningful refinements are made, such as inducing the expression of under-expressed transporters and enzymes to *in vivo* levels by cell biology methods, shortening culturing time, miniaturizing the cell culture apparatus, and automating the transport and metabolism experiments.

To better predict *in vivo* profiles accurately and efficiently utilizing *in vitro* data, methods to translate findings from *in vitro* assays to *in vivo* systems are needed (Hunt et al., 2009). The translation of information obtained from the Caco-2 monolayer to *in vivo* necessitates more sophisticated modeling approaches (Sun et al., 2008). Selection of scale-up factors with recognition of expression differences in transporters and enzymes between systems and integration with physiologically-based modeling approach are two examples of current challenges. On the other hand, using novel, agent-based modeling and simulation methods may help provide better mechanistic insight for translation (Liu and Hunt, 2005).

2.2.2. Isolated perfused rat liver

Liver perfusion was first described by Claude Bernard as early as the 1850s (Gores et al., 1986). An isolated perfused liver model with blood being the perfusate was used to identify roles of the liver in synthesis of plasma proteins (Miller et al., 1951). However, this technique was not extensively applied to drug metabolism and pharmacokinetics until the 1960s. Since then, the

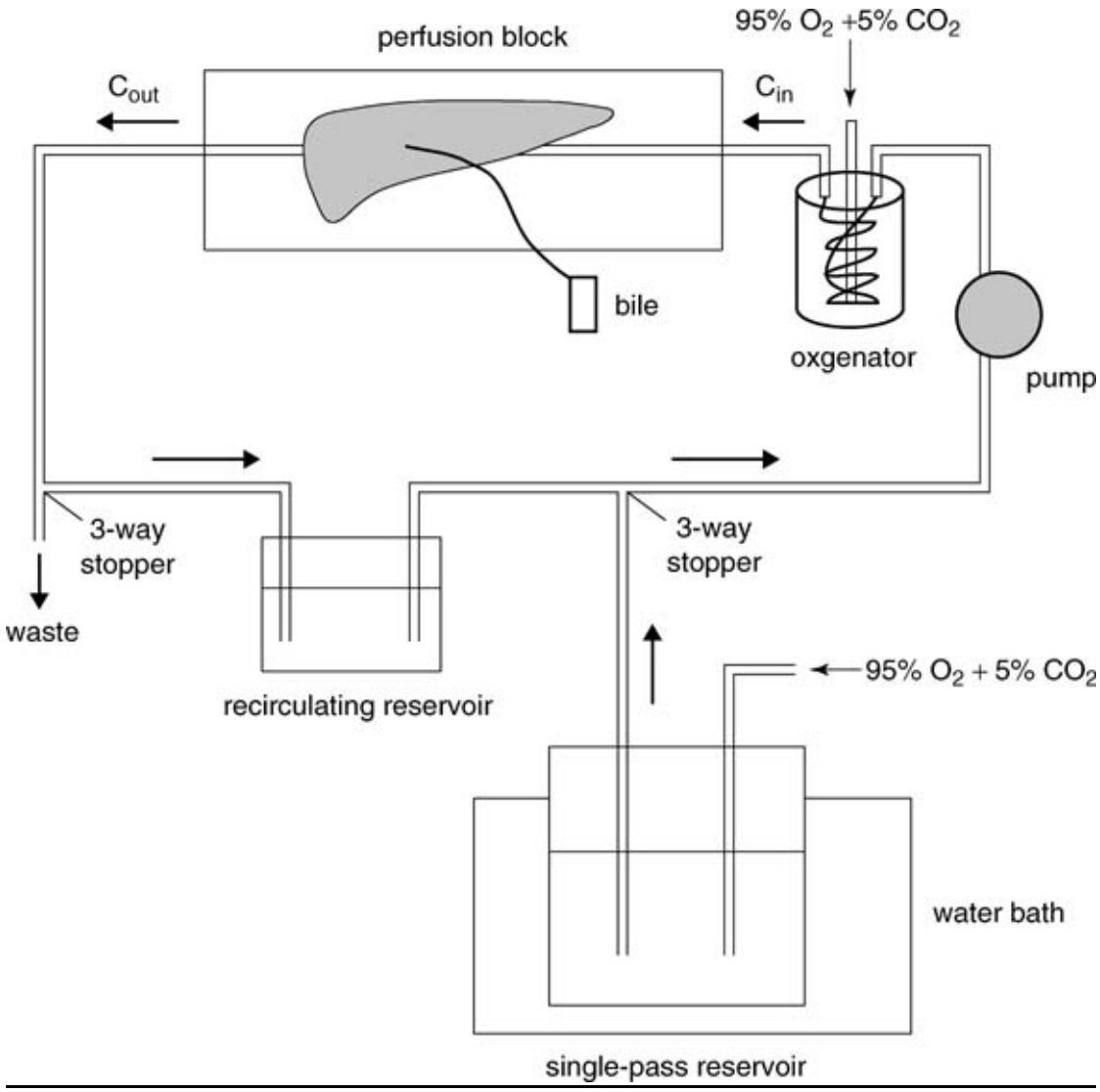


Figure 2.3 The schematic model for the isolated perfused rat liver. The liver is perfused with Krebs-Henseleit buffer at 37°C and saturated with 95% O_2 and 5% CO_2 gas mixture. Liver perfusion can be performed in single pass or recirculating by switching the 3-way stoppers. Bile is collected during the period of perfusion. C_{in} and C_{out} stand for concentrations in the inflow and outflow of perfusate, respectively. During recirculating experiments, samples of the perfusate are repeatedly taken from the reservoir.

isolated perfused rat liver (IPRL) has been extensively used as an intact organ model for determination of hepatic clearance and metabolism of drugs. The IPRL model can also be applied to determine physiologically based pharmacokinetics. An important advantage of using IPRL is that, in contrast to *in vivo* models, such as the bile fistula rat, the IPRL allows repeated sampling

of the perfusate and permits easy exposure of the liver to different concentrations of test substances. In addition, since the IPRL model avoids neural and hormonal interferences and excludes influences from absorption processes and non-hepatic elimination routes, such as renal excretion and respiration, it provides a relatively clean hepatic system to study metabolism and pharmacokinetics. It is especially useful to model the hepatic uptake associated with plasma protein binding and transport.

The basic perfusion system consists of a peristaltic pump, an oxgenator, and a perfusion block. The schematic diagram for an IPRL model is show in Fig 2.3.

The procedure of setting up an IPRL can be summarized as follows. The animal is anesthetized, and surgery is performed to expose the liver. The bile duct, hepatic vein, and portal vein are cannulated, and the liver is removed from the animal. The removed liver is maintained at 37 ºC. Perfusate, usually oxygenated Krebs-Henseleit buffer, is then pumped through the liver, entering from the portal vein and exiting from the hepatic vein. Bile flow and samples of the perfusate are collected during the perfusion for analyses (Liu et al., 2004). Key perfusion parameters are perfusate type, additives, and volume, perfusate oxygenation, perfusion flow and pressure, perfusion temperature, and perfusion duration (Wolkoff et al., 1987; Bessem et al., 2006).

Well-stirred and parallel-tube models are most commonly used to describe the elimination of drugs by the liver (Pang and Rowland, 1977). The equations are as follows:

$$Cl_H = \frac{Q \times f_u \times Cl_{int}}{Q + f_u \times Cl_{int}} \quad (\text{well-stirred model})$$

$$Cl_H = Q \times \left(1 - e^{-\frac{f_u \times Cl_{int}}{Q}}\right) \quad (\text{parrallel-tube model})$$

where f_u is the unbound fraction of drug in plasma; Cl_{int} is the intrinsic clearance of unbound drug; Q is the hepatic blood flow; and Cl_H is hepatic clearance of total drug.

Analysis of data from IPRL studies using the above models allows for estimation of hepatic extraction ratio, hepatic clearance, the intrinsic hepatic clearance, and fraction unbound. In addition, IPRL experiments are also useful in studying hepatic metabolism (Xiong et al., 2002; Yamada et al., 2008), hepatotoxicity (Yoshihara, 2000), protein synthesis (Liao et al., 1995), hepatic heterogeneity (Kato et al., 2001; Schwab et al., 2003), hepatic drug disposition (Liu et al., 2005), and influences of liver disease on drug disposition (Hung et al., 2005; Arab et al., 2007).

It must be recognized that the isolated perfused liver is a dying organ that has a life-span of approximately 3-4 hours (Gores et al., 1986). Viability assessments must be made to ensure hepatic function liver is not deteriorating faster than expected. Common measures include assessment of gross appearance, perfusate flow and pressure, bile flow, pH, lactate dehydrogenase activity, oxygen concentration, and potassium concentration. However, it is unrealistic to expect parts of the perfused liver and hepatic functions to deteriorate at the same rate. This may lead to intrinsic spatial and temporal heterogeneity in hepatic activities during the perfusion experiment, the effect of which is largely unappreciated and undocumented.

2.3. Traditional inductive pharmacokinetic models

Developing and then comparing the output of a pharmacokinetic model simulation to observed pharmacokinetic data involves multiple models. How they relate to each other, including the assumptions embedded in each, impacts the future usefulness of the parameterized pharmacokinetic model. To specify the differences and similarities between traditional pharmacokinetic models and the new class of synthetic models, one must identify and clearly describe the modeling method along with the variety of models used by each. I will expand on the descriptions from chapter 1 and begin with details of the traditional pharmacokinetic modeling process, focusing on key features and principles of the process.

After cogitating over the shape and appearance of the accumulated PK time series data, in light of available knowledge and expectations, the experimenter settles on an abstract, mechanistic description of the processes that is thought to be responsible for the data. The description may be novel. More often, it is one of several available descriptions in the literature in the form of a diagram-supported description in prose of, for example, the extended, convection-dispersion model or a tube model having equilibrating compartments. This model, induced from domain knowledge, is the *induced (mechanistic) model* being used to describe the data. There are aspects that the experimenter believes reflect influential events occurring within the biological tissue during the experiment. The left side of Fig. 2.1 illustrates that those ideas provide an abstract, conceptual mapping from the induced model to specific cellular dynamics. Because the mapping is from an intellectual concept (induced model) to real objects (parts of the liver and events within), the mapping cannot be made concrete (instantiated). Abstract, conceptual mappings require subjective interpretation and are difficult to falsify. (I argue in chapter 6 that hypothesis falsification, not validation, is a source of new knowledge.) To enable falsification, the hypothesis must be made concrete.

Typically, a goal of the pharmacokinetic experiment is to fit a pharmacokinetic model to the data. When the induced model has been taken from the literature, a corresponding set of equations will be available that describe possible time-series features of the model. An abstract, conceptual mapping will exist in the mind of the researcher between the mathematics and the experimental data.

The induced model cannot be executed directly. Execution following implementation requires that algorithms must be developed and coded. Implementation of the induced model, using a specific computer platform and compiler, generates an executable program. When executed, it is a working, concrete pharmacokinetic model thought to describe the experimental data. The aspect of the execution that is of interest to the experimenter is the measured states of

the model, i.e., the simulated profile. A concrete mapping exists between the simulated profile and the observed data. Consequently, a metric may be specified to measure the goodness of fit between the output of the implementation and the data. The parameter values assigned to the induced equation model are retained and reported.

When the researcher is satisfied with the goodness of fit, s/he is positioned to use the parameterized form of the equation model for whatever use was originally intended. Some parameters of some traditional induced pharmacokinetic models have been assigned descriptive names, such as clearance, mean transit time, and apparent volume of distribution. Fitted parameter values have proven to be useful summary descriptors of features of the pharmacokinetic data, providing scientists with a useful metric to compare features of different sets of data. However, there is no direct, concrete mapping between those valued parameters and particular features of the actual biological system during the experiment. Interpretations of those parameters are offered by a scientist. They draw and depend upon the scientist's mental model about the biology, along with the conceptual mapping from features of the induced model to the biological mechanisms.

In the following sections I present four types of frequently used equation-based pharmacokinetic models, from the most empirical sums of exponentials to the most mechanistic physiologically-based pharmacokinetic modeling.

2.3.1. Sum of exponentials

Linear pharmacokinetic data can be fitted to sums of exponential terms. For example, after an orally administered dose, a concentration-time profile with an absorption phase, an initial distribution phase, and a terminal elimination phase is observed as shown in Fig 2.4 below. One can fit a triexponential equation

$$C(t) = \mathrm{A}e^{-\alpha t} + \mathrm{B}e^{-\beta t} + \Gamma e^{-\gamma t}$$

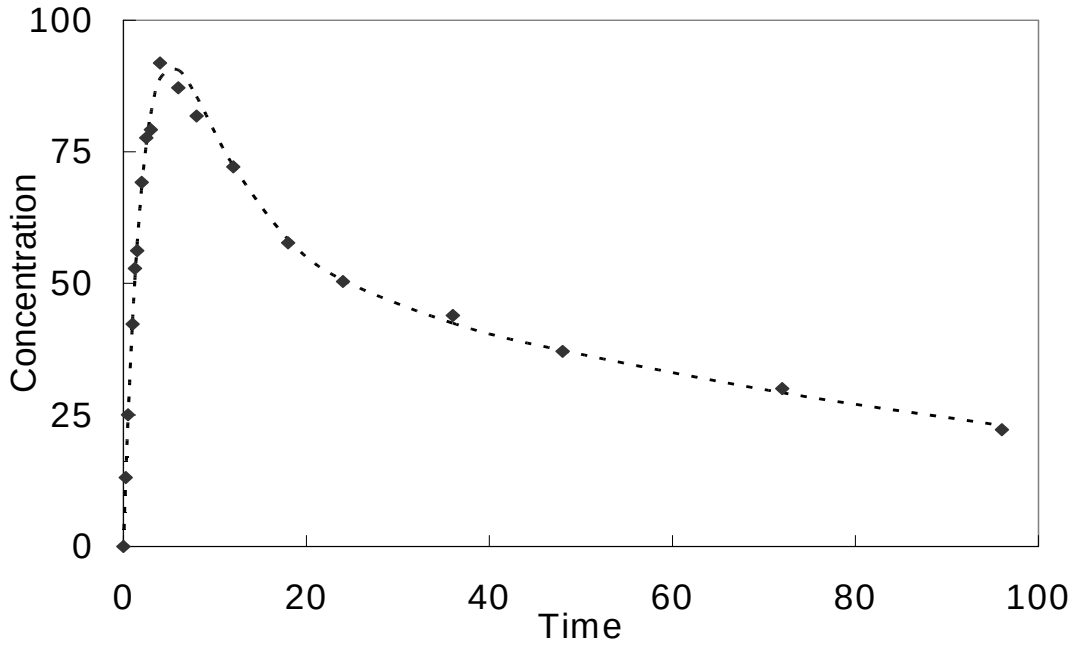


Figure 2.4 A typical time-concentration profile from a pharmacokinetic study. The profile showed an upward slope, followed by a peak, and a two-phase decline. One may fit a tri-exponential equation; the dotted line shows the equation $C(t) = 125*\exp(-0.15t) + 60*\exp(-0.01t) - 185*\exp(-0.4t)$.

Here, $C(t)$ in the concentration; $A, B, \Gamma, \alpha, \beta,$ and γ are parameters of the model and are selected to produce a best fit.

Fitted empirically, the equation is a model of data; it is a description of the observed data. Fitting to the equation, or the lack of fit per se, does not lead to validation or falsification, for there is nothing to be falsified, nothing to be tested. This model of data has no value in an experiment of a different kind, either. A "model of data" has value only as a simulator of the observed data; that is, it describes the data but contributes nothing to understanding. Its parameters carry no physiological meaning (Rescigno and Beck, 1987; Rescigno, 2004).

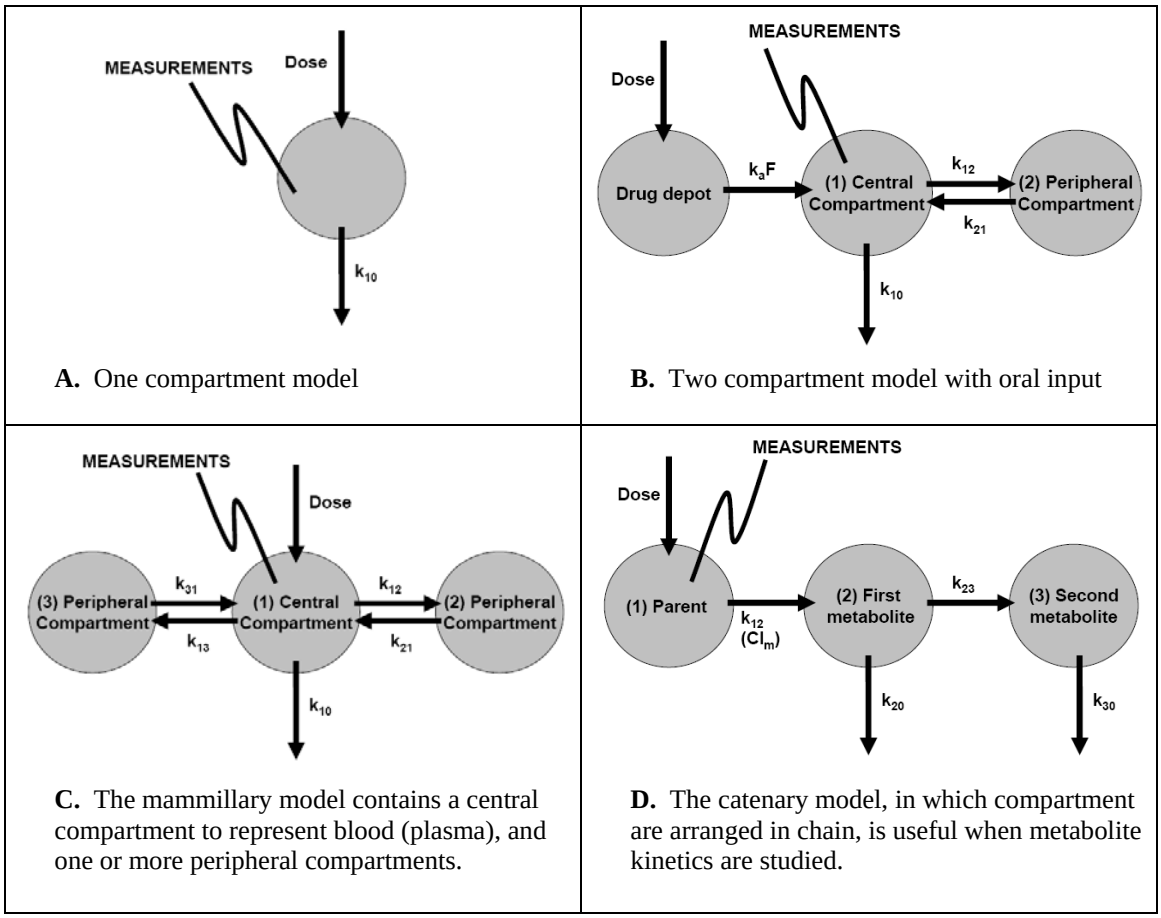
2.3.2. Compartmental models

In pharmacokinetics, a compartmental analysis involves the development of a model in which an administered drug is, conceptually, thought to enter into, exit from, and transit between

a finite number of homogeneous idealized “compartments;” this process gives rise to the observed time-course profile of the drug. Most commonly, a set of linear ordinary differential equations are used to describe specifically the input, exit, and transfer between compartments. The resulting pharmacokinetic model is a mathematical representation of the passage of the drug through the body and contains variables (e.g. dose, times of doses, and times at which blood samples are taken) and constants called parameters (e.g. clearance and volume of distribution) that quantify the drug disposition. In a pharmacokinetic model, the compartments may represent different sections of a body within which the concentration of a drug is assumed to be uniformly equal.

The compartment represents a defined volume for which there is no actual anatomical counterpart. To improve the mapping from conceptual compartmental model to the equations requires making the following assumptions: a) instant, homogeneous distribution of materials (drug) within a "compartment"; b) the exchange rate of materials among the compartments is related to the densities (concentrations) of these compartments; c) materials do not undergo changes while inside, or transmitting among the compartments, when materials undergo chemical reactions, the resultant materials are modeled as different compartments; and d) the volumes of the compartments are assumed to be constant over time, although this may not be totally true in biological systems.

The most common compartmental approach is to represent the whole body as a single homogenous compartment, which is termed the one-compartment model (Fig. 2.5A). Additional compartments may be added on in order to describe other patterns that may be seen in the observed concentration-time data. For example, a very slow elimination phase might be represented by a “deep tissue compartment,” which could map, conceptually, to adipose tissue. The resultant model is a two-compartment model. On the other hand, oral input of drugs requires an input compartment representing the gastrointestinal tract. The time-concentration profile,



A. One compartment model

B. Two compartment model with oral input

C. The mammillary model contains a central compartment to represent blood (plasma), and one or more peripheral compartments.

D. The catenary model, in which compartment are arranged in chain, is useful when metabolite kinetics are studied.

Figure 2.5 Four commonly used compartmental models

shown in Fig. 2.4, may be modeled with the three-compartmental model shown in Fig. 2.5B. Most compartmental approaches of this nature seldom have more than three to five distinct compartments. Two subgroups of compartmental models are of particular interest. The mammillary model (Fig. 2.5C) contains a central compartment to represent blood (plasma). Samples are taken from the central compartment while peripheral compartments equilibrate with the central compartment. The catenary model (Fig 2.5D), in which compartments are arranged in a chain, is useful when metabolite kinetics are studied.

The simplest one-compartmental model (Fig 2.5A) can be expressed as follows,

$$\frac{dX}{dt} = -k \times X$$

where *X(t)* is the quantity of material present in the compartment at time *t*, *k* is the first-order rate of elimination.

Dividing the amount *X* by the volume of the compartment *V* (assuming *V* is constant) yields an equivalent form in terms of concentration *C(t)*:

$$\frac{dX}{dt} = -k \times V \times C = -Cl \times C$$

where *V* (or V_d) is the *volume of distribution* and $Cl = kV_d$ is the *clearance*. Hence, volume of distribution V_d is a scaling factor between amount *X* and concentration *C*, and the clearance is the proportionality constant between rate of amount eliminated *dX/dt* and concentration *C*.

The integral of the two differential equations are

$$X(t) = X_0 \times e^{-kt}$$

$$C(t) = C_0 \times e^{-kt} = C_0 \times e^{-\left(\frac{Cl}{V}\right)t}$$

A general *n*-compartment model usually takes the form a of set of *n* linear ordinary differential equations:

$$\frac{dX_i}{dt} = -\sum_{j=0, j\neq i}^{n} k_{ij} X_i + \sum_{j=1, j\neq i}^{n} k_{ji} X_j$$

where X_i is the amount in the compartment *i*, k_{ij} is the transfer from compartment *i* to *j*, and k_{i0} is the elimination rate from compartment *i*.

The integral solution to this set of differential equations is in general sums of exponentials:

$$X_i(t) = \sum_{j=1}^{n} L_{ij} e^{-\lambda_{ij} t}$$

where L_{ij} and λ_{ij} are constants and depend upon the initial conditions and the parameters k_{ij} of the model.

For a general treatment of linear mammillary models, Benet (Benet, 1972; Nakashima and Benet, 1989) presented a general solution using general input and disposition functions, the methods of partial fractions in solving Laplace transforms and multiple dosing functions.

Importantly, this approach does not necessarily convey physiological meaning to the compartments, for example, the compartments represented by a 2-compartment model may not necessarily reflect any physiologically distinct tissues (e.g. the blood and the heart), but rather, they are representative of lumped tissues (e.g. all fast perfusion tissues).

The advantage of the mathematical pharmacokinetic models is that a compartmental or a physiologically-based pharmacokinetic model (below) has good descriptive, predictive, and prescriptive qualities (Rescigno, 2010). The models allow the concentration-time course to be described accurately in the subjects in the study. In addition, these approaches are generally independent of the size of the dose and dose interval, and therefore, concentration-time course in other individuals who may receive the drug in the future may be predicted. If the compartmental model contains mechanism-based parameters like clearance and bioavailability, and if its parameters can conceptually map to physiological variables, then most likely the concentration-time profiles may be able to be predicted in patients with different disease conditions or different administrations of dose. For example, for reduced clearance in renal disease patients or intravenous versus oral dosing, therapy may be tailored to achieve desired clinical goals.

2.3.3. Noncompartmental models

Noncompartmental analysis reduces the mapping assumptions that must be made in modeling concentration versus time data. An advantage of noncompartmental analysis is that it requires fewer assumptions than those necessary with compartmental analysis (DiStefano, 1982). While compartmental analysis requires knowledge (or assumptions) about the compartment into which the drug is inputted, the compartment from which the drug is eliminated, and the specific arrangement and transfer between compartments, noncompartmental analysis requires only that all measurement, sources, and sinks are from the one (central) pool. With that, there can be any number of recirculation or exchange between any numbers of pools, none of which has to be identified with any physiological structures, as illustrated in Fig. 2.6. It also avoids some of the common problems seen with compartmental analysis. For example, sometimes compartmental analysis would reveal the kinetics of the drug in some subjects is best described using a two-compartment model, while it is best described with a three-compartment model in other subjects.

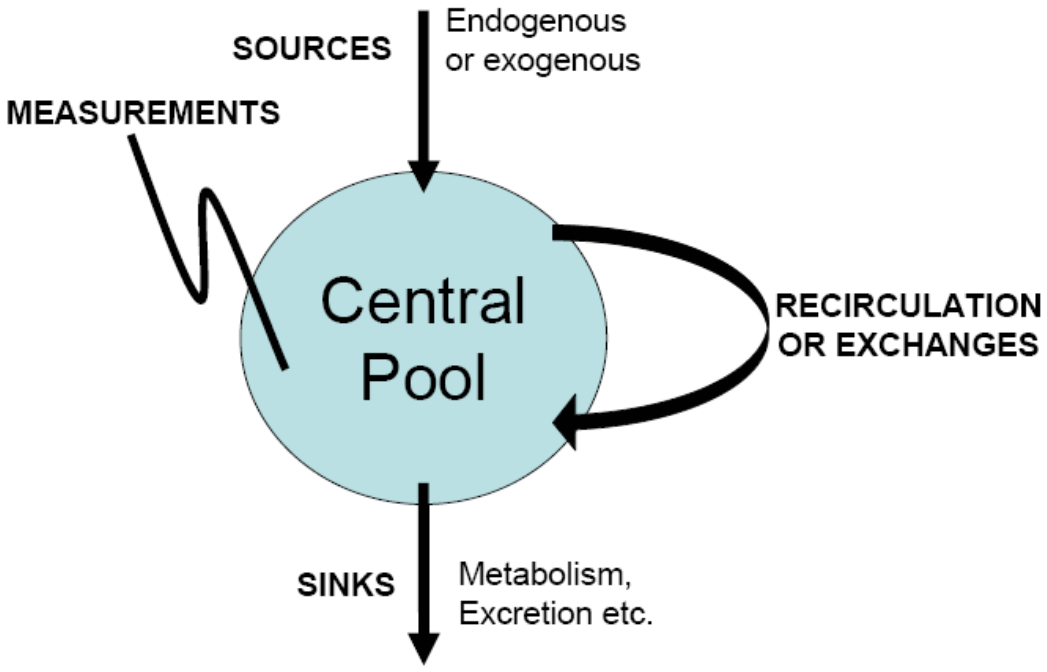


Figure 2.6 The simple structure assumed in noncompartmental analysis.

Should the data for all subjects be "forced" into a single model? Or should every subject be modeled based upon the fit of their individual plasma concentration-time data? These and related questions pose real problems in the use of compartmental analysis – problems that are largely avoided with noncompartmental analysis. However, it is important to realize that noncompartmental analysis is not "model-independent:" it required a specific model structure as described above. There are still assumptions that are made with noncompartmental analysis; for example, all sources and sinks are equivalent, which may or may not be correct. Often, the available data does not allow one to definitively determine the accuracy of the assumptions. The number and implication of those assumptions are, however, less than that seen with compartmental analysis (DiStefano and Landaw, 1984).

The most frequently derived quantities from noncompartmental analysis are the *total clearance* (*Cl*), *mean residence time* (*MRT*), and *volume at steady state* (V_{ss}). Their respective equations are given below.

$$Cl = \frac{D}{AUC} = \frac{D}{\int_0^\infty C(t)dt}$$

$$MRT = \frac{AUMC}{AUC} = \frac{\int_0^\infty t \times C(t)dt}{\int_0^\infty C(t)dt} = \frac{1}{K'_{el}}$$

$$V_{ss} = Cl \times MRT = \frac{D \times \int_0^\infty t \times C(t)dt}{\left(\int_0^\infty C(t)dt\right)^2}$$

where *D* is the dose given, *AUC* is the area under the concentration-time curve, *AUMC* is the area under the first moment of concentration-time curve, and K_{el}' is the apparent elimination rate. It should be noted that the *MRT* and V_{ss} equations are only true for intravenous bolus dosing.

2.3.4. Physiologically based pharmacokinetic models

Physiologically-based pharmacokinetic (PBPK) modeling is a mathematical modeling technique for predicting the absorption, distribution, metabolism, and excretion of a compound in humans and other animal species. (Nestorov, 2003) PBPK modeling is frequently used in pharmaceutical research and development and in health risk assessment and toxicokinetics.

PBPK models try to rely *a priori* on the anatomical and physiological structure of the body. These are usually also multi-compartment models, but the compartments correspond,

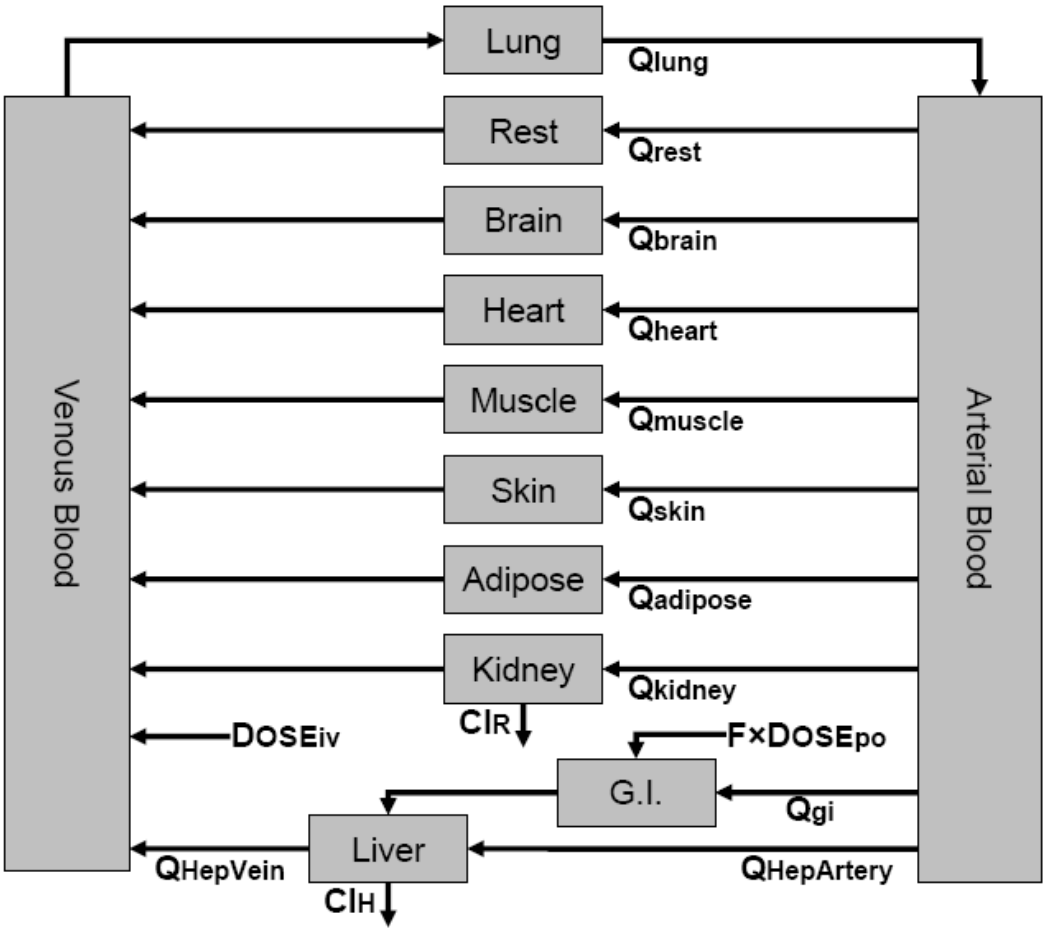


Figure 2.7 Schematic of a typical physiologically-based pharmacokinetic model. It is important to note that flow between compartments reflects physiologic blood flow. *F* is bioavailability, *DOSE*po is oral dose, *DOSE*iv is intravenous dose, *Q*'s represent blood flow, and *Cl*'s are the clearance of the respective organ.

conceptually, to predefined organs or tissues. The compartmental interconnections maps to movement of the compound of interest in physiologic blood (and lymph) flow. A system of differential equations is derived based on known anatomy and physiology, whereas its parameters represent blood flows, drug elimination (or absorption) rate at specific tissues, tissue volumes, and so on. Information for each of these parameters may be available in scientific publications. Often, the description of the body is simplified to strike a balance between model simplicity and the needed complexity sufficient for predefined model use. Besides the advantage of incorporating *a priori* information about parameter values, these models also facilitate inter-species transpositions or extrapolation from one mode of administration to another (e.g., intravenous to oral).

The PBPK models map the complex drug transport scheme onto a physiologically realistic compartmental structure (Fig. 2.7). The major structural elements of those models are physiologically realistic body tissues, fluids, organs, and/or systems. In this sense, the structure of a PBPK model is predetermined and is largely independent of a particular drug of interest. Therefore, in contrast to the conventional pharmacokinetic model, the structure of a PBPK model is derived from the anatomical and physiological structure of the organism studied and not from the available drug-related data. This can be considered as the major distinctive feature to discriminate between the two model classes. (Nestorov, 2003)

The PBPK model is a special form of pharmacokinetic models that belongs to the more general set of the “compartmental” models. Similar to conventional compartmental pharmacokinetic models, PBPK models compartmentize the organism into a number of subunits, called compartments. Unlike conventional compartmental pharmacokinetic models, which are sometimes called “empirical” or “data based,” PBPK models are often stated to be “mechanism based,” meaning that they reflect aspects of our knowledge about the mechanisms of the underlying pharmacokinetic processes. In fact, neither type is purely empirical or mechanistic.

Even the simple one- and two-compartment models are based on mechanistic assumptions, e.g., that the drug is transported by the blood circulation system, eliminated from there or from a peripheral compartment, and so forth. Therefore, the distinction between the conventional and the PBPK models is made based on the relative predominance of mechanistic or empirical elements in the process when compartmentization is carried out. In PBPK models, the model structure is determined before the modeling process starts; in empirical compartmental models, it is determined based on the best fit to the experimental data.

Emulating the structure of the living organism studied, the whole body PBPK model is a connection of tissues, fluids, organs, and systems; the link, effecting the compound distribution, is the blood circulation (Fig. 2.7). Therefore, at the first modeling step, the tissue connectivity within the blood circulation loop is specified. With small variations, the overall structural scheme usually follows the anatomical whole body circulation structure common to all mammalian species. The major model structure decision is to decide which body tissues, fluids, organs, and systems to include as components of the PBPK model. (Nestorov, 2003)

As usual, PBPK model specification should compromise between at least two conflicting driving forces: on one hand, the need to include tissues and organs that are essential for the pharmacokinetics and pharmacodynamics of the compound, for disease progression, as well as drug administration, and on the other hand, the need to avoid unnecessary complexity, computational intractability, and assumptions due to lack of data and information.

Generally, the following body tissues, fluids, organs, and systems are included in a PBPK model:

- Blood, subdivided between the venous and arterial pools.
- Main eliminating organs, i.e., liver for metabolized drugs and kidney for excreted drugs.
- Adipose tissue, for lipophilic compounds.

- Tissues that are expected to account for significant portion of the drug amount, e.g. muscle.
- Other tissues of special interest, including
    - Site of drug administration, e.g. skin for subcutaneous administration.
    - Other eliminating tissues, such as lung and intestine.
    - Potential sites of drug action.

For each of the tissues, in the majority of cases, the simplest perfusion rate-limited tissue model, which represents each tissue or organ by a single well-stirred compartment, is assumed. The underlying mechanistic assumption in this case is that the drug distributes instantly in the whole volume of the tissue from the incoming blood flow, and there are no drug concentration gradients within the particular tissue. In some cases, the more complex permeability rate-limited tissue model is assumed. Two particularly frequent models are 1) the vascular and extravascular compartments, in which permeation of capillary membrane is rate-limiting, or 2) the extracellular and intracellular compartments, in which permeation of cellular membrane is rate-limiting. Finally, one may need to assume that concentration gradients exist and a dispersion tissue model is used.

Once the PBPK model equations are written, their parameters need to be specified and/or estimated. The PBPK models have two groups of parameters. Physiological parameters are the parameters characterizing the anatomical structure and physiological processes of the animal species researched. Examples include bodyweight, tissue volumes, and tissue blood flows. Because the physiological parameters most often are assumed to be independent of the compound administered, they are also known as “drug-independent parameters.” These parameters are commonly extracted from sources independent of the particular pharmacokinetic study and are considered to be fixed, meaning that they are not estimated. Compound-specific parameters, on the other hand, are parameters characterizing processes such as binding, partition, permeability,

and metabolism. Sometimes, part or all of those parameters are known from previous studies or other prior knowledge. Most often, however, the aim of the PBPK study is to obtain the values of some compound-specific parameters. In such cases, the latter are estimated from appropriate experimental data.

However, one should keep in mind that the initial “drug-independent” PBPK model, using literature, organ size, and blood flow values as its parameter values, is a concrete hypothesis (or a collection of assumptions) about the referent anatomy and physiology that is rarely, if ever, challenged. The referent mammal will differ from this idealized PBPK model in several ways. Thus, the resultant “drug-specific” parameterizations from data-fitting procedures necessarily conflate two things: 1) the behavioral phenotype of the drug in the idealized PBPK model, and 2) the unknown differences between the referent organism's actual physiological features and those specified for that initial idealized PBPK model.

PBPK modeling is thought to have an extended domain of applicability compared to that of classical empirical function-based, pharmacokinetic models, especially in pharmaceutical research and toxicity risk assessments. (Nestorov, 2003) However, there are important barriers to seeing PBPK modeling developed and implemented to its full potential. First, developing a PBPK model requires a large amount of information, both precise data and extensive knowledge, that is not always available, especially during the early drug development phase. Second, PBPK modeling requires complex and multidimensional model structures, and is thus methodologically, numerically, and computationally much more demanding compared with conventional empirical pharmacokinetic modeling. As a consequence, the successful development and implementation of a PBPK model is seen to require the investment of significant experience, effort, time, and resources.

2.4. Synthetic analogues as in silico pharmacokinetic systems

I now summarize the key features and principles used in developing and instantiating a *synthetic*, mechanistic model of a referent pharmacokinetic system. Two detailed implementations, the recirculating in silico livers (RISL) and the in silico experimental Caco-2 cell monolayers (ISECC), are presented in detail in the following chapters.

The process of building synthetic analogues to referent pharmacokinetic systems is generalizable. The process is illustrated on the right side of Fig. 2.1. Because all models designed to have extended life cycles are usage oriented, an essential first step is to specify the planned uses of the model. That includes specifying a set of model-related capabilities that the analogue system would be expected to exhibit; Table 2.1 is one example. These capabilities are important because synthetic analogues will be used and refined iteratively. In addition, an initial set of targeted, phenotypic attributes is specified that the analogue is expected to exhibit and possibly mimic. Initially, that set may simply include having components that map qualitatively to specific anatomical features and measures of the simulation that map quantitatively to observed profile levels of a drug and its metabolite(s). The expectation is that the set of targeted attributes can be expanded and the analogue can be iteratively revised so that its new phenotype includes the expanded set of attributes.

The next task is to specify an initial level of mechanistic resolution. It should be no more detailed than is needed to achieve the initially targeted attributes. An *abstract mechanistic description* of the system is created that is likely different from an inductive model such as a PBPK model. This initial inductive process is similar to that for traditional PBPK modeling, except that current knowledge about the experimental system and biological tissue and the available data play an even more prominent role in providing detailed specification of that mechanistic description.

Table 2.1 List of capabilities that synthetic analogues are expected to exhibit

Analogue capabilities needed to bridge the gap in Fig 1.1

As engineered devices, what characteristics must a gap-spanning synthetic analogue exhibit to achieve the vision presented in chapter 1? They must exhibit the eight capabilities listed below. The challenge is to design, build, and validate such analogues.

1. ***Transparency***: Simulation details, as they unfold, need to be visualizable, measurable, and comparable to those of referent wet-lab and animal systems.
2. ***Experimental indistinguishability***: Measurements of analogue attributes made during execution are, to wet-lab domain experts (at a comparable level of abstraction), indistinguishable experimentally from measurements made on a referent wet-lab model during experimentation.
3. ***Mappings***: Observables in silico, including components, are designed to be consistent with *in vitro* and *in vivo* observables. Doing so enables clear, iteratively concretizable mappings (Fig. 2.1; right side) between *in vitro* and in silico components and mechanisms.
4. ***Local mechanisms***: As in wet-lab systems, the simulated physiological behaviors and measurements that emerge during execution (Fig. 2.1, mapping A) are the consequences of local mechanisms—local component interactions. Components can be heterogeneous.
5. ***Reusability***: Analogue components can be made quasi-autonomous, when needed, and thus can be easily reconfigured to represent different mechanistic hypotheses, aspects of the referent wet-lab system, and experimental conditions.
6. ***Flexibility***: Because system aspects of interest will change, it must be relatively simple to increase or decrease the granularity of any analogue aspect in order to simulate an additional phenotypic attribute or change usage and assumptions, as new wet-lab observations become available or the space of targeted attributes expands.
7. ***Adaptability***: Analogous to how object-oriented programming makes software objects more adaptable, the components and the system itself should be constructed so that they can be adapted easily to function as components in other analogues.
8. ***Components articulate***: To study alternate mechanistic hypotheses and to better understand normal-to-disease transitions and differences between individuals, it must be easy to construct—to *synthesize*—alternative systems and transform them during simulations. Doing so requires that it be easy to join, disconnect, and replace components, both vertically and horizontally, without having to significantly reengineer the system.

Abstract software components, intended to represent the referent components at an appropriate level of granularity, are designed, coded, and verified, guided by the abstract mechanistic description. The product of the process is a collection of software components. A clear mapping is hypothesized to exist between these components and how they can be plugged together, and physiological and anatomical details in the referent. The relative similarity of those assembled components to wet-lab components and mechanisms is controlled in part by parameterizations. The stochastic nature of those parameterizations provides a built-in degree of uncertainty that can stand as a counterpart to the inherent uncertainty in repeated wet-lab experiments. The mapping from this *assembly of components* to the biological components and mechanisms is not conceptual; it can be made increasingly concrete by iterative refinements.

The components are plugged together by the researcher within a software framework that is also capable of managing the conduct of automated experiments. Two examples are SWARM package (www.swarm.org) and the MASON Multiagent Simulation Toolkit (cs.gmu.edu/~eclab/projects/mason/). The resulting *compiled source code* will be both framework and designer dependent. One cannot directly compare an executable model to data. One must take measurements from the executing model just as one does during execution of a liver perfusion experiment. The two sets of measures may then be compared. The recordable events occurring during execution are caused by the autonomous operation of the components. Together, they comprise a working implementation. Because the abstract mechanistic description is not intended to be a perfectly consistent theoretical construct, in contrast to a system of equations, we refer to its implementation as an *analogue* in order to emphasize that we will learn about its behavior largely through experimentation rather than theoretical inference.

The analogue is intended to represent the dynamics that are believed to occur within the liver during an experiment. Hypothesized mappings exist between analogue dynamics and the dynamics of hepatic components during a perfusion experiment. Additional mapping exists

between analogue dynamics and the unique dynamics and properties of the components representing drugs and metabolites, for example, which have been introduced during a simulated perfusion experiment.

Because most event details, as they unfold, are based on random number draws, simulation details are effectively nondeterministic, and can be viewed, recorded, and measured in a variety of ways. Measurements intended to mimic those carried out during the wet-lab experiment provide time series data. Analogous to measurements from experimental models, results from simulations are a filtrate of events that occurred within the hardware during execution of analogue. The expectation is that during validation, a subset of component parameterizations will be identified that lead to simulated aspects that are similar in measurable ways to the experiments. Metrics are then specified to measure the degree of similarity between simulated output and the wet-lab data. Those metrics provide a concrete mapping between them.

Note that the synthetic model is a concretized (instantiated) mechanistic hypothesis. Execution shows that it either does or does not function as expected, or that it either does or does not cause the emergence of phenomena that fall within a pre-specified range of referent phenomena. If it failed to produce the targeted phenomena, the mechanistic hypothesis can be rejected; otherwise, it can be subjected to further validation or falsification.

2.5. Summary

The computational pharmacokinetic models reported in the literature, traditional inductive, or synthetic, string together several distinctly different yet equally important types of models. Ideally, the specifics of each of the models and the assumptions on which each is based are made clear in order to facilitate comparisons among pharmacokinetic models and the inferences enabled by their use. Figure 2.1 shows how approaches of inductive and synthetic pharmacokinetic modeling follow a somewhat different process and use different methods. The list of capabilities desired of synthetic pharmacokinetic models reflects the different intended uses

of that class of models and their components, which focuses in part on gaining deeper insight iteratively into the relevant causal mechanistic details underlying and accounting for the unique, individual-specific, pharmacokinetic and pharmacodynamic properties of different drugs.

3. Relational Analogues of Pharmacokinetic Systems

3.1. Introduction

In this chapter I detail my effort in using the synthetic modeling and simulation method to build In Silico Pharmacokinetic Systems that are analogues of biological components in common pharmacokinetic experiments. The synthetic modeling and simulation method provides a means of developing a scientific, experimental approach to unraveling and understanding some of the complexities of pharmacokinetic phenomena observed in experiments. I plugged together validated, quasi-autonomous software components to form abstract, but mechanistically realistic analogues of wet-lab experimental into which one could add objects representing drugs, alone or in combination. It is noteworthy that when I started this work, there were no methods or protocols on how to proceed. I was moving into uncharted scientific territory. This chapter presents implementation details and logic, as well as evidence of verification and validation. They provide foundations on which the Recirculating In Silico Liver (RISL) and the In Silico Experimental Caco-2 Cell Monolayer Cultures (ISECC) are built.

To distinguish clearly in silico components and processes from corresponding biological components and processes, I use SMALL CAPS when referring to the former, and *italics* denote in silico analogues' parameters, variables and internal states.

It should be emphasized that I have not tried to describe how the specific, detailed events responsible, at multiple levels, for the pharmacokinetic data observed occurred within those referent experiments. That cannot be done: the precise knowledge needed to do so is simply not available. What I did is diagrammed on the right of Fig. 2.1 in chapter 2. I built an abstract, analogue systems in software, using autonomous components, with the expectation that measures of their behaviors, following the referent experimental protocols, would be sufficiently similar to the referent data so that in silico mechanisms could stand as a plausible hypothesis for what may have occurred during those experiments. Individual in silico mechanisms were not intended to be 1:1 physical or chemical descriptions of referent events (again, I lack the detailed knowledge). Because the in silico pharmacokinetic system is abstract, so must be its mechanisms. Rather, my intent has been that in silico mechanisms map logically and intuitively to referent events. For example, if it is known that a drug had to cross at least one biological barrier to access metabolic enzymes in order to undergo metabolism, then I specified that the same must be true within the in silico pharmacokinetic system. Logically, I also wanted to draw as extensively as possible from accepted knowledge and theory in building mechanisms, and I wanted my constructed objects to behave consistent with current knowledge and theory. For example, if it is known that a referent compound in its ionized state cannot cross biological barriers at physiological pH, then I expected the same behavior from an in silico COMPOUND. Initial component verification (the components works the way it has been designed to work) and face validation (component function based on performance during simulation is judged reasonable by domain experts) therefore consisted of providing evidence that the in silico component (and its mechanism) exhibit behavior consistent with current information and knowledge. The efforts below built on earlier verification and

validation (against wet-lab data) experiments (Liu and Hunt, 2006), (Garmire et al., 2007), and I incorporate those results by reference.

In addition, I make the case for building these analogues as internally relationally grounded. So doing was a new undertaking. The components of biological mechanisms are grounded to each other. The grounding of cellular components to each other and their environment is independent of any measures. From that fact, I can infer that analogues that bridge the gap in Fig 1.1 in chapter 1 will exhibit similar grounding. We measure wet-lab phenomena using metric devices. We cannot use those same devices to measure events during simulations.

A relationally-grounded model is one in which events, parameters and measurements are always relative to other components within the model; they are free of any referent system dimensions and independent of referent experimental data. The components and processes in synthetic models need not have units assigned. Instead, in relational grounding each constituent is grounded to a proper subset of other constituents. Relational grounding enables synthesizing flexible, easily adaptable analogues; it avoids the complications to combine metric space-grounded components and models.

However, a separate mapping model is needed to relate analogue to referent phenotypic attributes. An in silico experimental protocol is devised to take measurements during simulations and to map them to referent experimental data. A separate method is required to map in silico parameters to biological counterparts. Different wet-lab experimental protocols require different in silico experimental protocol that can use the same relational analogue. For verification and validation purposes, that requires that a separate scaling mapping to relate in silico events to referent events. Doing so required the additional verification experiments detailed in this chapter.

3.2. Objective and approach

My objective was to devise a fully relational analogue that would represent a single parameterized set of multi-level mechanisms that together would provide one plausible explanation for the observations in the two referent experiments (Lau et al., 2004) and (Mouly et al., 2004), while achieving a stringent measure of similarity. I sought an actual, relational and experiment-independent working mechanism—*an analogue*—comprised of quasi-autonomous processes and components, as well as local interactions between them. When measured during execution, the analogue would give time course data similar to wet-lab data presented in referent experiments.

Traditional PK modeling uses continuous macroscopic descriptions such as concentration and reaction rates. I parameterized processes using unit-less, quantities such as fraction of dose and probabilities of events. The following provides detailed description of model structure, each of its components, parameters, internal logic and implementation. I highlight the relational features in my approach.

During analogue design, a focus was also on simplicity. I adopted a parsimonious standard modeling and simulation guideline: always prefer the simpler mechanism until evidence is obtained that falsifies the simpler mechanism. Consequently, I thoroughly explored parameterizations for the simplest case before making the mechanism more complicated.

Take the RISL as an example. How simple or complex must a device be to represent the liver during recirculating perfusion experiments, when sampling times are several minutes apart? The answer depends on intended use and especially the referent data being simulated. Hunt et al. (2006) and Yan et al. (2007; 2008) describe an In Silico Liver comprised of similarly constructed LOBULES. Each LOBULE is comprised of several multi-level agents called Sinusoid Segments arranged into a network that forms three zones. That level of detail (system granularity) was supported by the referent wet-lab data against which their simulations validated. That data came

from single pass liver perfusion studies: outflow was collected and analyzed every few seconds. However, for experimental designs similar to the recalculating system used in the referent experiments of the RISL (Lau et al., 2004), more abstract, simplified representations have sufficed (Booth et al., 1996; Booth et al., 1998; Tirona and Pang, 1999; Matsuura et al., 2001), because rates of drug level change within lobular spaces are slowed when perfusate recalculates compared to single pass perfusions. I therefore inferred that neither zonation nor different Sinusoid Segment types would be needed to achieve my objective. Results of subsequent in silico experiments supported that inference. The resulting analogue has several features in common with the In Silico Transwell Device (Liu and Hunt, 2006; Garmire et al., 2007). The In Silico Liver cited above was designed to exhibit ten capabilities, similar to the ones listed in table 2.1 in chapter 2. The following five are especially relevant:

- It must be easy to *reconfigure* and *reuse* an analogue to represent different histological, physiological, or experimental conditions.
- In order to represent particular specifics of different experiments, it must be relatively simple to change analogue usage and assumptions, and increase or decrease detail, without requiring significant re-engineering.
- To facilitate the two preceding capabilities, it must be easy to join, disconnect, and replace in silico components: the components *articulate* easily.
- It must be straightforward to separately validate components, and that is facilitated by making simulation details visualizable and measurable. Consequently, the analogues must be *transparent.*
- The analogues must be usable for simulating the disposition, clearance, and metabolic properties of a wide variety of *compounds*, separately or in the same experiment.

3.3. Space and discretization

To represent physical space, I specified intuitive, toroidal grids of ELEMENTS, and identified each with a (x,y) coordinate. Each ELEMENT had eight neighboring ELEMENTS. The grid structure defined the analogues' relative spaces. I did not specify the location, shape, capacity, dimension or physical volume of each ELEMENT. Rather, they were all relative. Location of each ELEMENT was relative to its eight neighboring ELEMENTS, as well as stationery objects in the neighborhood. Shape was indefinite.

ELEMENTS can be made different from others within the same space, should the need arise. ELEMENTS can have different internal properties (e.g., simulating a more lipophilic or acidic region), enabling, for example, heterogeneous environment, such as the one implemented in ISECC.

The potential of discretization decisions and specifications to create artifacts that damage the model's ability to mimic the referent were taken into consideration during design and final specification of in silico components. The following are illustrative. The extracellular, intra-sinusoidal region is an infinitely divisible, continuous space. The same is true for the extracellular, apical compartment of a Transwell device. The path of a compound within would appear smooth at any achievable level of resolution. Within the corresponding in silico spaces, the spaces are represented using 2500 elements arranged in two-dimensional, toroidal grids having identical granularities: from one simulation cycle to the next, a COMPOUND'S location is fixed. The number of grid spaces and the time interval to which a simulation cycle maps needed to be specified so that the discretized space and movement have been sufficiently smoothed to avoid artifacts; such discretization decisions can impact the ability of the analogue to mimic the referent. Processes that are very fast (relative to the sampling interval), such partitioning and binding are treated as events, and so discretization of that process into an instantaneous event has no impact on the ability of the model to mimic the referent. METABOLISM and TRANSPORT are

somewhat slower and are believed to involve several steps. However, relative to the key events being simulated—the change in perfusate levels of COMPOUND and METABOLITE at intervals of minutes, they too can be treated as events, for which the detailed chemistry is ignored. Such discretization does not create any artifacts that damage the analogues' ability to mimic the referent. Importantly, it is straightforward to increase granularity (detail) when that is needed.

3.4. Selecting the number of Monte-Carlo simulations

Each Monte Carlo run (analogue execution) is an experiment. Early in the simulation

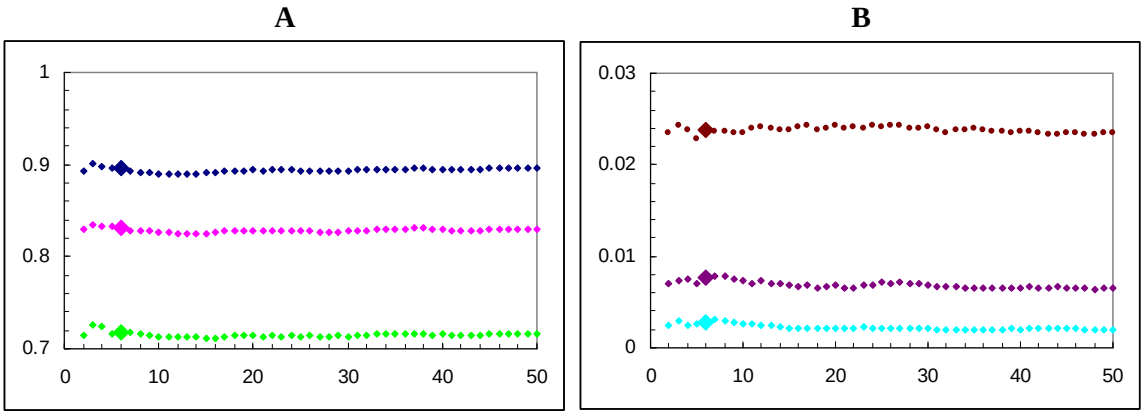


Figure 3.1. Cumulative averages of 50 experiments. Cumulative average of up to 50 repeated runs. **A** (from top): PERFUSATE DIGOXIN after 5, 10, and 20 simulation cycles, B (from bottom) PERFUSATE METABOLITE after 5, 10 and 20 simulation cycles. X-axis, number of repeated runs included in the cumulative average; Y-axis, cumulative average of fraction of administered DOSE. The graphs show that the average of only 6 runs (oversized data point) is an acceptable approximation of average of 50 runs.

exercise, I decided on the number needed to represent the mean analogue behavior for a given parameterization. Here, I show the data supporting my decision to use at least six Monte-Carlo experiments.

Using an early RISL, I ran 50 repeat experiments. For each, I made measurements of PERFUSATE DIGOXIN and METABOLITE levels at simulation cycles 5, 10, and 20. I then calculated their cumulative average. Results are presented in Fig. 3.1. They shows that the cumulative averages do not change significantly after n = 6.

3.5.Probabilistic parameter values

Most events, such as a COMPOUND binding to ENZYME, were probabilistic. When an event option arose, a pseudo-random number (PRN) was drawn from a uniform 0–1distribution. The PRN was compared to that of the probability parameter to decide what action to take. For example, if the PRN was less than the value assigned to *assocProb*, the event occurred. Else, it did not.

By drawing on the body of knowledge available for each event, I can select a likely initial range for each parameter value. For example, if a drug is known to have an above average affinity for CYP3A4, then I can infer the following. The probability of binding for any drug molecule that is in close proximity to CYP3A4 within a 10 second interval will likely be in the range 0.6–1.0. Absent any other information, I can start with a value of 0.8, for example, and then adjust it to improve the degree to which simulated and referent observables match. The width of the initially considered range contracts or expands based on prior knowledge or lack thereof. Because of the networked nature of the many independent, probabilistic events (the outcome at the end of each simulation cycle emerges as a property of them all), precise values of each parameter are not as critical as is often required with traditional, inductive, equation based models. A 5% change in the value of the key analogue probabilistic parameters did not typically produce a statistically significant change in a measured system level behavior, especially when the number of replicate simulations is kept small, six in this case (examples are provided in the section parameter sensitivities, and in Fig 3.6).

If uncertainty about the parameter's value is complete, I can simply assign a new random value for each event. So doing marginally increases simulation variance, but it also provides an objective means of dealing with uncertainties, which is an advantage of this modeling and simulation approach. By drawing on prior experience in validating similar analogues, several of

my initial parameter estimates proved adequate to achieve my objective. Because the reference data were unique, I elected not to undertake parameter fine-tuning.

By using a uniform distribution with variable frequency parameters throughout the simulation, I added uncertainty that maps to my ignorance about many of the details that are below my level of resolution. Assumptions about frequency values were separated from, and external to the analogues. Although PRNs were drawn from uniform distribution, the probabilistic parameter values controlled the frequency of event occurrence that mapped to biological reaction rates and biases.

3.6. Stationary objects: TRANSPORTERS, BINDERS, and ENZYMES

I designed and used three classes of stationary, active, quasi-autonomous objects. One mapped to subcellular processes responsible for transport. The second, ENZYME (or CYP) mapped to subcellular processes (that include CYP3A) responsible for drug metabolism. The third class was BINDERS; they mapped to anything within a cell or perfusate that might bind or sequester compounds. However, the details of TRANSPORTERS, BINDERS, and ENZYMES operations were below the level of resolution and so do not map to biochemical counterparts.

Each simulation cycle, each TRANSPORTER, BINDER, ENZYME follows specified logic to select the action or actions to take based on its local environment. Researchers have validated and used similar ENZYME and TRANSPORTER agents in different contexts (Liu and Hunt, 2006; Sheikh-Bahaei and Hunt, 2006; Garmire et al., 2007; Yan et al., 2008; Garmire and Hunt, 2008). CYPS can have multiple active and regulatory sites. However, for simplicity, the analogues RISLs and ISECCs do not use the latter. Each CYP scans each COMPOUND in its neighborhood to determine which have non-zero values of *assocProb*; for those that do, there is an opportunity during that same cycle for binding. A CYP can bind any COMPOUND that has a non-zero *assocProb*, whether it is a substrate or not. However, only substrates are METABOLIZED. Each CYP may bind multiple

COMPOUNDS; the maximum number is controlled by the parameter *maxSites*. Following that, each CYP has an opportunity to METABOLIZE a bound COMPOUND.

The binding sites within a CYP are identical and independent. However, details of CYP operations are below the level of mechanistic resolution and so do not map to biochemical counterparts. At the start of a simulation, all CYPS are assigned randomly to elements within S3 (intracellular space). Each simulation cycle, each CYP steps through its assigned logic to determine what action to take. The logic (diagrammed in Fig. 3.2B) and its use can be summarized as follows. A CYP examines its adjacent neighborhood. Neighborhood size is controlled by a parameter (*sitesN*). The probability that a CYP-DRUG binding event will occur for any DRUG within a CYP'S neighborhood is governed by *assocProb*, the value of which maps to affinity.

The logic used by TRANSPORTERS (PGP and OAT), diagrammed in Fig. 3.2A, can be summarized as follows. Each cycle, they first scan for COMPOUNDS in their neighborhood. Binding is next. A PGP or OAT, like CYP, can bind anything that has a non-zero *assocProb*. A PRN is drawn from a uniform [0-1) distribution. If its value is less than *assocProb*, the COMPOUND is bound. That process repeats for the next unoccupied site, and so on. The maximum number of bound COMPOUND is specified by the parameter *maxSites*. Then, they determine if each COMPOUND bound earlier will be released. That probability is controlled by the parameter *releaseProb*. For an inhibitor, that probability is small. When it is decided to release the bound COMPOUND, the PGP (or OAT) can release it into the space of origin, such as S3 (S1 for OAT), or into the destination adjacent space, such as S5 in RISLs, and S1 in ISECCs (S3 for OAT). However, only substrates are TRANSPORTED (RELEASED) to the destination space. The probability of releasing SUBSTRATE into the space of destination is controlled by (TRANSPORT) *efficiencyProb*.

Unlike traditional models, which parameterize binding and dissociation with affinity and

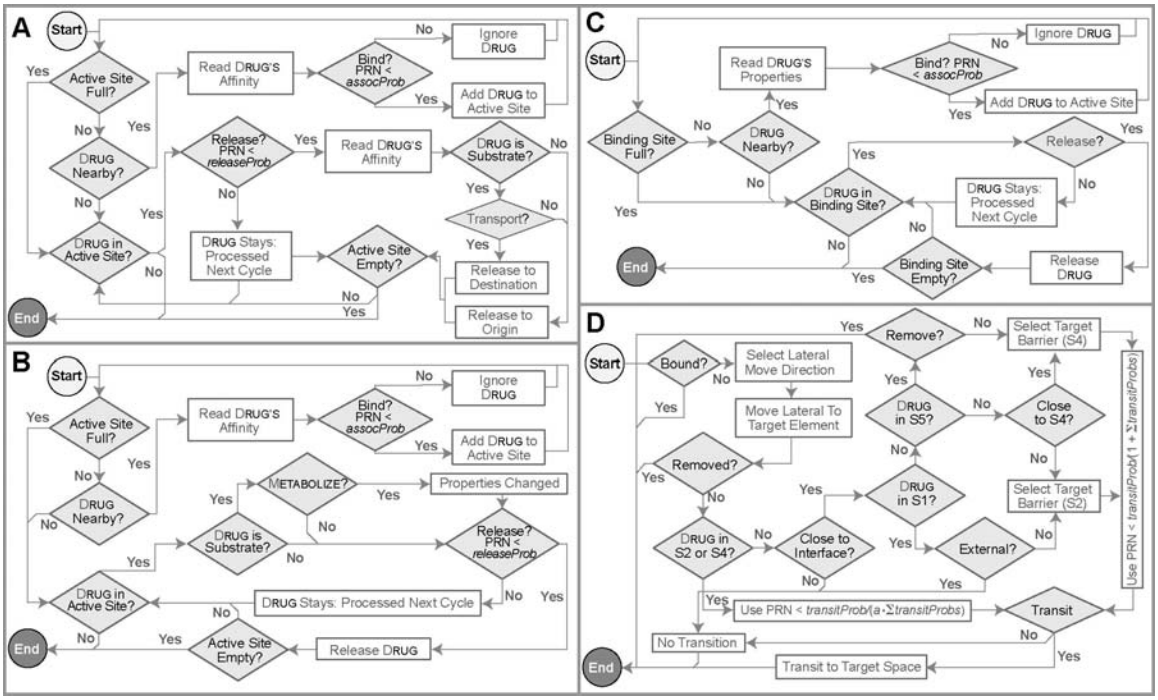


Figure 3.2. Internal logic used each cycle by TRANSPORTERS (**A**), ENZYMES (**B**), BINDERS (**C**) and COMPOUND (**D**). The internal logic of TRANSPORTERS, ENZYMES and BINDERS is similar. First, each decides whether it can bind A COMPOUND within its local neighborhood (neighborhood size is specified by *siteN*). BINDING is probabilistic. The value of *assocProb* maps to the compound's affinity for its binding partner. When a *PRN* < *assocProb* binding occurs. When multiple COMPOUNDS are within a neighborhood, the process continues until all COMPOUNDS have had one opportunity to bind. **A**: A TRANSPORTER first decides whether or not to release the bound COMPOUND. If *PRN* < *releaseProb*, it selects a location to place the COMPOUND. When the COMPOUND is a SUBSTRATE and *PRN* < *efficiencyProb* the COMPOUND is transported to a *destination* (across the barrier). Otherwise, the COMPOUND is released to the space from which it was bound. **B**: Each CYP first decides whether or not to METABOLIZE a bound COMPOUND. When *PRN* < *efficiencyProb* it is METABOLIZED. If *PRN* < *releaseProb*, the bound object may be released. **C**: A binder's internal logic is similar to that of a CYP, except there is no METABOLISM step. **D**: The logic used by mobile objects to move about is divided into two phases: 1) lateral movement within the same space, which simulates dispersion, and 2) transition to a neighboring space, representing passive permeation. A COMPOUND will use different transit rules depending on the space in which it is located. A COMPOUND bound to stationery objects is not given an opportunity to relocate during that cycle (even when it gets released during that cycle). In RISLs, COMPOUNDS that transition to BILE (S5) are not given an opportunity to return; they are REMOVED. In ISECCs, COMPOUNDS follow additional movement rules when being in S3 (intracellular); the rules are detailed in chapter 5.

dissociation constants expressed in concentration terms, which are invariably macroscopic, I implemented a microscopic, relational view. I held that molecular events were probabilistic, and molecules had no knowledge about the whole system. Local concentration was replaced by (amount of) substrate, which was in (close enough to) the stationary objects' local neighborhood, whose size (measured in ELEMENTS) was controlled by the parameter *sitesN*. Binding affinity mapped to probability of association (*assocProb*), the probability of having a binding event within the current time step when a COMPOUND was in the stationary object's neighborhood. Dissociation constant mapped to probability of release (*releaseProb*), the probability of release of a bound COMPOUND from a stationary object within the current time step. These two probability parameters governed BINDING and RELEASE: if PRN < *assocProb* (*releaseProb*), BINDING (RELEASE) event occurs, otherwise, there was no BINDING (RELEASE).

Similarly, I used the dimensionless parameter *efficiencyProb* to represent a dimensioned rate of reaction. The frequency at which METABOLISM occurred was controlled by this tunable parameter. Only compounds that were flagged substrate (*isaSubstrate*) of the CYP were METABOLIZED. If PRN < *efficiencyProb*, METABOLISM occurred, otherwise, there is no METABOLISM. For TRANSPORTER, *efficiencyProb* controls the frequency of substrate released into the space of destination. If PRN < *efficiencyProb*, the substrate is released to the destination space; otherwise it was returned to the original space. When TRANSPORTER found and bound a non-substrate, such as an INHIBITOR, it was always released in a subsequent simulation cycle to the space from which it was bound.

Although binding affinity and *assocProb* were related, a quantitative direct scaling between them, as well as between other analogous pairs, was outside the scope of this dissertation.

3.7. Probability scaling

Consider a probabilistic event *X*. The probability of its occurrence during a simulation cycle is $P_{old}(X) = 0.3$. Assume that one simulation cycle (the "old" cycle) maps to 1 second.

Suppose I desire to span more wet-lab time with one simulation cycle. I want one new simulation cycle to map to 4 seconds: the process is to be scaled up by a factor (*SF*) of 4. What is the new probability $P_{scaled}(X)$ for the same event within the new simulation cycle?

Increasing $P_{old}(X)$ by *SF* yields an incorrect result:

$$P_{scaled}(X) = 4 \times P_{old}(X) = 1.2 > 1$$

Instead, I allow four trials of the probabilistic event within one, new cycle. If the new simulation cycle maps to an event interval that is four times longer than with the original simulation cycle, then the original event is now being given four options to occur. After four options, the probability of $P_{scaled}(\text{not } X)$ is therefore

$$P_{scaled}(\sim X) = \left(1 - P_{old}(X)\right)^4$$

So,

$$\begin{aligned} P_{scaled}(X) &= 1 - P_{scaled}(\sim X) \\ &= 1 - \left(1 - P_{old}(X)\right)^4 = 1 - \left(1 - 0.3\right)^4 = 0.76 \end{aligned}$$

It is easiest to understand when *SF* is a positive integer. The same can be generalized to any positive number of scaling factor (*SF*).

$$P_{scaled}(X) = 1 - \left(1 - P_{old}(X)\right)^{SF}$$

For $SF > 1$, $P_{scaled}(X) > P_{old}(X)$, so the process is sped up; for $0 < SF < 1$, the $P_{scaled}(X) < P_{old}(X)$, so the process is slowed down.

The relationship is used to scale probabilities relative to the time to which a simulation cycle maps and other COMPOUND-PROTEIN interactions. The parameter *timestepfactor* is the factor used to scale all transit probability initial estimates.

The same relationship was used for time-variant parameters when I required that the probabilistic event of interest needed to "speed up" or "slow down." In RISL presented in chapter 4, the time-variant parameters *closeToInterface*, which specifies fraction *transitional*, and CYP'S *efficiencyProb* are both time-dependent. In the next chapter, fig. 4.4 shows them being scaled down from their initial value using the above relationship, with the value of *SF* decreasing.

3.8.Mobile objects and their movement

A COMPOUND (DRUG or METABOLITE) was a mobile, autonomous object that mapped to a small fraction of the drug added to perfusate (RISL) or dosing compartment (ISECC) at t = 0, or, in the case of METABOLITE, was generated during simulations. I included only objects that proved necessary to achieve my objective. Unlike traditional PK models, new components can be added without interfering with the function of those already present. As in previous reports (Hunt et al., 2006; Yan et al., 2008a) one COMPOUND mapped to a large but unspecified number of actual referent molecules. Amount of COMPOUND is expressed in terms of fraction of total COMPOUND used. Doing so avoided specifying an amount scale.

3.8.1. Passive dispersion

Each COMPOUND had a chance during each simulation cycle to disperse passively within its current space. In order to simplify implementation, within the same space, all ELEMENTS, each specified by a (x,y) coordinate, were treated as being identical, and having no particular shape, volume or absolute location. I used toroidal spaces to eliminate edge effects: a COMPOUND going out of a space to the right (or top) reappeared on the left (or bottom) and vice versa. During that simulation cycle, a COMPOUND in an ELEMENT has equal chance to moving to any one of its eight neighboring ELEMENTS, or to staying within its current ELEMENT. Doing so has enabled us to track specific COMPOUNDS, should that be desired for later verification or face validation purposes.

During a simulation cycle, each COMPOUND generates an integer PRN between 1-9, inclusive. A COMPOUND currently in ELEMENT (x, y) will move to the new location in the diagram below.

Coordinate	x – 1	x	x + 1
y – 1	PRN = 1	PRN = 2	PRN = 3
y	PRN = 4	PRN = 5	PRN = 6
y + 1	PRN = 7	PRN = 8	PRN = 9

An alternative implementation would have been to relocate each COMPOUND randomly during each simulation cycle. The end-result is the same. The disadvantage of the latter is that it would be more difficult to track a particular COMPOUND.

COMPOUND relocations are independent events. Each ELEMENT can hold any number of COMPOUNDS. When moving within a space, a COMPOUND does not consider the number of COMPOUNDS in the target ELEMENT. Because COMPOUNDS are placed randomly at the start of an experiment, there is no need for that additional complication.

Movement, or relocation, within a grid was relative to the location of stationery objects and their respective neighborhood. To adhere to the principle of local interactions, COMPOUND relocation depended upon local environment, and did not use global information, except for simulation time. Hence, unlike many other model approaches, I did not use a specific diffusion algorithm. Moreover, using a diffusion algorithm would attach a dimension scale to the spaces, and prevent them from being relationally grounded.

3.8.2. Passive transition

Each COMPOUND gets a chance to PARTITION passively into or out of a MEMBRANOUS space (S2 or S4) during each simulation cycle. Although numerous models are available to characterize *in vitro* passive permeability, none have been validated in a complex *in vivo* system such as the recirculating isolated perfused rat liver. Absent reliable values, and given the

relational analogue design, I needed a method to arrive logically at initial rates of passive transition. I formulated the following approach by adapting established physicochemical models. The following is a list of specifications made to enable arriving at initial estimates, which may or may not need to be adjusted subsequently during analogue tuning.

- In the referent system, a compound must be sufficiently close to the aqueous-membranous interface such that within the time interval corresponding to a simulation cycle the compound can reach the membrane (or aqueous space). In the analogues, I labeled each compound with a state variable transitional, and used a parameter (*closeToInterface*) to specify whether the compound was sufficiently close to an interface.
- Compound exists in multiple physicochemical states in aqueous media, some of which have greater intrinsic permeabilities. Being charged compromises permeability. When estimating initial transit probabilities for the analogues, I take into account the referent compound's expected differential permeability.
- Chemical structure (functional groups), together with composition and properties of the resident medium, such as pH, ionic strength etc, affects the relative abundance of physicochemical states, thus influencing overall permeability. In the analogues, I used parameter *insilicoPH* for the spaces as an analogous measure of the tendency of compound being less permeable (ionized) in that particular space. Each compound also has in silico *pKa*'s that reflect the referent compound's tendency to become less permeable (ionized) in particular media.
- Compound size also influences rates of permeation: the larger the compound, the less permeable. I used the referent compound's molecular weight as a surrogate measure of size in affecting permeation in the RISL.
- Within the time interval corresponding to a simulation cycle, compound within a small volume on either side of a permeable cell barrier will have equilibrated, all other factors

being held constant. As a result, the ratio of concentrations within the two adjacent spaces relates directly to the probability of finding a molecule in each space by the end of the next simulation cycle, which in turn gives the probability of transit during the cycle.

I use terms and methods such as the Henderson-Hasselbalch equation in order to draw as often as possible from established knowledge and accepted methods. I fully appreciate that applicability of these models have not been validated within complex biological systems (such as recirculating isolated rat liver) and that many of the measures (such as pH within of the intracellular matrix) are not well-defined and cannot be measured or validated during a perfusion experiment. Using these models and methods to make initial parameter estimates may handicap the analogues with unverifiable assumptions. However, I accepted that risk because I believe that I can limit risk by drawing on established theory as the logical first step. To date, the resulting initial estimates of transit probabilities have proven to be adequate for achieving my goal of implementing plausible mechanisms. Multiple approaches are feasible, and as part of future projects, I intend to explore alternatives. Garmire et al. (2007) presented an alternate approach, one that was designed not to be fully relational, yet it works well in the context of the model used.

As presented, I estimated probabilities of transiting from one space to another (*transitProb*) from physicochemical models, not from extrapolation from literature values of experimental permeability, which are scaled and experiment-dependent. An algorithm to adjust probabilities for time-step duration was implemented. All probabilities were relative to time-step duration. Here, using a relational movement design allowed us to change the model of estimating percolation and permeation probabilities independent of other analogue components. They can be separately validated and modified as needed. This important flexibility allowed us to reuse the same validated components in a different referent system.

3.8.3. Initial estimates of transit probabilities

The following is the physicochemical model used to obtain initial estimates of transit probabilities. These transit probabilities govern the likelihood of passive transition from one space to another when the COMPOUND is in the *transitional* state (specified as being close to interface).

Consider a multi-protic acidic compound A (H_xA) in space S, having pH = pH_S. Define

$$R_S = \frac{\text{fraction ionized}}{\text{fraction unionized}} = \frac{\left[A^{x-}\right] + \left[HA^{(x-1)-}\right] + \ldots + \left[H_{(x-1)}A^{-}\right]}{\left[H_xA\right]} = \sum_{k=0}^{x-1} \frac{\left[H_kA^{(x-k)-}\right]}{\left[H_xA\right]}$$

From the Henderson-Hasselbalch equation,

$$\frac{[HA^-]}{[H_xA]} = 10^{pKa_1 - pH_S} \text{ and } \frac{[A^{x-}]}{[H_xA]} = 10^{pKa_1 + pKa_2 + \ldots + pKa_{(x-1)} - (x-1)\,pH_S}$$

So,

$$R_S = 10^{(pKa_1 - pH_S)} + 10^{(pKa_1 + pKa_2 - 2\,pH_S)} + \ldots + 10^{(pKa_1 + pKa_2 + \ldots + pKa_{(x-1)} - (x-1)pH_S)}$$

$$R_S = \sum_{m=1}^{x-1} 10^{\left(\left(\sum_{n=1}^{m} pKa_n\right) - mpH_S\right)}$$

$$\text{Fraction unionized} = \frac{1}{R_S + 1} \quad \text{and} \quad \text{fraction ionized} = \frac{R_S}{R_S + 1}$$

Let P_N be the partition coefficient of the neutral or most permeable specie, H_xA.

Define the distribution coefficient D_S between a space containing a cell membrane M and aqueous space S,

$$D_S = \frac{\text{amount of all species of drug A in membrane space M}}{\text{amount of all species of drug A in aqueous space S}}$$

Assuming the partition coefficient of all ionized species to be kP_N, $k < 1$.

$$D_S = \frac{P_N\left(\frac{1}{R_S+1}\right)+kP_N\left(\frac{R_S}{R_S+1}\right)}{\left(\frac{1}{R_S+1}\right)+\left(\frac{R_S}{R_S+1}\right)} = P_N\left(\frac{1}{R_S+1}\right)+kP_N\left(\frac{R_S}{R_S+1}\right)$$

Analogously, for COMPOUND A, the in silico analogue of real drug A, I have analogous in silico parameters, *logP* (a measure of P_N), *pKa*, and *insilicoPH*. Using these in silico parameters reflecting physicochemical properties, I calculate an in silico D_S for COMPOUND A. In the analogues, I used $k = 0.001$.

Define *transit probability (transitProb)* from an AQUEOUS SPACE S (such as S1, S3, and S5 in Fig. 2) to a space M containing a MEMBRANE (such as S2 and S4 in Fig. 2), $p(\mathrm{S}\rightarrow\mathrm{M})$, be the probability that COMPOUND A transitions from S to M within the current simulation cycle.

Specify that the interval during which the referent transition would have actually occurred is smaller than a simulation cycle. Also specify that the in silico effective size of COMPOUND A is inversely proportional to the square-root of the value of the molecular weight property assigned to COMPOUND A ($\mathrm{MW_A}$).

$$P(\mathrm{S}\rightarrow\mathrm{M}) = \frac{D_S}{D_S+1}\times\frac{1}{\sqrt{\mathrm{MW_A}}}$$

For COMPOUND A in M flanked by two AQUEOUS SPACES S and S′,

$$P(\mathrm{M}\rightarrow\mathrm{S}) = \frac{D_{S'}}{D_S+D_{S'}+D_SD_{S'}}\times\frac{1}{\sqrt{\mathrm{MW_A}}}, \text{ and}$$

$$P(\mathrm{M}\rightarrow\mathrm{S}') = \frac{D_S}{D_S+D_{S'}+D_SD_{S'}}\times\frac{1}{\sqrt{\mathrm{MW_A}}}$$

These initial transit probability values may be adjusted (scaled) for use with simulation cycles that map to different wet-lab times by the parameter *timestepfactor*, using the probability scaling method described above. Within the RISL, *timestepfactor* = 2 (3.3 in ISECC).

$$P_{scaled} = 1 - (1 - P_{old})^{timestepfactor}$$

3.8.4. Relationship between COMPOUND (physicochemical) properties and initial transit probability estimates

I investigated the consequences of changing COMPOUND (physicochemical) property values on initial measures of transit probability to verify that transit probability estimates behaved consistent with theory. I used DIGOXIN, RIFAMPICIN, QUINIDINE and DIGOXIN METABOLITE (these were used in the RISLs). I adjusted one of their properties (for example, *MW*) over a range of values, while holding the others (*logP*, and *insilicoph*) constant. When conducting the experiments, I removed all transporters and time-variant processes. I measured initial permeability across a single barrier: the number of COMPOUNDS that reached S3 from S1 after 20 simulation cycles. The results presented in Fig. 3.3 are consistent with the above physicochemical models used to derive the initial transit probability estimates.

3.8.5. Movements of compounds

What is the probability that during a 10 second interval within a perfused liver, or within a Transwell system, that a compound will be close enough to an interface to transition (partition) passively from one space to another? An answer can only be an approximation. In the analogue, that probability is specified by the parameter *closetoInterface*. If close, the probability of actually transiting is specified by *transitProb*. After making several assumptions about microarchitectural details of cell monolayers, Garmire et al. (2007) used estimates of the relative size of cellular spaces and the known properties of compounds to estimate a compound's relative intracellular location and likelihood to partition. Doing so required making several assumptions. I elected not

to make those assumptions for two reasons: I lacked knowledge about specific cellular and lobular microarchitectural details and I wanted to make all analogue mechanisms relational. In addition, I have conflated numerous cellular and lobular microarchitectural features (that cannot be known with precision) and represented them using three abstract spaces. *ClosetoInterface* and *transitProb* are tunable parameters, the values of which are constrained by the referent compound's listed properties. Initial estimates were obtained for each referent compound using values of molecular weight, log*P*, and pK_a as described above, analogous to those used by Garmire et al. (2007). Thereafter, these parameters were tuned, along with the others, to improve similarity.

The sequence in which each mobile object's transit decision was made followed the same pattern for each space. The process is described first for a COMPOUND in barrier space S4 and then for a compound in CELL interior space S3 that is close to an interface. p[S4→S3] is the value of *transitProb* during a cycle for the transition from S4 to S3. First, a PRN is drawn from [0–1). If PRN < p[S4→S5]/(*a*(p[S4→S5] + p[S4→S3]), then the COMPOUND transits from S4 to S5; else, if PRN < 1/*a* ($a \geq 1$), it transits from S4 to S3; else, it stays in S4. Because the relative properties of S1, S3, and S5 in the analogues, as are the relative properties of S2 and S4, the logic is the same for S2. Following exploration, I specified $a = 1.2$ for RISL and 1.05 for ISECC; that value is arbitrary; using a value slightly larger than 1.0 helped ensure that a COMPOUND would not still be found in S2 or S4 after more than three cycles. Once it is determined that a COMPOUND in S3 is close to an interface (I do not specify which one), a PRN is generated: if PRN < p[S3→S4]/(1+ (p[S3→S4] + p[S3→S2])), the COMPOUND transits from S3 to S4; else, if PRN < (p[S3→S4] + p[S3→S2])/(1+ (p[S3→S4] + p[S3→S2])), it transits from S3 to S2; else, it stays in S3. The logic would be similar for a COMPOUND in S5 transiting to S4.

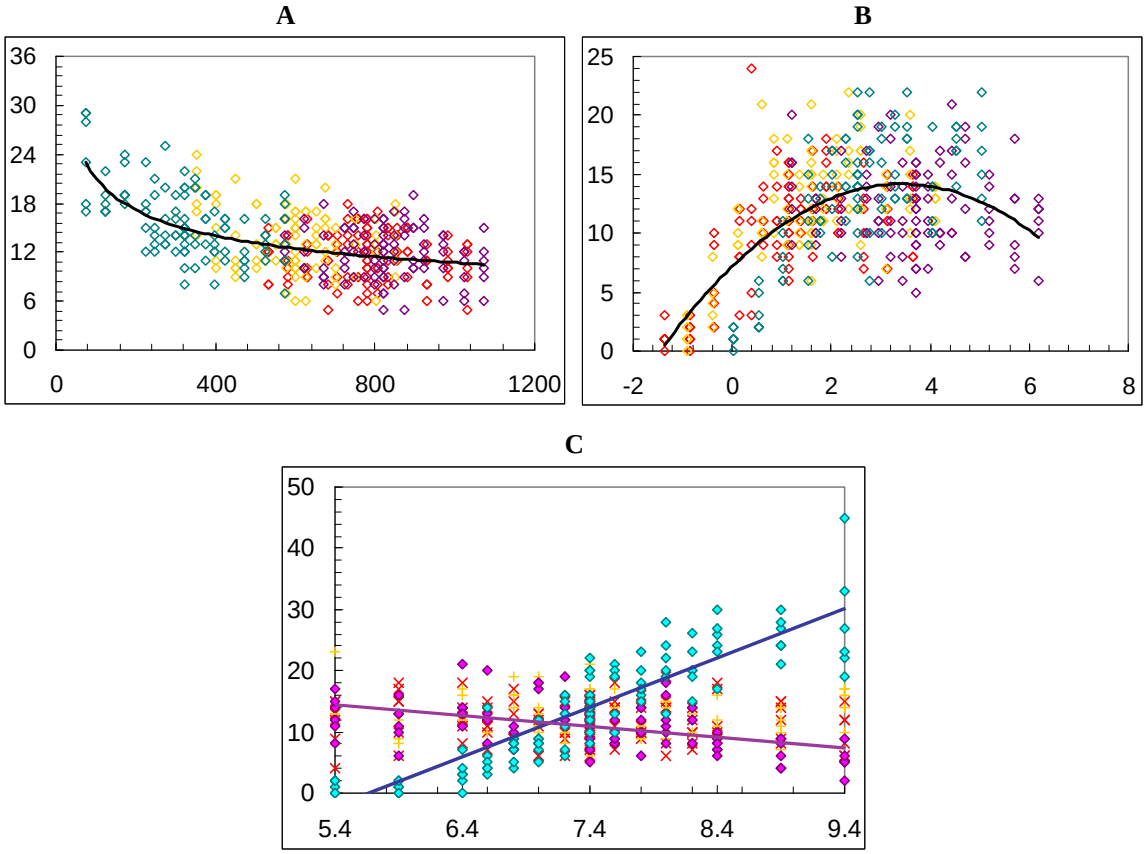


Figure 3.3. Relationship COMPOUND (physicochemical) properties and initial transit probability estimates. Each symbol is one RISL experiment. Y-axis: number of COMPOUNDs in S3 after 20 simulation cycles when the initial dose is 1,000. **A:** Graphed are relationships between *MW* (x-axis) and RISL measures of initial permeability. Symbols: QUINIDINE (light blue), METABOLITE (yellow), DIGOXIN (red), RIFAMPICIN (purple). The curve is a fit of an inverse power relationship to the pooled data. **B:** Graphed are relationships between *logP* (x-axis) and initial permeability (y-axis), as specified in **A**. Symbol colors are the same as in **A**. The curve is a parabola fit to the pooled data; the peak occurs at about *logP* = 3.3. **C:** Graphed are relationships between *insilicoPH* and initial permeability. Symbols: DIGOXIN (red crosses), METABOLITE (yellow pluses), QUINIDINE (light blue diamonds), RIFAMPICIN (purple diamonds). The two trend lines are for only the QUINIDINE and RIFAMPICIN data. QUNIDINE behaves as would a weak basic having pKa = 8.6: PERMEABILITY increases with increasing *insilicoPH*. RIFAMPICIN behaves as would an amphoteric compound having a weak acid group (pKa = 1.7) and a weak basic group (pKa = 7.9): PERMEABILITY decreases with increasing PH. DIGOXIN and its METABOLITE behave as would compounds for which the fraction ionized does not change over the *insilicoPH* range.

3.8.6. Simulated perfusate flow (for RISL only)

In perfusion experiments, perfusate is pumped through the organ by an experimental apparatus. From the experimenter's macroscopic perspective, perfusate, and content therein, moved. However, the compound within the perfusate has no knowledge about the perfusion experiment. From its perspective, its location is relative to objects inside or outside the liver. Consistent with that perspective, I assigned different states to each COMPOUND: 1) *external* (to the liver), 2) within extracellular hepatic spaces but not likely to transition across cell membranes within the time step (*internal*), and 3) within the liver and sufficiently close to cell membranes so that transition would be possible (*transitional*). Since each ELEMENT represented a fraction of the perfusate, with no specified location, perfusate flow was not simulated using movement within the S1 space. Instead, changing state simulated both perfusate flow and rapid equilibration between perfusate and hepatic spaces. The parameter *internalProb* specified the probability that a COMPOUND in S1 would be designated *internal* within a simulation cycle. The COMPOUNDS that were randomly assigned to S1 were simply specified as being in one of three states, *external*, *internal*, and *transitional*; and every COMPOUND in S1 had an opportunity to change its state each cycle. Changing *internalProb* simulated a change in perfusate flow rate. The parameter *closeToInterface* specified the probability that a COMPOUND in S1 would be *transitional* (available to transition to S2). The balance of COMPOUNDS in S1 was *external*.

I considered several other options to simulate flow other than the one described above. One option is to simulate flow within the space (e.g., S1) representing perfusate, analogous to how it was done in Hunt et al. (2006). S1 could be made circular and much larger than the HEPATIC spaces. COMPOUNDS could move around the circular space simulating flow. Transitioning into HEPATIC spaces would occur only in the portion of S1 that overlaps with the EXTRACELLULAR HEPATIC space. If I had simultaneous wet-lab measures of concentrations entering and exiting the liver against which to validate extraction, such a design would merit

consideration, but I do not. In fact, perfusate flow was sufficiently fast to make reliable measures of that type difficult. Having a circular S1 containing circulating COMPOUNDS does match the reality of the referent system, but it is not an essential feature for inclusion in an abstract analogue.

Another option would be to layer three spaces above S2. Call them S1.1, S1.2, and S1.3, with S1.3 being adjacent to S2. Only COMPOUND in S1.3 would have the option to transition to S2 within a simulation cycle. COMPOUND in S1.1 would map to compound that is in perfusate external to the liver. COMPOUND in S1.2 would map to compound that is internal, within rapidly equilibrating extracellular spaces. Adjusting the exchange rate between S1.1 and S1.2 would map to perfusate flow. For simplicity, I elected to conflate those three spaces: I represented the merged set using just one space, S1.

3.8.7. Transporter-mediated transition and enzyme-mediated metabolism

In Figs. 3.4A and B I present verification evidence demonstrating the linear relationship between TRANSPORT activity and number of TRANSPORTERS and the value of the SUBSTRATE'S *assocProb* (affinity). I present data in Fig. 3.4C and D that show the consequences of active transport with concurrent passive transition. Finally, in Fig. 3.5, the graphs show measures of CYP function are consistent with Michaelis-Menten kinetics.

3.9. Parameter sensitivities

A classical sensitivity analysis studies how output variation from a mathematical model can be apportioned, qualitatively or quantitatively, to different sources of variation in input. Such an analysis investigates model robustness when the study includes some form of mathematical modeling. The analogues are not classical mathematical models. The goals and objectives are different. Nevertheless, I did not want to include components that were not needed to achieve one or more attributes. To that end, as part of verification exercises, I measured the consequences of inactivating components. During the many tuning experiments, I also observed the

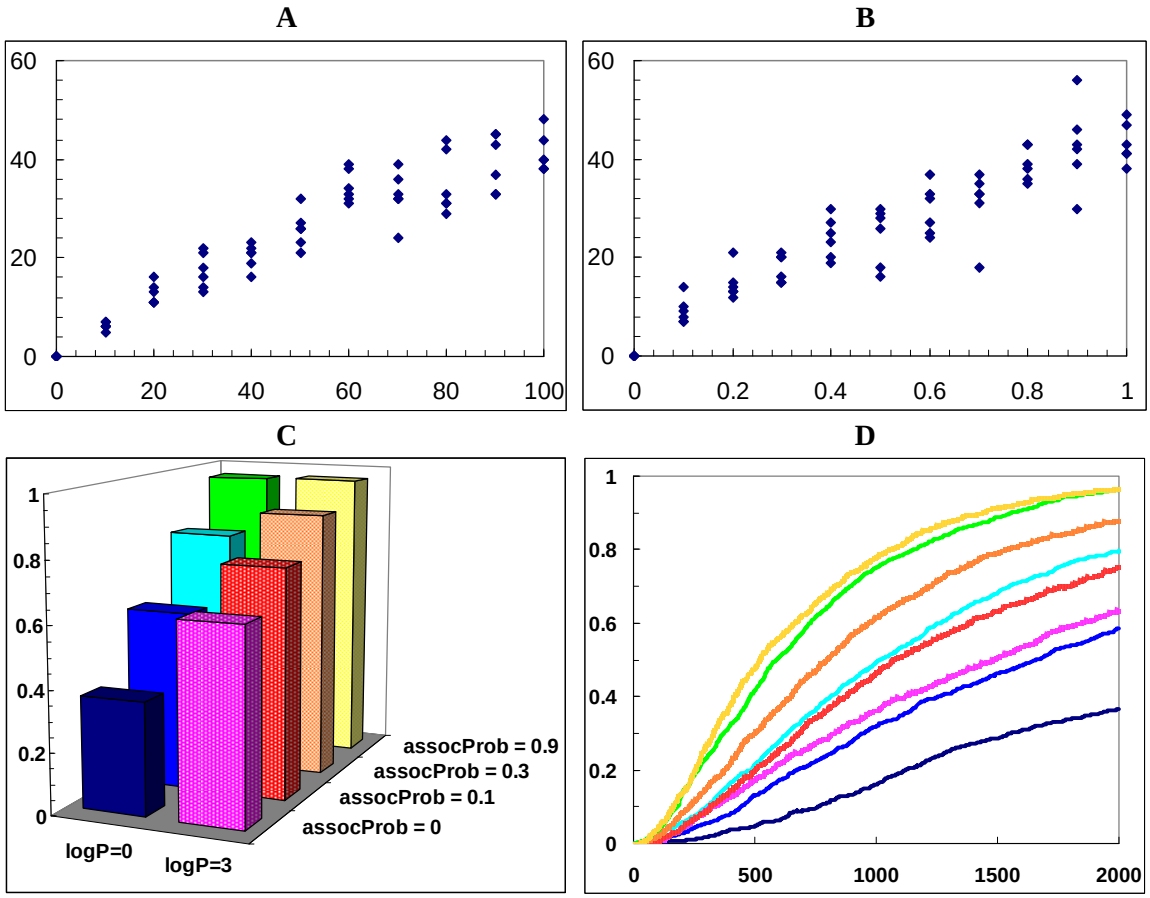


Figure 3.4. TRANSPORTER verification and consistency with simple Michaelis-Menten kinetics. **A:** The data show a linear relationship between PERMEABILITY as fraction of dose (y-axis) and the number of TRANSPORTERS. **B:** The data show a linear relationship between PERMEABILITY (y-axis) and the value of the COMPOUND'S *assocProb* (affinity; x-axis). **C:** The graph shows the combined effects of active transport and passive transition on the permeation of COMPOUNDS that are substrates for the TRANSPORTER and can also undergo passive transition. The graph shows the PERMEABILITY after 2000 simulation cycles. Being a TRANSPORTER substrate has a greater effect on less permeable (lower *logP*) COMPOUNDS. Also, when TRANSPORTER *assocProb* is high, TRANSPORTER-mediated permeation dominates. **D:** The graph shows the permeation time-profile (x-axis: simulation cycles) of the simulations in **C** The coloring scheme is the same.

consequences of changing the values of most of the parameters. In Fig. 3.6, I present data demonstrating the effect of a minor alteration of each of two key parameters: *transitProb*(S1→S2) and DIGOXIN'S *assocProb* to CYP. I ran these experiments using the same conditions and parameter values used for the final RISL analogue to be presented in chapter 4, except that the

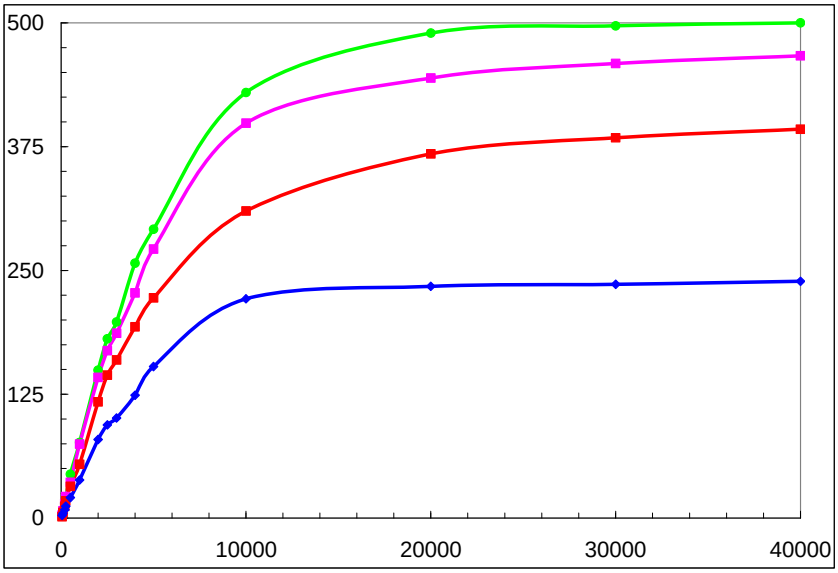


Figure 3.5. ENZYME function consistency with simple Michaelis-Menten kinetics. The graphs show the rate of METABOLISM (y-axis: number of METABOLITES formed after 10 simulation cycles) when starting with different numbers of COMPOUNDS (x-axis). The duration of each experiment was 10 simulation cycles. All components were confined to S3; 50 CYP were used. Graphs from top to bottom: 1) no inhibition; 2) competitive inhibition using 250 additional COMPOUNDS that are INHIBITORS; 3) competitive inhibition using 1,000 COMPOUNDS that are INHIBITORS; 4) noncompetitive inhibition caused by reducing *efficiencyProb* to 40% of it original value. The data are consistent with simple Michaelis-Menten kinetics.

value of each of the two parameters was increased 5%. The results demonstrate that changes in parameter values of a few percent do not significantly alter simulated profiles.

3.10. Software

The analogues of in silico pharmacokinetic systems were assembled within the Swarm platform using its libraries (http://swarm.org). I coded in Java Swarm. Most experiments used a single processor and ran under Microsoft Windows XP (Redmond, WA) with Java Software Development Kit and Java 2 Runtime Environment installed (Sun Microsystems, Santa Clara, CA). Source code was compiled with Java 2 Software Development Kit SE version 1.4.2_13 and executed with Java 2 Runtime Environment version 1.4.2_13 (www.java.com).

All pseudo-random number generation used Swarm's Mersenne Twister algorithm; the initial seed was extracted from the machine's clock. All PRNs were drawn from uniform

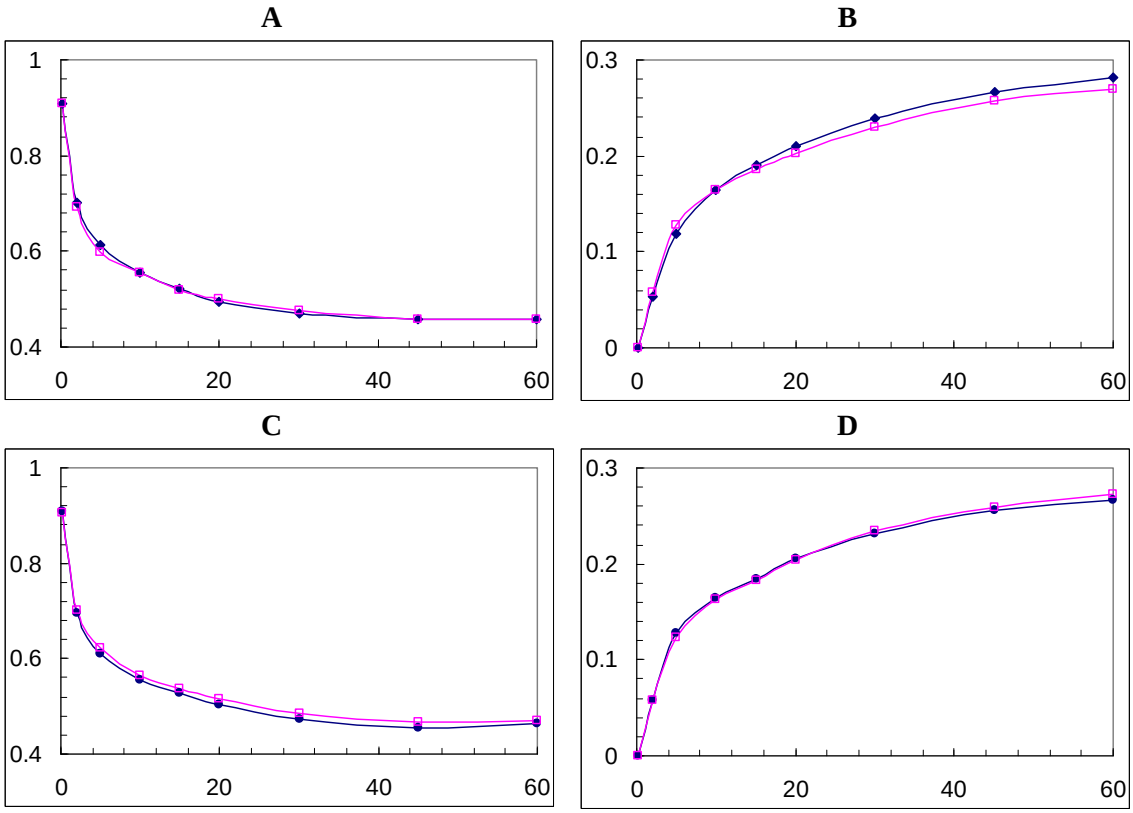


Figure 3.6. Influence of small changes in two of the more important parameters. X-axis: time after adding DIGOXIN to PERFUSATE; y-axis: fraction of dose; closed symbols: results from final RISL analogue; open symbol: results from increasing parameter value by 5%; left:. **Top:** Graphed are the results showing the effect of 5% increase in *transitProb*(S1→S2) on **A:** DIGOXIN and **B**: METABOLITE in PERFUSATE. **Bottom:** Graphed are results showing the effect of 5% increase in DIGOXIN'S *assocProb* for CYP on **C** DIGOXIN and **D**: METABOLITE in PERFUSATE .

distributions. Random integers were drawn from a uniform distribution between [min, max], inclusive. Random floating-point numbers were always drawn from a uniform [0.0, 1.0)—including 0.0, excluding 1.0.

3.11. Discussion: thinking about relational grounding

Given the variety of experimental conditions, protocols and apparatus that are commonly used in pharmacokinetic experiments, how can we enhance substantially our ability to integrate and reconcile results and conclusions across experiments? While experimental systems differ between experiments, the biological mechanisms giving rise to them remain similar. A model

that describes not just the observations, but also the underlying mechanisms will help in understanding biology across different experiments. Traditional, equation-based models are not suitable for this task because they are induced from the observables: they describe the data.

Relationally-grounded models shift the focus of modeling and simulation from achieving precise, quantitative predictions of experimental data to discovering and achieving an understanding of plausible mechanisms causing the referent phenomena. The approach avoids offering precise predictions without physiological meaning or mechanistic understanding as is frequently presented by traditional equation-based modeling approaches (Rescigno, 2004). The analogues described herein are representations of our current understanding of mechanisms across different experimental protocols while keeping assumptions about observations and experiments external. As I will show in chapters 4 and 5, analogues may even be useful in translating to different experimental contexts that share (some of) the same biological mechanisms. Such analogues facilitate integration and reconciliation between experiments, thereby providing a means for achieving deeper insight into the referent. Using these analogues, I was able to discover new, plausible mechanistic details of hepatic drug interactions with RISLs in chapter 4 and evaluate the plausibility of mechanistic hypotheses with ISECCs in chapter 5.

To make an analogue somewhat independent from experimental protocols, measurements, and observation perspectives, one must recognize that experimental protocols and systems of conventional units act as filters (as models). Such filters are implicit in the investigator's mind and reflect his/her knowledge, interpretation and assumptions about the biology and the experiments. It is part of the experimental model I referred to in chapter 1. Biological consequences to be observed are selected by the protocol, and observations are filtered through the aspects on which we focus to become values with physical units attached, which are recorded as measurements. This filtering process is a built-in feature of conventional wet-lab experiments because investigators must select certain aspects to be observed; and for those aspects, the

scientist cannot observe the microscopic events and interactions individually and therefore must take a macroscopic view of the system. With that, ideas of how microscopic events aggregate to become macroscopic observations must be in place. For example, assuming the drug is homogeneously mixed in the target medium, one may take a small aliquot of known volume and assay for amount of drug in that sample. The resulting quantity, i.e. concentration, is then used. Systems of conventional units then come into play: for example, one cannot observe numbers of molecules (dimensionless), so a macroscopic concentration (dimensioned) is used; one cannot observe the likelihood of a microscopic biological reaction (dimensionless), so a reaction rate (dimensioned) is used etc.

However, with this new class of model, the analogue—the part that describes biology—can be made separate from the filter and from the part that describes observations. Recognizing and separating this filter, relational model construction must follow these principles: 1) separate biological behavior from measured observations; 2) focus on representing biological behavior, not describing the data; 3) represent components, mechanisms and interactions from the local, individual, microscopic perspective, not from a systematic or macroscopic perspective; 4) avoid systems of conventional units and quantify relationships relative to each other.

Although it is useless to attempt modeling every aspect and detail of the biology, the model can still became a useful analogue by following an iterative model building and validating/falsifying process, as demonstrated in chapters 4 and 5. A level of abstraction was defined under which details were represented by objects, agents, and/or axioms. A minimum set of components, mechanisms, and interactions, which were believed to be causing the biological behavior, was used to assemble a working analogue. When the set became evidently inadequate, additional details (complexities) were evaluated iteratively, validated (or falsified) and incorporated (or rejected.) Even though the analogues presented were relatively simple, they

became scientifically useful representations of the referent while having only a few concrete 1-to-1 mappings between their attributes and mechanisms and those of the referent.

Moreover, because dynamic mechanistic, anatomical, physiological, and molecular biology details are represented during execution, models of this class have the potential to evolve into an executable representation of what we know (or think we know) about biological systems: executable biological knowledge embodiments. I expand on that idea in chapter 6.

4. Discovering Plausible Mechanistic Details of Hepatic Drug Interactions

4.1. Introduction

Given the variety of transporters and enzymes that can be involved, how can we enhance substantially our ability to confidently anticipate the consequences of hepatic drug-drug interactions in advance of costly wet-lab experiments? To do so, we need improved knowledge of multi-level spatiotemporal mechanistic details. We need new classes of models with capabilities beyond those of the current pharmacokinetic (PK) and pharmacodynamic variety, plus methods to more realistically represent critical spatiotemporal details, all within single model systems. Finally, we must be able to explore realistic, concurrent, interaction consequences of two or more drugs when given key physicochemical properties, which is infeasible using current PK models. An essential first step in realizing those needs is to demonstrate a fine-grained, computational analogue in which measures of simulated interactions of two drugs are experimentally indistinguishable from those of referent wet-lab experiments. When the analogue's assembled components map realistically to hepatic components, and the measured consequences of in silico and wet-lab mechanisms are indistinguishable, we can posit that the

responsible mechanisms of both systems are similar, as diagrammed in Fig. 4.1. Achieving these important goals has been an objective of this project.

The referent wet-lab experiments for this project used an ex vivo, perfused rat liver protocol to measure interactions between digoxin, rifampicin, and quinidine. The mechanisms are relatively well understood (Lau et al., 2004). Digoxin, a P-glycoprotein substrate (de Lannoy and Silverman, 1992), is transported basally in hepatocytes by Oatp2 (*slc21a5*) (Noe et al., 1997). Because rifampicin is an Oatp2 inhibitor and quinidine is a P-glycoprotein inhibitor, the disposition properties of digoxin are altered when it is co-administered with either.

The synthetic method of modeling and simulation (Hunt et al., 2006; Yan et al., 2008a; Yan et al., 2008b), also called executable biology (Fisher and Henzinger, 2007; Hunt et al., 2008), offers advantages over traditional methods for exploring and testing hypotheses about mechanisms. Traditional PK models formulate hypotheses about data, whereas the assembled components of a synthetic model are a testable mechanistic hypothesis about spatiotemporal details underlying that time course data. Execution tests that hypothesis. Measures of analogue attributes can overlap corresponding measures of referent attributes. Measures of attribute similarity provide a quantitative means of accepting or rejecting the hypothesis.

I used the synthetic method to construct and verify liver analogues suitable for simulating recirculating liver perfusion experiments. They were similar to, but simpler than, the one used in (Yan et al., 2008a). The result was the Recirculating In Silico Liver (RISL) diagrammed in Fig. 4.2. Objects representing Oatp2s and P-glycoproteins were included within separate (membrane) spaces. Cell interiors and perfusate contained objects responsible for nonspecific drug binding. Different objects represented drug-metabolizing enzymes. Spaces were designated to represent bile and perfusate. Three different mobile objects carried information identifying each as representing digoxin, rifampicin, or quinidine (Lau et al., 2004). RISL components used that information to distinguish between them.

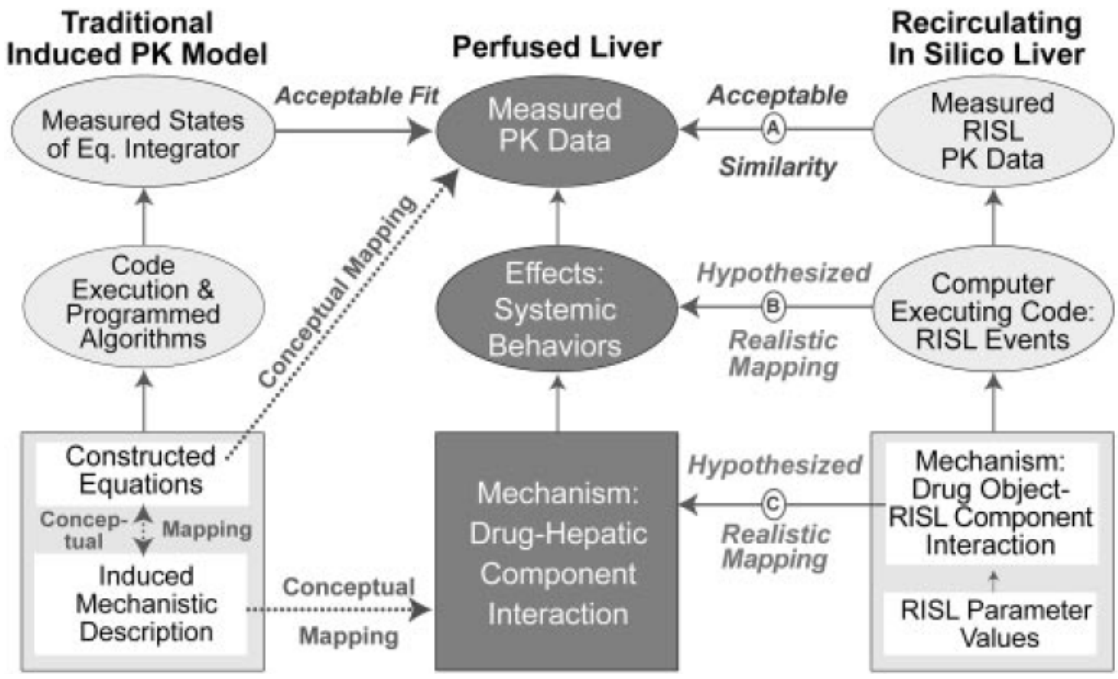


Figure 4.1. Relationships between wet-lab, perfused liver experiments, traditional PK models, and the RISLs. **Center**: rat livers in an experimental context are the referent systems. During experiments, lobular components interact with transiting drug molecules to cause changes in digoxin's concentration-time profile (and that of its primary metabolite). The system's behaviors during the experiment are reflected in the collected data. **Left**: the researcher identifies patterns in the wet-lab data. From those and prior PK knowledge, a mechanistic description of what is thought to have occurred during the experiment is induced, thus establishing an abstract, conceptual mappings from that description to hepatic mechanisms. The researcher offers a set of PK equations believed capable of describing the time course patterns identified in the data. A discretized, validated model of the equations in software is constructed and executed to simulate parameterized equation output. Metrics specify the goodness of fit of the simulated output to the data, and that establishes a concrete mapping from simulated output to wet-lab data. **Right**: The abstract mechanistic description (Fig. 4.2) is different from that on the left side. Software components are designed, coded, verified, and assembled and connected guided by that mechanistic description. The product of the process is a collection of mechanisms rendered in software. A clear mapping—C—is intended to exist between components and how they plug together, and hepatic physiological and microanatomical details. Relative similarity is controlled in part by parameterizations. Importantly, mapping C can be concretized iteratively. Compilation and source code execution gives rise to a working analogue. Its dynamics are intended to represent abstractly (mapping B) corresponding dynamics (believed to occur) within the liver during an experiment. Mapping B can also be concretized iteratively. Measures of simulated dynamics provide time series data that are intended to mimic corresponding measures of wet-lab perfusion experiments. Achieving SM-1, -2, and –3, enables mapping A to be made concrete.

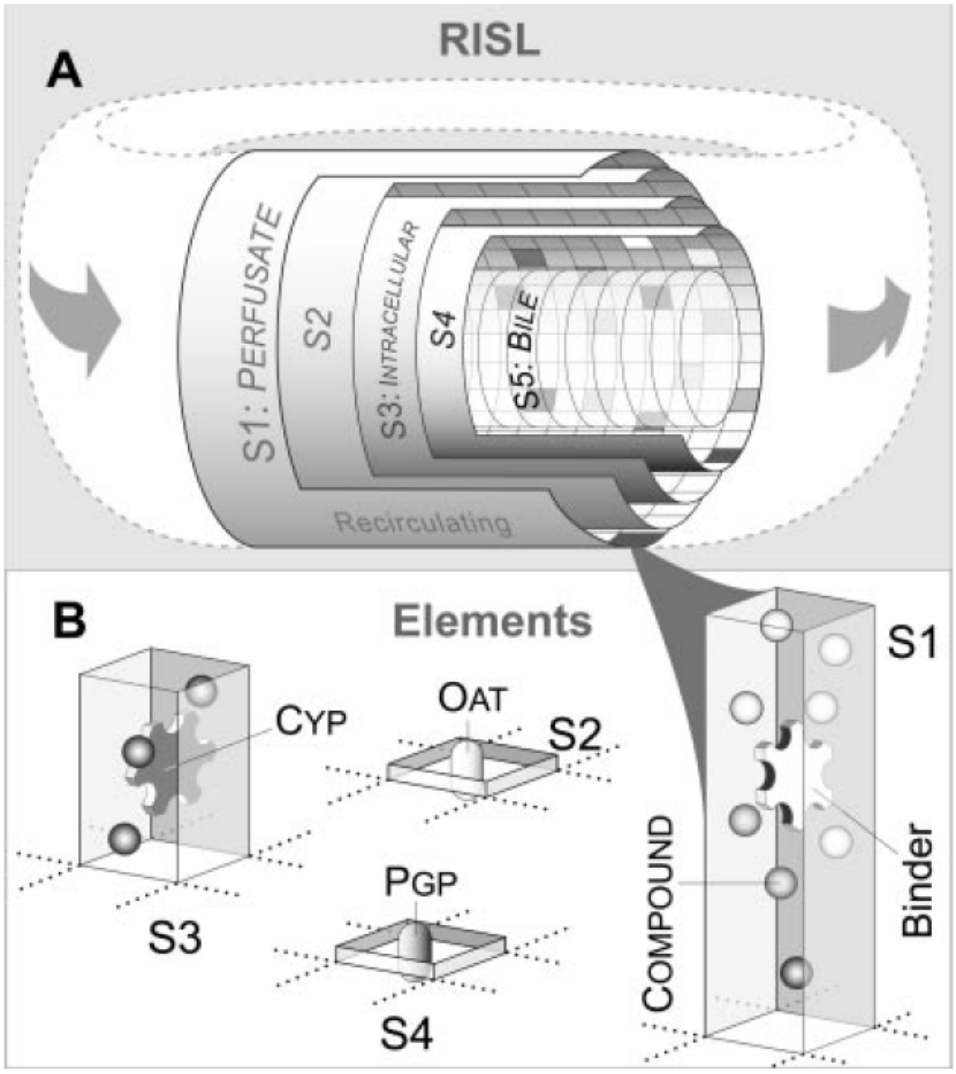


Figure 4.2. Illustration of the RISL and its components. **A:** The RISL is represented as five concentric spaces (S1–S5) with S1 representing recirculating perfusate and rapidly equilibrating intrahepatic spaces. Each space is a 50x50 toroidal square grid. S2 represents barriers between S1 and S3. S3 maps to intracellular spaces. S4 represents barriers between S3 and S5. S5 maps to bile. ELEMENTS are containers for other objects. One is located at each grid location. Within a space, they have the same relative capacity, but capacity differs between spaces. The grid spaces shaded differently illustrate that each element can contain different objects. **B:** Shown are illustrations of four ELEMENTS, one each from S1-S4, with examples of the object they may contain. Different ELEMENT sizes illustrated different capacities. ELEMENTS can contain mobile and fixed objects. The spheres represent mobile COMPOUNDS; the different shadings illustrate the different states in which they can exist. A fixed OAT object is shown in the S2 element, and a fixed PGP object is shown in the S4 element. A stylized illustration of a multi-site CYP is shown in S3; one COMPOUND is shown bound. A stylized illustration of a multi-site binder is shown in S1.

I started with a simple, time-invariant RISL having five spaces along with components representing digoxin, two inhibitors, CYP3A enzymes, Oatp2s, and P-gps. RISL disposition profiles of rifampicin and quinidine were similar to the reported liver perfusion profiles. Disposition and metabolite formation profiles for digoxin, administered alone, were also acceptable matches to liver perfusion profiles, using a quantitative similarity measure. However, corresponding profiles following a pre-dose of rifampicin or quinidine were completely dissimilar to referent values. Because implemented mechanisms were too simple, that initial RISL failed to simulate observed drug interactions. Adjusting that RISL's spatial, component, and interaction details, failed to uncover plausible explanations. A protocol was followed for adding additional mechanistic complexity. It used three, increasingly stringent measures of similarity. Postulating time-dependent changes in specific details, including simulating hepatic injury, erosion in the quality of regional hepatic perfusion and increasing loss of hepatic viability dramatically improved the quality of match. It took the combination of four different time-dependent changes for the RISL profiles to achieve the most stringent similarity measure, yet changes in hepatic physiology were not evident during wet-lab experiments.

Because wet-lab model systems are artificial, measures are taken to minimize change, yet some may go undetected and can influence observables. The methods used uncovered such changes thereby providing deeper insight into plausible drug-drug interaction mechanisms, even though the nature of some was unexpected.

4.2. Methods

To distinguish clearly in silico components and processes from corresponding hepatic components and processes, I use SMALL CAPS when referring to the former, *italics* denote RISL parameters, variables and internal states. In contrast to inductive models, which often focus on prediction, the RISL and its method have been designed for discovering and testing plausible,

detailed mechanistic explanations of hepatic drug disposition and interaction data. A specific RISL instantiates a hypothesis (Fisher and Henzinger, 2007; Hunt et al., 2008). Execution and comparison of results to referent counterparts tests the hypothesis.

4.2.1. Objective and approach

My objective was to discover a single parameterized set of multi-level hepatic mechanisms that would provide one plausible explanation for the three pairs of the time series data in Fig. 4.3, while achieving a stringent measure of similarity as described below. One pair of time series data describes the PK and metabolism of digoxin administered alone. The other two are measures of the PK and metabolism of digoxin when the livers were predosed ten minutes earlier with rifampicin or quinidine. I was not seeking a traditional, differential equation explanation of time course data. Rather, I sought an actual, working mechanism—*an analogue*—comprised of quasi-autonomous processes and parts, which when measured during execution would give time course data similar to wet-lab data in Fig. 4.3. My technical approach used the synthetic modeling and simulation method (Hunt et al., 2006): I plugged together validated, quasi-autonomous software components to form an abstract, but mechanistically realistic analogue of a liver undergoing perfusion into which one could add, alone or in combination, objects representing the three different drugs: digoxin, rifampicin, and quinidine. My experimental approach followed an iterative sequence: RISL synthesis, testing and evaluation, validation or falsification, assessment, cogitation, and system revision.

Hunt et al. (2006) and Yan et al. (2008a and b) describe an In Silico Liver comprised of similarly constructed LOBULES. Because perfusate was recycled in the wet-lab experiments, the data did not require that level of detail, so I designed, validated and used a simpler in silico liver. The rationale for doing so is explained in chapter 3. During RISL design, my focus was on simplicity. Parsimony is important. Traditional PK modeling strives to avoid overparameterization. By analogy, synthetic modeling strives to avoid overmechanization. I

decided to abstract away the complex lobular architecture and interconnected sinusoidal networks and use only five separate spaces in the initial RISL. One (S1) would represent perfusate and extracellular fluid. A second (S2) would represent the barrier (not limited to just cell membranes) between perfusate and cell interiors that would be represented by a third space (S3). A fourth (S4) was needed to represent the barrier between bile and cell interiors. The fifth (S5) represented bile. Although there was no referent bile data, S5 was needed as the location for COMPOUND elimination. The temporal resolution of the data in Fig. 4.3 was insufficient to distinguish and validate separate representations of a layer of endothelial cells, the space of Disse, and hepatocytes. Consequently, I conflated their representation. The resulting five spaces can be viewed as a tube comprised of a layer of CELLS (S2, S3, and S4, representing primarily hepatocytes), flanked by PERFUSATE and rapidly equilibrating EXTRACELLULAR FLUID (S1) on the "basal" side and BILE (S5) on the "apical" side. Together the five spaces function as a simplified analogue of a hepatic lobule, or a portion thereof. As in Hunt et al. (2006), I pooled or averaged results from n Monte Carlo LOBULE variants to represent an entire RISL. Drawing on results of early simulations, I decided that $n \geq 6$ were adequate for the objective (experimental results supporting that choice are provided in chapter 3). The referent wet-lab experiments used $n = 6$. As in the above-cited In Silico Liver, I specified three categories of components: spaces, active objects and passive objects. They are listed in Table 4.1. Primary parameter values are listed in Table 4.2. Additional information is provided in chapter 3.

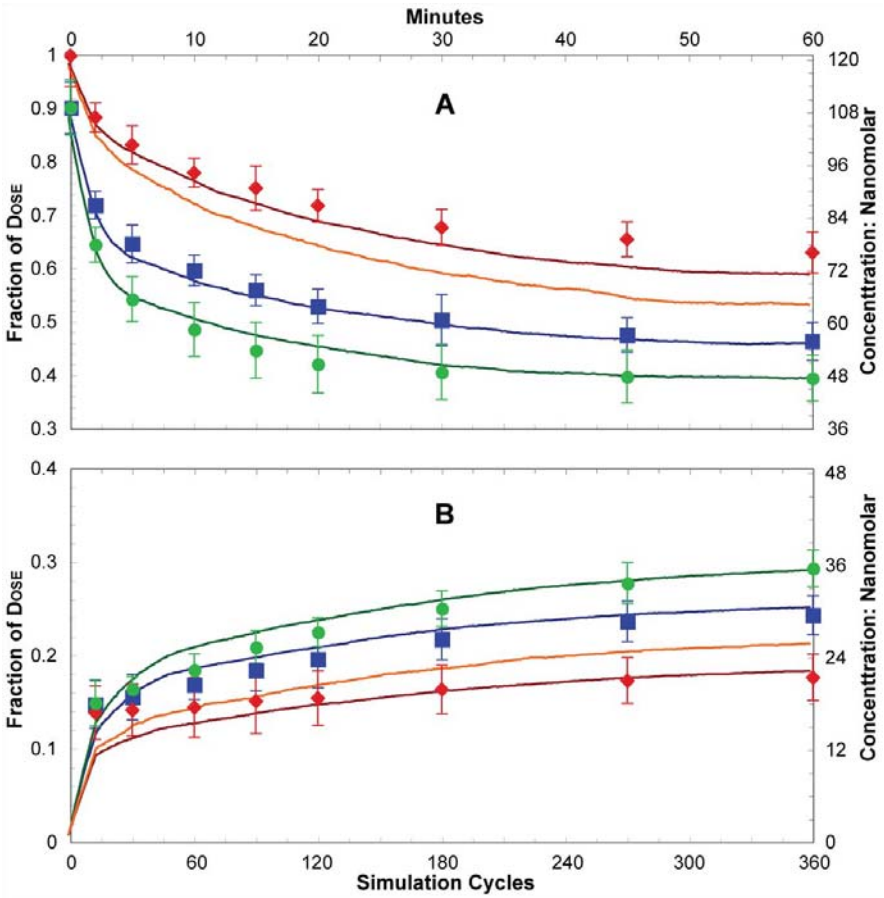


Figure 4.3. Simulated results using RISLs with the time-variant mechanisms graphed in Fig. 4.4. Symbols are values reported in Lau et al. (2004) of fraction of digoxin dose remaining in the perfusate for each of the three different sets of experiments (each set: $n = 6$); the bars are ± 1 SD for the wet-lab data. The curves are mean simulated values of 18 Monte Carlo executions of the RISL parameterized to exhibit time-variant mechanisms according to Tables 4.3–4.5. **Red**◆ : rifampicin pre-dose; **Blue**■ digoxin only, no predose, **Green**●: quinidine pre-dose; **Orange**, both inhibitors pre-dose. **A**: Shown is the mean DIGOXIN PERFUSATE levels. In all three cases, simulated DIGOXIN values achieved SM-3; most were within a SD of referent values. **B**: Shown are the mean METABOLITE PERFUSATE levels. Again, most simulated values were within a SD of referent wet-lab data. Consequently, all simulated METABOLITE values achieved SM-3. Collectively, the six sets of simulated data in **A** and **B** also achieved SM-3. The time-variant RISL differs in four ways from those simple, linear, time-invariant RISLs. 1) A METABOLITE-specific TRANSPORTER was added to S2 to increase the rate of return at early times of METABOLITE to PERFUSATE. 2) The accessibility of S2 to mobile DRUGS in PERFUSATE (S1 in Fig. 4.2) decreased with time (Fig. 4.4A), decreasing the rate of HEPATIC permeation by all three DRUGS. 3) The efficiency of CYP decreased with time (Fig. 4.4D). 4) In addition, a portion of CYP, OAT, and PGP became inactive (Fig. 4.4C) in each simulation cycle. A small number of less functional CYPs were added to S1 (perfusate) to mimic enzymes released from injured hepatocytes during system preparation.

4.2.2. Model structure and components

S1–S5, sketched in Fig. 4.2, have been implemented as two-dimensional, toroidal grids having identical granularities. Grid granularity (grid dimensions) is parameter controlled. Simulation run time is proportional to grid size, and measurement variance increases with decreasing grid size. The exploratory nature of parameter tuning and iteratively revising analogues necessitates running many simulations. Balancing these considerations, I settled on a 50x50 grid size for all experiments.

At each grid location, I placed an object that functions as a container for other objects. I call it an ELEMENT (of that space). Each of the 2,500 S3 ELEMENTS mapped to a portion of a

Table 4.1. Comparing the perfused liver with the RISL, its in silico analogue

Biological Aspects; System Features	Analogue Components and Features
Form; space	GRIDS; ELEMENTS that function as containers
Function; mechanism	Sequence of events; data transforms in object methods and behaviors
Time	Simulation cycles
Drug (and inhibitor) molecules	Mobile objects: DRUGS, COMPOUNDS
Perfusate; experimental apparatus	SPACE S1; COMPOUNDS are in one of three states
Liver = many similar lobules	A number of simulations, averaged
Cell contents	ELEMENT objects assigned to SPACE S3
Bile	ELEMENT objects assigned to SPACE S5
Cell content–sinusoid interface	ELEMENT objects assigned to SPACE S2
Cell content–bile interface	ELEMENT objects assigned to SPACE S4
Metabolic enzymes and transporters	Stationary objects: CYP, OAT, PGP, MT, BINDERS
Mechanisms	Sequence of events
Passive dispersion	COMPOUNDS change location within SPACES
Passive transport	COMPOUNDS move between SPACES
Transport (active or facilitative)	COMPOUNDS moved by TRANSPORTERS
Metabolism	COMPOUNDS converted to METABOLITES
Binding	COMPOUNDS attaching to BINDERS
Measures taken during experiments	Measures taken during experiments
Hypothesis; hypothesis testing	A specific analogue; a set of simulations

Table 4.2. Primary parameters of the Recirculating In Silico Liver (RISL) system

System Parameters	Remarks	Values
systemSize	Size of SPACES S1–S5	50 x 50
numSolute	Dose: number of SUBSTRATES: DIGOXINS	1,000
numInhibitor	Dose: number of INHIBITORS: RIFAMPCIINS or QUINIDINES	2,000
adminTimeSolute	Administration time of SUBSTRATE	60
adminTimeInhibitor	Administration time of INIHIBITOR	0
insilicoPH	In silico pH of SPACES	7.4
numPgps	Number of PGP	30
numOats	Number of OAT	100
numCyps	Number of CYP (ENZYMES)	30
numBinders	Number of BINDERS	75
numMTransporters	Number of METABOLITE TRANSPORTER (MT)	100
timestepfactor [a]	Scaling factor for passive permeation	2.0
internalProb	Flow rate: probability of a compound in S1 being *internal* to the liver; simulates flow from bulk perfusate into sinusoidal spaces	0.13
maxSites	Number of substrate binding sites per stationary object	2-10
efficiencyProb [b]	Efficiency parameter for CYP, PGP, and OAT	
sitesN [b]	Size of neighborhood: number of adjacent ELEMENTS scanned each cycle by CYP, PGP, and OAT	
defunctProb [b]	Probability of going *defunct* within the simulation cycle	

[a] See chapter 3 for a detailed explanation.

[b] Denotes time-variant parameters: see Table 4.5

lobule's intracellular content. ELEMENTS in S2 and S4 mapped primarily to portions of cell membranes, including hepatocyte apical and basal membranes, along with some of the adjacent material. All TRANSPORTERS were placed randomly and uniformly within these two spaces. S2 also mapped to features of the sinusoidal endothelial layer and the adjacent space of Disse. However, such detail was below RISL resolution. Each S5 ELEMENT mapped to a portion of bile, and an ELEMENT of S1 mapped to portion of perfusate plus the rapidly equilibrating extracellular spaces (hereafter referred to as PERFUSATE). An ELEMENT can be identified and labeled with relevant chemical properties. For example, each element of S1 can be assigned whatever referent properties are required for the experiments planned, such as pH, ionic strength, etc. These properties can be made different for S5 (BILE).

ELEMENTS can contain mobile and stationary objects. I included only objects that proved necessary to achieve my objective. Unlike with traditional physiologically based PK models, new components can be added without interfering with the function of those already present. A COMPOUND (also called DRUG) was a mobile object that mapped to a small fraction of the drug added to perfusate at t = 0, or, in the case of METABOLITE, generated within hepatocytes. As in previous reports (Hunt et al., 2006; Garmire et al., 2007; Yan et al., 2008a; Yan et al., 2008b), one COMPOUND mapped to a large but unspecified number of actual referent molecules. A RISL used four types: DIGOXIN, METABOLITE (of DIGOXIN), RIFAMPICIN, and QUINIDINE. Their RISL properties are listed in Table 4.3 along with referent's properties. To enable each RISL stationary object to distinguish different types of mobile objects, each COMPOUND carries identification; it could include values representing several of the referent compound's properties. RISL stationary objects were given logic (diagrammed in chapter 3 Fig. 3.2) that enabled them to use a COMPOUND'S properties to control how they interacted with COMPOUNDS. For examples, see (Yan, et al., 2008b. Sheikh-Bahaei and Hunt, 2006). Parameter values were tuned separately for each COMPOUND. Each COMPOUND, selected randomly, has one opportunity to relocate during each simulation cycle. Six simulation cycles mapped to one minute of wet-lab time. A COMPOUND relocation event was subject to the COMPOUND'S local environment. That process was a variation on those used previously and cited above.

4.2.3. TRANSPORTERS, BINDERS, and CYP

An RISL uses three classes of active, stationary objects. One maps to transporters. The second, CYP, maps to the enzymes responsible for digoxin metabolism, primarily CYP3A. The third class is BINDERS; they map to anything within a lobule or perfusate that might bind or sequester compounds. The logic used by each is diagrammed in chapter 3 Fig. 3.2 along with additional validation evidence. I implemented initially two types of transporters: PGP and OAT. PGP maps to P-glycoprotein and any other transporters responsible for efflux of digoxin from

hepatocytes to bile. OAT maps to Oatp2 and any other transporters that are responsible for the active cellular uptake of digoxin. However, the details of PGP and OAT operations were below the level of resolution and so do not map to biochemical counterparts.

CYP binding sites within a CYP are identical and independent. As with PGP and OAT, details of CYP operations are below the level of resolution and so do not map to biochemical

Table 4.3. Compound specific parameters values

COMPOUND Parameters	Remarks		Values			
			DIGOXIN [a]	RIFAMPICIN [a]	QUINIDINE [a]	METABOLITE [b]
MW [c]	molecular weight of the solute		781	822	324	601
logP [c]	logP - common logarithm of octanol/water partition coefficient		1.14	3.60	2.53	1.60
pKa [c]	pKa closest to *insilicoPH*		13.5	6.9	8.6	13.5
inihibitorType	uptake inihibitor? True/False		n/a	TRUE	FALSE	n/a
closeToInterface	Fraction sufficiently close to interface such that it is *transitional* [d]	S1	--[e]	0.75	0.75	0.10
		S3	0.10	0.02	0.04	0.10
		S5	0.10	0.05	0.05	0.10
transitProb	Initial estimates of trans-membrane transit probabilities, after scaling by *timestepfactor*	S1→S2	0.066	0.069	0.102	0.078
		S2→S1	0.123	0.007	0.099	0.068
		S2→S3	0.123	0.007	0.088	0.068
		S3→S2	0.066	0.069	0.104	0.078
		S3→S4	0.066	0.069	0.104	0.078
		S4→S3	0.123	0.007	0.090	0.068
		S4→S5	0.123	0.008	0.060	0.068
		S5→S4	0.066	0.069	0.106	0.078

[a] Physicochemical properties reflect those of the referent drug.

[b] Physicochemical properties reflect those of digoxigenin bis-digitoxoside.

[c] *Molecular weight*, *logP*, and *pKa* are used in calculating the initial, passive transit probabilities, as described in chapter 3. For *pKa*, only the value closest to 7.4 is listed.

[d] The *closeToInterface* value for all solutes in S2 and S4 was always 1. This reflects the fact that the referent barriers are sufficiently thin so that *all* solutes within them at the start of a 10 second interval (a simulation cycle) will have had an opportunity to transition by the end of that interval. For all spaces except S1, *closeToInferace* is the *fraction transitional*

[e] *CloseToInterface* for S1 is a time-variant parameter. It, together with *internalProb*, specifies the *fraction transitional* (Table 4.5).

counterparts. CYP parameter values are listed in Tables 4.2, 4.4 and 4.5. At the start of a simulation, all CYPS were assigned randomly and uniformly to elements within S3. Each simulation cycle, each CYP steps through its assigned logic to determine what action to take. It examines its adjacent neighborhood. Neighborhood size is controlled by *sitesN*. The probability that a CYP-DRUG binding event will occur for any DRUG within a CYP'S neighborhood is governed by *assocProb* (Table 4.4), the value of which maps to affinity. A CYP can bind any COMPOUND that has a non-zero *assocProb*, whether it is a substrate or not. However, only substrates are METABOLIZED. The frequency at which METABOLISM occurs is controlled by the tunable METABOLIC parameter *efficiencyProb*. After that, each CYP examines each bound COMPOUND; each is given an opportunity to be RELEASED within that same cycle. That process is governed by

Table 4.4. Parameters for COMPOUND-PROTEIN interactions

COMPOUND-PROTEIN Interaction Parameters		**Values**			
		DIGOXIN	**RIFAMPICIN**	**QUINIDINE**	**METABOLITE**
isaSubstrate substrate of ? TRUE / FALSE	MT	FALSE	FALSE	FALSE	TRUE
	PGP	TRUE	FALSE	FALSE	TRUE
	OAT	TRUE	TRUE	FALSE	TRUE
	CYP	TRUE	FALSE	FALSE	FALSE
	BINDER	FALSE	FALSE	FALSE	FALSE
assocProb reflects binding affinity to protein	MT	0	0	0	0.9
	PGP	0.4	0	0.9	0.4
	OAT	0.75	0.7	0	0.05
	CYP	0.5	0	0	0
	BINDER	0.3	0	0	0.5
releaseProb reflects dissociation from protein	MT	1.0[a]	1.0[a]	1.0[a]	0.9
	PGP	0.7	1.0[a]	0.04	0.7
	OAT	0.8	0.05	1.0[a]	0.9
	CYP	0.2	1.0[a]	1.0[a]	1.0
	BINDER	0.1	1.0[a]	1.0[a]	0.01

[a] Not used because *assocProb* = 0.

releaseProb (Table 4.4). When the COMPOUND is a substrate, the released object may be a METABOLITE.

Each PGP and OAT functions similar to CYP. Table 4.4 shows that *isaSubstrate*, *assocProb*, and *releaseProb* are tuned to different values for PGP, OAT, and, CYP as well as for the different COMPOUNDS. PGPS are assigned randomly and uniformly to S4. They look for COMPOUNDS in S3. When one is found and the value of *assocProb* is non-zero, it is given an opportunity to bind. However, only substrates are TRANSPORTED to S5; the COMPOUND'S property list is checked to determine if *isaSubstrate* is true for PGP. If so, it qualifies as a substrate and has an opportunity to be TRANSPORTED to S5. Analogously, OATS were assigned randomly and uniformly to S2. They look for COMPOUNDS in both S1 and S3. When one is found and *assocProb* ≠ 0, it is given an opportunity to bind. If *isaSubstrate* is true for OAT, it qualifies that COMPOUND as a substrate and it is given an opportunity to be TRANSPORTED.

Table 4.5. Time-variant parameters

Time-variant parameter			Initial value	Terminal value
Effective SURFACE AREA changes	*Fraction transitional* [a] in S1		0.13	0.052
	sitesN (neighborhood)	CYP	22	2
		PGP	9	4
		OAT	20	4
		MT	25	4
METABOLIC and TRANSPORT activity changes	*efficiencyProb* [a]	CYP	0.95	0.03
		PGP	1	1
		OAT	0.65	0.65
		MT	1	1
	defunctProb	CYP	0.0042	0.018
		PGP	0.0003	0.001
		OAT	0.0017	0.0057
		MT	0.0003	0.001

[a] The method for changing *fraction transitional* and *efficiencyProb* (Fig. 4.4A and 4.4D) is detailed in chapter 3.7, under Probability scaling.

Frequency of release into the space of destination is controlled by *efficiencyProb*. When PGP and OAT find and bind a non-substrate, such as an inhibitor, it is always released in a subsequent simulation cycle to the space from which it was bound.

A BINDER is an object that functions logically similar to CYP, except that it does not metabolize. It simply binds and later releases the COMPOUND. BINDERS can be added to any ELEMENT, but in these RISLs, they have been confined to S1 and S3. Briefly, a BINDER searches its assigned neighborhood for a COMPOUND. For each one found, its property list is scanned. If the COMPOUND qualifies for binding, it is given an opportunity to bind. If bound, it will be released in a later simulation cycle. Binding and release events for each site are controlled by the BINDER and COMPOUND-specific values of *assocProb* and *releaseProb* listed in Table 4.4.

4.2.4. Tunable parameter values and movements of COMPOUNDS

Most events, such as a COMPOUND binding to CYP, are probabilistic and have values in the 0–1 range. When an event option arises, a pseudo-random number (PRN) is drawn from a range such as 0–1. Its value is compared to that of the parameter to decide what action to take. For example, if the PRN is less than the value assigned to *assocProb*, the event occurs. Otherwise, it does not.

The percolation of compounds through lobular spaces during perfusion is simulated by individual movements of unbound COMPOUNDS. The logic governing COMPOUND movement is adapted from the validated logic used previously and is diagrammed in chapter 3 Fig. 3.2, where additional details are also provided. Briefly, in each cycle, unbound COMPOUNDS are given an opportunity to move. Next, TRANSPORTERS, CYP, and BINDERS are given opportunities to bind and then release COMPOUNDS. The order in which COMPOUNDS are updated is randomized within each cycle independent of the space in which they are located. A COMPOUND first has an opportunity to relocate laterally or stay in its current ELEMENT. The current and eight adjacent

ELEMENTS are each assigned one of nine integer values 1–9. A PRN is drawn randomly from [1–9]; the COMPOUND moves to the ELEMENT assigned the same integer value. Following that, the COMPOUND has an opportunity to move into an adjacent space.

To transit across a barrier, a compound must be sufficiently close to the interface. To represent that, I use the probability parameter, *closetoInterface*. The barriers to which S2 and S4 map are represented as being very thin (less capacity) relative to S1, S3, and S5. For simplicity, given the fact that a simulation cycle maps to 10 seconds, I specified *closetoInterface* = 1 for all compounds in S2 and S4. For a compound in S1, S3, and S5, if a PRN < *closetoInterface*, the compound is given an opportunity to transition; else, it stays put. The probability that a transition will actually occur, given the chance, is specified by another parameter *transitProb*. As done with *closetoInterface*, I estimated an initial value using the compound's properties (see chapter 3) for each transition; when needed, that estimate was tuned. The values listed in Table 4.3 proved adequate for achieving targeted Similarity Measures. The sequence in which each compound's transit decision was made followed the same pattern for each space which is discussed in chapter 3.

4.2.5. Simulated perfusate flow

Perfusion experiments started with the dose being added to the reservoir. Because the flow rate was 40 ml/min, a second or two later, drug-free perfusate entered the reservoir from the liver as drug-containing perfusate was entering the liver. There was a short interval in which drug rapidly distributed into easily accessible extracellular spaces. A few seconds later those spaces were mostly occupied by drug. Thereafter, perfusate compound levels continued declining, but within a 10-second simulation cycle, the difference between entering and exiting drug levels rapidly shrank to within the range of wet-lab analytical sensitivity. Following early exploratory simulations, I estimated that during that initial, rapid distribution phase, about 13% of the dose had been retained within hepatic spaces (was *internal*). At that stage, drug that had not

already transitioned into cells was essentially in one of three states: 1) *external* (to the liver), 2) within extracellular hepatic spaces but not likely to transition across cell membranes within the next 10 seconds (*internal*), and 3) within the liver and sufficiently close to cell membranes so that transition within the next 10-second was possible (*transitional*).

I considered several options to simulate the preceding scenario. One option would be to simulate flow within the space (e.g., S1) representing perfusate, analogous to how it was done in Hunt et al. (2006). S1 could be made circular and much larger than the HEPATIC spaces. COMPOUNDS could move around the circular space simulating flow. To keep the RISL relatively simple, all that was needed was to distinguish the portion of the PERFUSATE that was external to the LIVER at each sampling time. The COMPOUNDS that were randomly assigned to S1 were simply specified as being in one of three states, *external*, *internal*, and *transitional*; and every compound in S1 had an opportunity to change its state each cycle. Changing state simulated both perfusate flow and rapid equilibration between perfusate and hepatic spaces. The parameter *internalProb* specified the probability that a COMPOUND in S1 would be designated *internal* within a simulation cycle. Changing *internalProb* simulated a change in perfusate flow rate. The parameter *closeToInterface* specified the probability that a COMPOUND in S1 would be *transitional* (available to transition to S2). The balance of COMPOUNDS in S1 was *external*. For most of the RISL simulation results discussed below, *internalProb* = 0.13. BINDERS were randomly assigned to S1 and so a small fraction of COMPOUNDS was bound and thus neither available to partition into S2 nor be taken up by an OAT.

When an experiment called for pre-doses of either RIFAMPICIN or QUINIDINE interacting with DIGOXIN, the inhibitor was added to S1 first. The simulation was then run for an additional 60 simulation cycles (10 MINUTES), at which time the simulation was stopped, DIGOXIN was added, and the simulation was re-started.

4.2.6. Targeted attributes are objectives to be achieved

An early protocol task was to specify a target set of phenotypic attributes, and have in mind additional attributes that could be sequentially added to the target list (Hunt et al., 2006; Tang et al., 2007). The goal was to find RISL parameterizations such that measures of simulation attributes, such as METABOLITE levels in PERFUSATE, would be similar to targeted wet-lab attributes in some specified way. My initial set of four targeted attributes was simple (thus less demanding) and did not include the entire time-series profiles. They were 1) the fraction of the digoxin dose remaining in perfusate after 60 minutes, 2) the corresponding perfusate level of metabolite, 3) the increase in 60-minute metabolite levels in the presence of quinidine (assuming significant inhibition of P-glycoprotein), and 4) the decrease in 60-minute metabolite levels in the presence of rifampicin (assuming significant inhibition of Oatp2). A parameterized RISL that achieved those objectives was valid for the targeted attributes. Next, my goal was to expand the attribute list by adding all time-series observations. The final goal was to obtain one RISL, which when dosed with combinations of the three simulated drugs, would generate behaviors that were experimentally indistinguishable from the time series data in Fig. 4.3. At first glance, the preceding four criteria seem lax. That proved not to be the case. Explanations are provided in the next section.

4.2.7. Achieving similarity measures

Specifying a similarity measure (SM) is arbitrary. Three were used. They ranged from least stringent (SM-1) to most stringent (SM-3). I started with SM-1. Once that was achieved, I switched to SM-2, followed by SM-3. Once a SM was achieved, I could state that the RISL had been validated against the available data using that SM criterion. Of course, switching to a more stringent SM may have falsified that RISL.

The targeted attributes coupled all three pairs of time course data. That required one RISL be used to simulate outcomes of all three treatments. If a parameter change was made to

improve, for example, the match to metabolite data following treatment with rifampicin, all of the other five trajectories were also altered. The actual mechanisms responsible for the data in Fig. 4.3 were complex, multilevel, and fine-grained (more so than the RISL). Because the RISL is abstract and course-grained, it was not clear what level of mechanistic detail would be needed to achieve SM-3.

The wet-lab experiments achieved average mass balances of 89 to 91%. Consequently, a requirement of all SMs was that the simulation dose be within 10% of the stated, administered dose. Consequently, DOSE and other parameters were tuned within narrow ranges to achieve targets. The reported coefficient of variation for wet-lab values averaged 6.8% for perfusate digoxin values and 13.2% for metabolite values. To specify a target range, I assumed that repeated wet-lab measures would be normally distributed. I wanted RISL values to be well within ± 2 SD, based on the reported variances stated above. The target range for DIGOXIN was to be within 10% of mean, referent wet-lab values; and within 25% for METABOLITE. SM-1 focused just on the wet-lab values at 60 MINUTES: qualitatively, all three pairs of simulated measures needed to exhibit all four target attributes listed above; and quantitatively, mean RISL PERFUSATE DIGOXIN values at 60-minute needed to fall within 10% of the corresponding mean wet-lab values and within 25% for the METABOLITE values.

SM-2 required that SM-1 be met, and at least 46 of 51 (about 90%) of RISL values (DIGOXIN plus METABOLITE) must be within the target range. This criterion could be achieved by obtaining a good match to all three sets of digoxin time course values and only one set of metabolite time course values. SM-3 required that SM-2 be met, and that for each of the six time course profiles, no more than one measurement would be outside the above target ranges. A simulation result that achieved SM-3 was deemed experimentally indistinguishable: it would have been statistically indistinguishable from the results of a repeat experiment adhering to the original wet-lab protocol under identical conditions.

4.2.8. Tuning and refinement

I started with 50 of each type of stationary object. I divided BINDERS between S1 and S3 based on a ballpark estimate of the relative amount of protein in hepatic cells and in perfusate: I initially placed 95% in S3 and 5% in S1. Most parameters having a 0–1 range were set initially to 0.5. The inhibitors were initially given higher affinities (0.9 or 1.0) and low release probabilities (0.1 or 0.01). Thereafter, a standard parameter-tuning protocol was followed iteratively: 1) simulate, 2) validate or falsify: if the specified SM was not met, then the RISL was falsified. 3) When falsified, diagnose (e.g., an observed mismatch between in silico and wet-lab results may have been caused by specific defects in RISL mechanisms), 4) hypothesis (e.g., the discrepancy will be reduced by specific RISL parameter adjustments or a change in the analogue's structure), 5) testing (repeat steps 1 and 2: contrast new simulation results with those from the predecessor RISL and with wet-lab results), 6) return to step three, or when the SM is achieved and the RISL validates, stop.

4.2.9. Time-dependent parameter changes

Despite intensive exploration of the RISL parameter space, I failed to find a time-invariant parameter vector that would enable achieving SM-2. To do so, the only obvious option was to increase RISL complexity, one alteration at a time, until first SM-2 and later SM-3 could be achieved. That process is described in Results; it too followed the tuning and refinement protocol. One new component was needed: it mapped to efflux transporters for the digoxin metabolite, and was called MT. It resided in S2 and functioned identical to PGP; it moved only METABOLITES from S3 to S1. In addition, I found it necessary to enable four parameter values to change unidirectionally with time. Those changes required adding additional parameters to manage the time-dependent changes. Selection of parameters to become time-dependent was done iteratively, one at a time. Each change was preceded by a hypothesis: increasing (or decreasing this parameter over 60 MINUTES will measurably improve the degree of similarity

between simulated attributes and the wet-lab counterparts (sufficiently so that a statistical test was not needed to confirm that an apparent improvement was real). The simulation results provided a test of that hypothesis.

4.2.10. Software

The RISL was built from the Swarm platform and libraries (http://swarm.org). I coded in Java Swarm. Most experiments used a single processor and ran under Microsoft Windows XP with Java SDK and J2RE installed. Source code was compiled with Java 2 Software Development Kit SE v1.4.2_13, and executed with Java 2 Runtime Environment v1.4.2_13, (www.java.com). Output data files were processed, graphed and analyzed with Microsoft Excel and S-plus (Insightful). I repeated simulations six or more times. Results are reported as arithmetic mean values, unless otherwise noted. I assumed that the central limit theorem held for all observations.

4.3. Results

4.3.1. Overview

My plan was to proceed through three stages marked by achieving increasingly stringent SMs. The stage one objective was to discover a RISL composition (numbers of RISL components and their location) and a parameterization that would produce simulated results at 60 MINUTES that met SM-1. The expectation was that once it was achieved, RISL alterations and reparameterizations could be found that would move RISL behaviors closer to SM-2 and, later, SM-3. Prior to achieving a SM, the RISL was a hypothesis: average results of at least six Monte Carlo experiments will achieve the stated similarity criterion. Simulation results stood as a test of that hypothesis, and a validated mechanism provided new knowledge.

Because a RISL is abstract and relational, I do not require one-to-one correspondence of components or observables between it and the wet-lab system. The goal was that relative,

measured RISL behaviors be similar to those of the referent with the expectation that the two mechanisms will have features in common, as illustrated in Fig. 4.1. To facilitate that process, the original perfusate and metabolite data were transformed to fraction of the administered dose.

Only one RISL was used. The only difference between experiments was the nature of the inhibitor PREDOSE. Changing a parameter value to improve the match to one time course always changed the other five. That coupling shrank the region of parameter space that would enable meeting a SM. It also dramatically limited the shapes of the simulated time course profiles. For many RISLs, there was no region of its parameter space that would enable meeting SM-2 or SM-3.

The steps taken to achieve SM-3 are detailed below. The results are graphed in Fig. 4.3.

4.3.2. Separate disposition of RIFAMPICIN and QUINIDINE implemented

To mimic the wet-lab protocol for two of the three treatments, RIFAMPICIN and QUINIDINE were administered separately prior to DIGOXIN dosing. I then tuned RISL's drug-specific parameter values until I obtained simulated PK time course data for each that was similar to reported data. Time course data for RIFAMPICIN and QUINIDINE administered alone are provided in Supplement Fig. 4.6. Those profiles were strikingly similar to corresponding referent quinidine and rifampicin profiles (Lau et al., 2004) without having either undergo METABOLISM. That similarity indicates that RIFAMPICIN and QUINIDINE metabolism, if measurable, was very modest. Consequently, to preserve simplicity, I elected not to complicate the analogue further by implementing a small degree of clearance for either RIFAMPICIN or QUINIDINE.

A simpler way to represent interaction with DIGOXIN would have been to deactivate all OAT to simulate complete inhibition by RIFAMPICIN, and deactivate all PGP to simulate inhibition by QUINIDINE. However, by so doing, I would have abstracted away (thereby making

assumptions about) potentially important drug interaction details. Having the means to explore drug-drug interactions in some detail was a driving motivation for engineering the RISL as I did.

4.3.3. A simple hypothesis fails

I first explored a wide variety of time-invariant parameter settings and a range of relative ratios and amounts of OAT, PGP, and CYP. No parameter vector was found that was able to provide adequate similarity at early times (and thus possibly achieve SM-2) while still meeting SM-1. I refer to that RISL as the simple, linear hypothesis. I present those results in supplement section 4.5.3 I concluded that a somewhat more complicated set of mechanisms would be needed. Note that the RISL is capable of generating a wide variety of PK profiles (that could be fit using one, two, or more exponentials). The failure of the simple RISL to achieve SM-2 is because its mechanisms were too few and/or too course-grained. I next considered a variety of mechanistic changes, extensions, and additions separately with the goal of improving the degree of match at early times, while still achieving SM-1. None worked. Nevertheless, I describe two that were thoroughly explored to provide background and a context for the changes that followed.

4.3.4. Two mechanistic changes to improve RISL's behaviors

First, I considered that metabolite might compete with digoxin for uptake transporters, slowing digoxin uptake as metabolites accumulate in the perfusate. I implemented the mechanism, specifying a bidirectional OAT, with METABOLITE being an OAT substrate. I explored the expanded parameter space (results not shown). In the absence of an INHIBITOR PREDOSE (the control), early DIGOXIN levels did move closer to referent values. However, in the case of EFFLUX inhibitor (QUINIDINE) PREDOSE, more METABOLITE formed (compared to no pretreatment), causing more inhibition of DIGOXIN uptake, which was inconsistent with the data.

Might a different transporter facilitate rapid metabolite efflux from cells to perfusate causing early, rapid accumulation of metabolite in perfusate? I created a METABOLITE

TRANSPORTER and called it MT, and added it to S2. I explored a variety of parameterizations and found several that caused much higher, early PERFUSATE METABOLITE levels. Adding this new mechanism moved me much closer to satisfying SM-2, but not close enough.

In the above scenarios, RISL mechanisms were unchanged during execution. I speculated that in the referent experiments that might not have been the case: mechanisms and component behaviors may have changed during the course of the experiments. I selected the simple RISL from above, which included MT, that exhibited the properties closest to SM-2 and asked, what time-variant mechanistic changes might lead to achieving SM-2 (and even SM-3)? In the absence of a PREDOSE, more DIGOXIN accumulation in S3 would be needed and that would need to be coupled with more METABOLITE entering PERFUSATE. I began exploring how to achieve those two attributes.

4.3.5. Implementing mechanistic deterioration improved similarity

I pondered: *ex situ*, the liver might be deteriorating during the perfusion experiment, even though there was no reported evidence indicating so (Lau et al., 2004). Ischemic deterioration alone could cause a variety of influential changes, including changes in flow paths, cellular volume, accessible sinusoidal surface area, homeostatic metabolism, and microarchitecture (Straatsburg and Frederiks, 1997; Bailey and Reinke, 2000). Such change could go unobserved during the course of an experiment, and yet alter enzymatic and transport functions. Plausible, yet unobserved, changes include transporter visibility to digoxin being reduced because fenestrae frequency and cross-connections between sinusoids could have been reduced. Digoxin metabolism could decrease because some CYP450 (and/or cofactor) functions became defunct.

I used the inductive method to first identify and then explore several mechanistic changes in line with the above ideas. The following four were aggregated together and implemented.

Results following RISL implementation, refinement, and parameter tuning met SM-2 and later enabled achieving SM-3. Results that achieved SM-2 are presented in Supplement Fig. 4.8. Alterations such as changes in digoxin-accessible sinusoidal surface area, reduced flow in some lobules, decreased fenestrae size or relative density, and/or altered sinusoid microarchitecture were conflated and represented together as a decrease in the effective surface area accessible to COMPOUNDS. This change was implemented by decreasing the value of *closeToInterface* and thus the fraction of dose in S1 that was in the *transitional* state during a simulation. It was reduced using a rapidly decreasing scaling factor. I found it most effective (in achieving closer matches to earlier digoxin perfusate levels) to have most of the decline occur within the first five MINUTES. Behavior of the version used to achieve SM-3 is graphed in Fig. 4.4A.

Decreases in effective, accessible membrane surface area and cell volume were represented by changing COMPOUND visibility: the ability of OAT, PGP, and CYP to "see" and allow access to a COMPOUND in its local neighborhood, which was specified by the value of *sitesN*. I did that by randomly shrinking the size of the local neighborhood (by one ELEMENT) from the initial value (ranging from 9 to 20 ELEMENTS) down to 2–4 ELEMENTS within the first ten MINUTES. The average decrement was one ELEMENT every two simulation cycles. The behavior of the version used to achieve SM-3 is graphed in Fig. 4.4B.

Having COMPOUND visibility decrease in combination with the above time-dependent changes in HEPATIC accessibility improved similarity, but was insufficient to achieve SM-2. Two further time-dependent changes were needed. I enabled PGP, CYP, and OAT to "die" (or to become inaccessible) randomly. I did that by assigning a probability, *defunctProb*, to each PGP, CYP, and OAT (but not a BINDER). It specified the probability that the object would become defunct during any given simulation cycle. When *defunctProb* = 0, each PGP, CYP, and OAT always functioned (as adjusted by *efficiencyProb*). I initially set *defunctProb* to a near-zero value and then increased it gradually thereafter following a quadratic relationship: *defunctProb*(t) = *initial_defunctProb* •

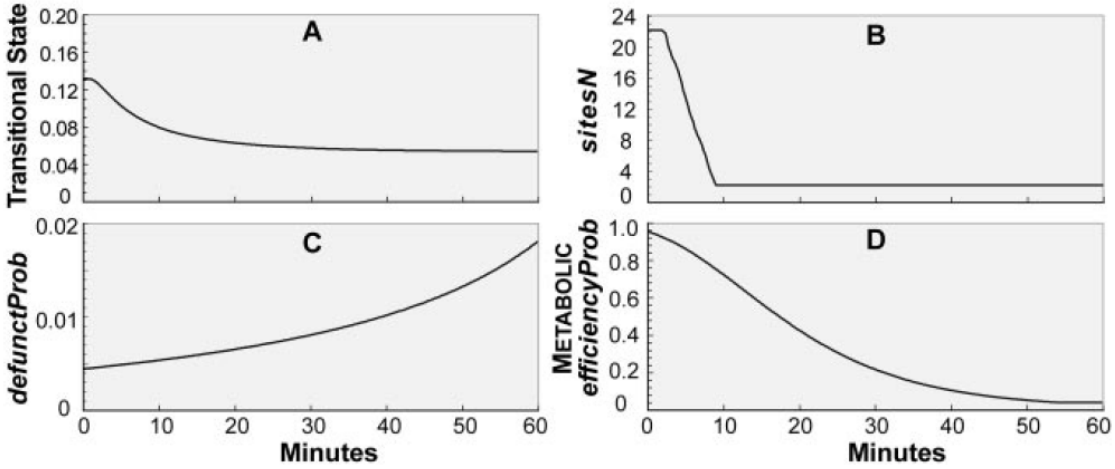


Figure 4.4. Values of time-variant parameters. Each graph shows time-dependent parameter values used to generate the results shown in Fig. 4.3. **A**: The graph shows the time-variant values of the fraction of COMPOUNDS in S1 that were in the *transitional* state, which is defined in the text. **B**: Show are mean, time-dependent values of *sitesN* for CYP averaged over 10 CYP; the pattern of change was the same for PGP and OAT. Factors such as accessible hepatocyte surface area and effective cellular volume map to the ability of OAT, PGP, and CYP to "see," and allow access to, a COMPOUND in its local neighborhood. The initial values of *sitesN* for OAT, PGP, CYP, and MT were decremented randomly. The average decrement was one every two simulation cycles. The initial and minimum *sitesN* values were 22 and 2 for CYP, 9 and 4 for PGP, and 20 and 4 for OAT. **C**: Shows are time-dependent values of *defunctProb* for CYP. The pattern of change was the same for PGP and OAT, and was specified using the quadratic relationship specified in the text. Initial *defunctProb* values: 0.0042 (CYP), 0.00003 (PGP), and 0.0017 (OAT). In each case, the *defunctProb* value at 60 minutes was about three to four times that of the initial value. **D**: Shown are time-dependent values of METABOLIC *efficiencyProb*: initial = 0.95; minimum = 0.03. An algorithm described in chapter 3 was used to adjust the final value along with the early and subsequent exponential rate of decline. *EfficiencyProb* for PGP and OAT remained constant.

(*defunctRate* + t)/(*defunctRate* – t), where *initial_defunctProb* and *defunctRate* are two tunable parameters. The behavior of the version used to achieve SM-3 is graphed in Fig. 4.4C.

The decline in the ability of hepatocytes to metabolize digoxin or transport a compound (into or out of the cell), for whatever reason, was represented by decreasing the efficiency with which a CYP carried out its function. METABOLIC efficiency was specified by *efficiencyProb*; it controlled the probability that a CYP would function as intended. Its value was set initially to *efficiencyProb* = 0.95. I decreased *efficiencyProb* as the simulation progressed using an

exponentially decreasing scaling factor to "slow" down the process. I added an additional condition: the minimum value of *efficiencyProb* was specified arbitrarily to be 0.03. The behavior of the version used to achieve SM-3 is graphed in Fig. 4.4D.

These four time-variant mechanisms together helped me achieve SM-2, but not SM-3. My best result using these four time-variant parameters is graphed in Supplement Fig. 4.8. In achieving SM-2, the RISL was able to produce two important features: rapid decline of digoxin level in the perfusate, and rapid accumulation of metabolite early in the experiment.

4.3.6. Further refinement of the consequences of liver injury achieved the most stringent similarity measure

After meeting SM-2, I strove to achieve SM-3. However, after exhaustive exploration of that RISL's parameter space, I failed to find a parameter vector that would move a METABOLITE profile closer to referent values without negatively affecting other profiles. For example, when I added more ENZYMES to S3 in an attempt to move the RIFAMPICIN PREDOSE METABOLITE time course up, the METABOLITE curves for DIGOXIN alone and for the QUINIDINE PREDOSE moved further up and outside the target range. Inspection of the original wet-lab data showed that, although the rate of digoxin uptake varied greatly due to uptake transporter inhibition by rifampicin, the amount of metabolite that formed early during perfusion was comparable. There was no measurement for metabolite at time = 0. In order to achieve SM-3 it was necessary to introduce further complexity to generate early levels of METABOLITE. As would be expected, doing so necessitated re-tuning most of the other parameter values.

What might have caused the early, high levels of metabolite in perfusate? I arrived at three alternative theories. 1) During pre-dose perfusion, prior to addition of digoxin, mild liver damage caused some cell lysates, including metabolic enzymes, to be released into perfusate. Even if metabolic activity released into perfusate was short-lived, it could have been sufficient to cause some digoxin metabolism. 2) Some digoxin metabolite may have already been present in

the administered source material. 3) An endogenous factor (of hepatic origin) may have co-eluted with the metabolite during assay or cross-reacted with assay reagent. No evidence was available to rule out any of the three and detailed duplication of the same experiments was impracticable. Assertion of any one theory was expected to enable achieving SM-3. Because liver damage can release the contents of some hepatocytes into the perfusate, I elected to assert and explore that theory first. As I show below, doing so enabled achieving SM-3. Consequently, there was no reason to find parameterizations that would enable the other two theories to also achieve SM-3. To explore the plausibility of the first theory, I added different numbers of differently parameterized ENZYMES to PERFUSATE prior to addition of DIGOXIN, and tested the consequences of a wide variety of parameterizations. The following specifications contributed to achieving SM-3. I added just four CYP to PERFUSATE, about 13% the total CYP in the S3. To reflect the different environment of these ENZYMES, I arbitrarily specified different ENZYME parameter values that were lower than CELLULAR ENZYMES: METABOLIC *efficiencyProb* = 0.3 (rather than 0.97) and *sitesN* = 12 (rather than 15). I specified a much higher *defunctProb* (0.025 rather than 0.0025). With that small addition and re-tuning of other parameters, I achieved SM-3. The results are presented in Fig. 4.3. The time-variant parameters are specified in Table 4.5; their values are graphed in Fig. 4.4.

4.3.7. Significance of time-variant mechanisms

A crucial assumption when experimenting with perfused livers is that, absent perturbing events that are part of the experimental design, hepatic functions will remain effectively constant over the duration of the experiment. By being forced to switch from time-invariant to time-variant mechanisms, I was representing a perfused liver whose behavior was inconsistent with that assumption. By turning off all injury-induced, time-variant mechanisms, and executing experiments, I observed plausible results for “normal” livers having constant function throughout the experiment. Those results are graphed in Fig. 4.5. In Supplement Figs. 4.10 and 4.11, I

present eight examples of the consequences of turning off different aspects of the time-variant mechanisms. Together, these sets of RISL time-series data show how each time-variant mechanism contributes in distinctive ways to differences between Figs. 4.3 and 4.5.

4.3.8. Two concurrent inhibitors

To demonstrate the ability to do de novo in silico experimentation with the RISL, I executed the RISL with both inhibitors pre-administered. I observed an anticipated profile of the interaction: the anticipated time-course lay between the no-predose and the rifampicin-predose curves. I believe it is a reasonable prediction because interference with a more up-stream process (uptake) would produce a greater impact of disposition than that of a more down-stream process

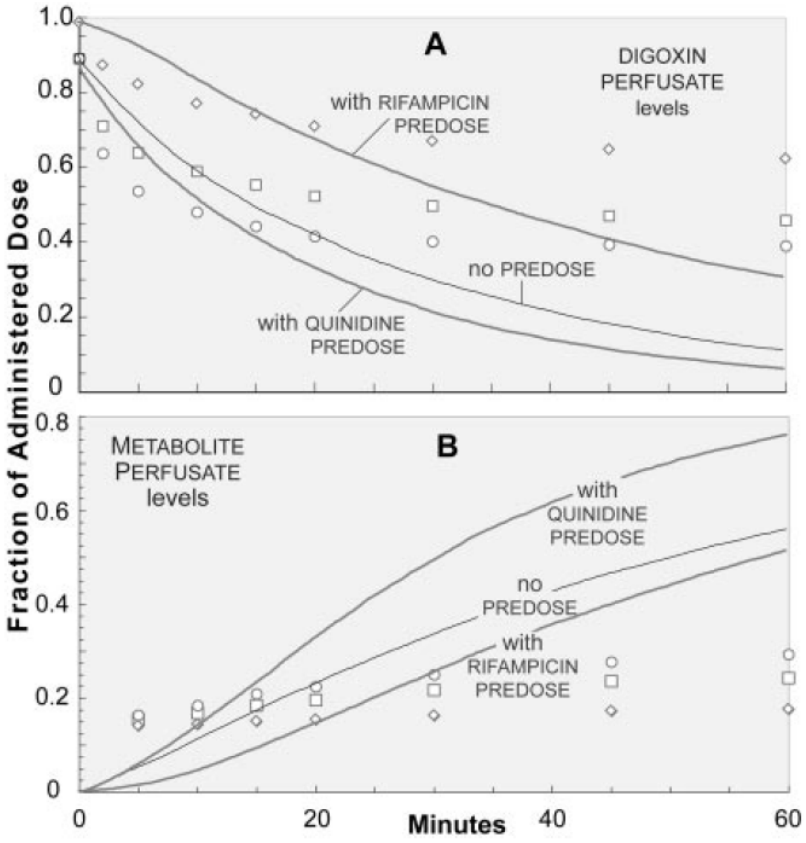


Figure 4.5. Simulation results from an "ideal experimental liver." This RISL Shown are mean results (n = 12) for an RISL and experimental conditions that are the same as in Fig. 4.3, except that all time-variant parameter changes in Fig. 4.4 were turned off and no ENZYME was added to PERFUSATE.

(efflux). The result is presented in Fig. 4.3.

Discovering the mechanisms used for Fig. 4.3 was an iterative and exploratory process that involved induction from both the referent data and—*importantly*—the structure and behavior of prior RISLs. The result is an example of a hybrid inductive–synthetic model that used a middle-out inductive strategy, appropriate for model usage.

4.4. Discussion

I first discuss characteristics of the RISL class of models relative to traditional PK models and then conclude with observations on the results. An objective of perfused liver experiments is to gain new knowledge regarding details of hepatic disposition and metabolism. Hypotheses about those details are induced from the data. Fitting inductive mathematical models to the data is often used as evidence in support of particular hypotheses. To date, designing and conducting new wet-lab experiments has been the only practicable means to falsify experimentally those hypothesized mechanisms. The methods presented provide an additional experimental means of discovering and testing the plausibility of drug interaction details even when the mechanistic details change during the course of experiments.

A purpose in conducting experiments that provide time course data is often to shed light on prevailing mechanistic hypotheses about the dynamics. The methods described here provide a means to leverage the investment in those experiments by constructing and studying mechanistic analogues contemporaneously with wet-lab experiments. A traditional, inductive, PK model hypothesizes an explanation of patterns in PK data (Rescigno, 2001). The mathematics of physiologically based PK models describe data features predicted to arise from conceptualized mechanisms, which are typically described in sketches and prose. There is an unverifiable, conceptual mapping between equations and envisioned mechanisms, as illustrated on the left side of Fig. 4.1. The RISLs and the methods used are different. The synthetic modeling and simulation method (Hunt et al., 2006; Fisher and Henzinger, 2007; Hunt et al., 2008), enables one

to assemble software components into mechanisms as illustrated in Fig. 4.1. They provide an independent, scientific means to challenge, explore, and better understand any inductive mechanism and, importantly, the assumptions on which it rests.

Measures of RISL executions such as Fig. 4.3 provide a test of the mechanistic hypothesis instantiated in the RISL. An acceptable similarity between in silico and wet-lab data, mapping A in Fig. 4.1, is evidence that mapping B may exist between the dynamics occurring during simulation and corresponding dynamics thought to occur within perfused livers, even though the actual events and processes in the two systems are different. That mapping can be realistic, consistent with discretizations and the degree of abstraction. To the extent that mapping B is realistic, I can posit that the implemented mechanisms may have also had biological counterparts: mapping C.

RISL design features make it easy to reuse a mechanism in developing an explanation of other experimental details, past or future. As additional data or knowledge becomes available, the components comprising the mechanisms on the right side of Fig. 4.1 can be reused, much as I have done here, to explore alternative mechanisms and improve realism. Experimenting with RISLs enables the scientist to better understand the analogue mechanisms, with the expectation that that improved insight can translate to the referent system.

RISLs are designed to be relational for maximum flexibility. Their mechanisms are abstract analogues of their referents, not precise one-for-one renderings. For mapping A in Fig. 1 to be quantitative, as in Fig. 4.3, an additional model—a method of scaling—is needed to relate RISL observables directly to the wet-lab PK data. To make mappings B and C semiquantitative, an additional translational model is needed for each. By making the RISL relativistic, I can keep translational models separate from the RISL. If I was to move any one of these scaling models into the RISL, I would immediately reduce the RISL's flexibility, which is scientifically undesirable. The levels of temporal, spatial, and mechanistic granularity (which control

resolution) are somewhat arbitrary: they need to be sufficiently fine so that the measured consequences of RISL execution meet a stringent SM. Granularity can be easily increased or decreased when that is needed. Because everything within the RISL is relativistic, an algorithm can be implemented when needed to automatically adjust parameter values to accommodate new levels of granularity so that the consequences of mechanisms can remain essentially the same.

The synthetic method of modeling and simulation is not intended for making quantitatively precise predictions about observables. Traditional, equation based PK models can already do that very well. The synthetic method is best suited for studying the observable consequences of interacting components and testing hypotheses about mechanisms. As demonstrated here, because any number of distinctly different compounds can be studied within the same experiment, RISLs and their methods are ideally suited for studying mechanisms of drug-drug interactions, and the results of this study have demonstrated their utility for doing that.

From a simulation perspective, it is demanding to expect one RISL mechanisms to simulate time course data for DIGOXIN and its METABOLITE that closely matches observed data from three different treatments. Because all of the mechanistic details are intertwined, a mechanistic change made to improve one profile automatically changes the others. Even though the RISL was relatively (to traditional PK models) fine-grained and hierarchical, having an RISL with only time-invariant mechanisms was too simple: it achieved SM-1, but not SM-2. Explorations of more complicated mechanisms lead to adding a METABOLITE TRANSPORTER to S2, including ENZYMES in PERFUSATE, and making the four time-variant changes presented in Fig. 4.4. 1) Accessibility (transitional state) of RIFAMPICIN, QUINIDINE, and DIGOXIN to HEPATIC surfaces (S2) decreased. 2) CYP became less efficient. 3) The probability that an OAT, PGP, or CYP would remain active in the next simulation cycle declined; and 4) the relative size of the neighborhood in which RIFAMPICIN, QUINIDINE, and DIGOXIN were visible to OAT, PGP, and CYP shrank. When implemented together within a RISL, I identified a parameterization that enabled

successfully achieving SM-3 (Fig. 4.3). I submit that analogous changes in hepatic mechanisms occurred during the referent experiments: the livers were dying and/or increasing portions were being poorly perfused. Having the evidence presented influences how we think about and interpret the wet-lab results. Comparable insight cannot be achieved from purely inductive PK models. That is because components of traditional, physiological based PK model are concepts. As such, and as illustrated on the left side of Fig. 4.1, there can be no concrete mappings between model components, their parameterizations, and referent counterparts (Rescigno, 2004).

An important underlying assumption when conducting liver perfusion experiments is that liver function remains relatively constant over the course of experiments. An important observation is that it was necessary to posit time-variant mechanisms in order to account for the data in Fig. 4.3. Together, the specific time-variant mechanisms in Fig. 4.4, in the context of the RISL analogue in Fig. 4.2, comprise the one abstract explanation for the referent data that I discovered. Different variations on this same mechanistic theme may provide equally satisfactory explanations. If I were to use a different liver analogue, the In Silico Liver in Yan et al. (2008a) for example, alternate, similarly plausible time-variant mechanisms may be discovered. How mechanistic explanations may change between different models of the same system is an avenue for future research. In addition, there may be several, somewhat different parameter vectors for the implemented mechanisms that do equally well at achieving SM-3, but my search did not locate them. Taken together, the following five observations are circumstantial evidence that the mechanistic explanations offered are plausible. A single, time-variant RISL produced: 1) simulated DIGOXIN and METABOLITE PERFUSATE levels that were quantitatively similar to referent data when digoxin was administered alone; 2) Simulated DIGOXIN and METABOLITE PERFUSATE levels following separate pretreatments with QUINIDINE and RIFAMPICIN achieved SM-3; 3) Measures of Monte Carlo variants of the RISL used for Fig. 4.3 had the same time course features observed in the wet-lab data. However, because of the built-in stochastic

uncertainty, they lack the precision of traditional PK models; 4) Simulated QUINIDINE and RIFAMPICIN data were similar to referent time course data (see Supplement Fig. 4.6); 5) Modest changes in parameter vector values altered the relative positions of the six simulated perfusate profiles (not shown), but not their shapes. The RISL can therefore stand as an abstract analogue of what may have occurred during the liver perfusion experiments.

4.5. Supplement

4.5.1. Fraction of administered dose calculations

In wet-lab experiments, the investigator selects specific experiment features to measure because observing all potentially important aspects of the experiment is infeasible. In the wet-lab experiments, the measure of the time course of relative abundance of administered drug was concentration, because only a tiny perfusate aliquot would be consumed. It is not uncommon for a small portion of administered dose to be lost and not recovered in later measurements.

An obvious advantage of RISL experiments is that the system is completely transparent. We can "see" where everything is and what each component is doing at the end of any simulation cycle. We can measure the exact relative abundance of administered COMPOUNDS, as fraction of total administered DOSE.

In order to compare referent and RISL PK profiles, we needed to either transform wet-lab concentration measurements into fraction of administered dose in perfusate, or transform RISL measurements to concentrations. The latter would have been complicated. I elected the former because it required only specifying a common denominator, in units of nanomolar, for all concentration measurement.

There are several options in selecting that common denominator:

- It was reported that 10 μg of digoxin was used, and that the volume of perfusate used was 110 mL. That gave 116 nM.
- It was also reported that digoxin, dissolved in sufficient diluent, was added to the perfusate to give an initial concentration of about 110 nM.
- Finally, it was reported that the initial, perfusate concentration without the rat liver was about 121.5 nM.

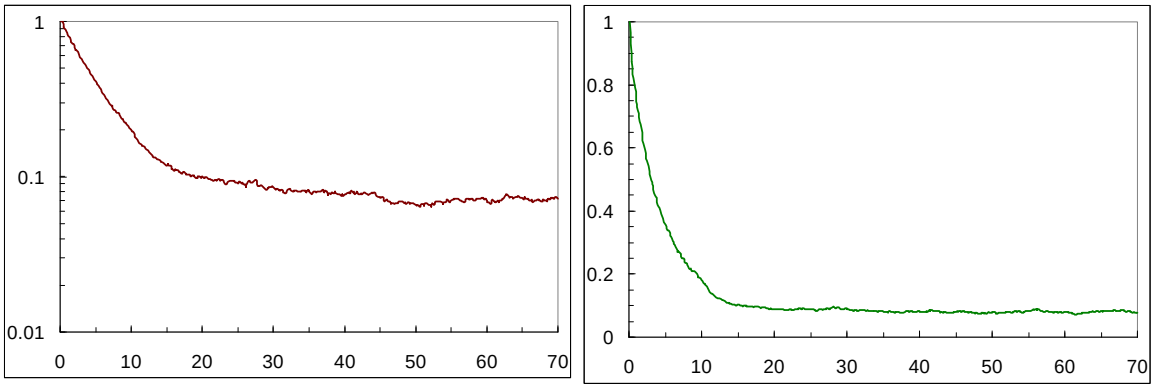


Figure 4.6. RISL PK profiles for RIFAMPICIN and QUINIDINE. X-axis: time after addition of INHIBITOR; y-axis: fraction of administered dose. **A:** RIFAMPICIN, the uptake inhibitor; **B:** QUINIDINE, the efflux inhibitor. Each profile is similar to its referent in Fig. 3 of Lau et al. (2004). DIGOXIN was administered 10 minutes after the INHIBITOR.

I elected to use 121.5 nM as the common denominator because it required making no assumptions about the actual volume being used. Using a different common denominator would shift the values to which I am comparing my simulation results and, in turn, change the assessment of whether or not a given set of RISL results meets a SM.

4.5.2. Inhibitor time-course profiles

It was reported that each inhibitor was added (separately) to the perfusate of an already-established, perfused liver 10 minutes prior to addition of digoxin. Figure 3 in the Lau et al. (2004) shows the reported PK profiles for each inhibitor. Presented in Fig. 4.6 are separate time-course profiles for the two INHIBITORS during a typical simulation experiment. They are qualitatively and quantitatively similar to the reported PK profiles. I elected not to fine-tune these profiles further and not to include them in the set of targeted attributes, because doing so would not have provided additional insight into the referent data in Fig. 4.3. I assert that the observed qualitative and quantitative similarities between the data in Fig. 4.6 and the referent data are adequate for the specified objectives.

4.5.3. A simple hypothesis fails

Fig. 4.7 shows results of the simple, linear RISL. It satisfied the SM-1 criterion, but failed to satisfy SM-2; that RISL's composition and time-invariant parameter values are given in Table 4.6. It included OAT, PGP, and CYP. All parameter settings and relative ratios and amounts of OAT, PGP, and CYP were time-invariant. No parameter vector was found that was able to provide adequate similarity at early times (and thus possibly achieve SM-2) while still meeting SM-1.

4.5.4. Metabolism in the simple time-invariant RISL

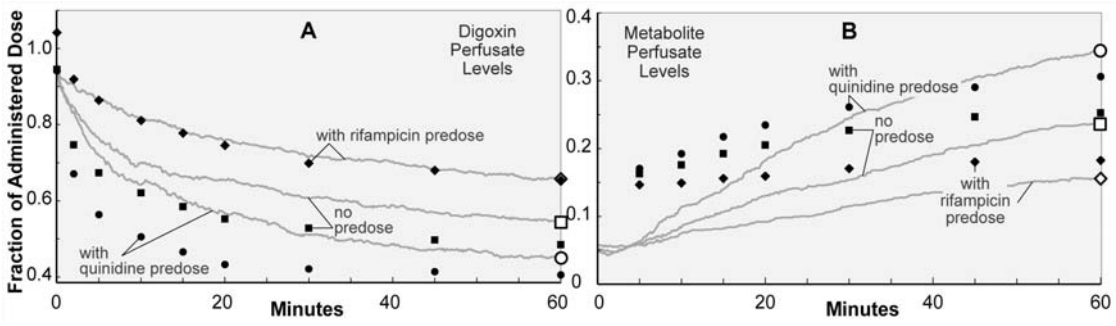


Figure 4.7. Comparison of wet-lab experimental data to simulated results using the simple, linear RISL described in the text. All graphed values are averages of six simulation experiments. All RISL mechanistic events were time invariant and linear relative to amounts. Key parameter values are listed in Table 4.6. Those not shown are the same as in Table 4.2. **A**: Solid symbols are values reported in Lau et al. (2004) converted to fraction of digoxin dose (0.4 to 1.0) remaining in the perfusate for each of the three treatments. The three open symbols at 60 min., along with the gray curves, are mean results for each of the three indicated treatments, simulated using the simple RISL. The simulated values at 60 min. (but not at earlier times) achieved the first Similarity Measure (SM-1). **B**: The closed symbols and treatments are the same as in **A**, but the values are the digoxin metabolite levels in perfusate, as fraction of reported dose (0.0 to 0.4). As in **A**, the three open symbols at 60 min., along with the gray curves, are mean results for each of the three treatments simulated using the simple RISL. The simulated values at 60 min. (but not at earlier times) achieved SM-1. However, as described in the text, no time-invariant parameterization of this RISL was found that enabled achieving SM-2.

To achieve SM-2, I speculated that some level of metabolite may have been present in perfusate shortly after digoxin was added (t = 0). One plausible cause was that hepatocytes were injured during earlier manipulations to establish the system, resulting in metabolic enzymes being released from hepatocytes into perfusate prior to the time digoxin was administered. This mechanism was later implemented. Other plausible causes are: 1) presence of analytically indistinguishable impurities; 2) digoxin hydrolysis produced measurable amounts of primary metabolite (formation of the metabolite from hydrolysis of digoxin is acid catalyzed); 3) a liver

Table 4.6. Key parameter values for the simple time-invariant RISL in Fig. 4.7

Parameters	Remarks		Values
numPgps	Number of PGP		25
numOats	Number of OAT		50
numCyps	Number of CYP (ENZYMES)		15
numBinders	Number of BINDERS		25
timestepfactor	Scaling factor for passive permeation		10
assocProb	DIGOXIN'S association probability to	CYP	0.15
		OAT	0.7
		PGP	0.5
		BINDER	0.15
releaseProb	DIGOXIN'S dissociation probability from	CYP	0.2
		OAT	1
		PGP	1
		BINDER	0.15
closeToInterface	Fraction sufficiently close to interface such that it is *transitional* for DIGOXIN	S1	0.01
		S3	0.08
		S5	0.1
maxSites	Number of substrate binding sites per stationary object		4
efficiencyProb	Efficiency parameter for stationary objects		1
sitesN	Size of neighborhood (number of ELEMENTS) scanned each cycle by a stationary object		4

derived contaminant co-elutes (HPLC) with the metabolite during assay. In Fig. 4.7, I specified that 5% of the administered dose was metabolite at t = 0. Consistent with that hypothesis, I decreased the initial concentration denominator from 121.5 nM to 115 nM. That change explains why wet-lab values in Figs. 4.7, 4.8 and 4.3 differ slightly. Figure 4.7 is the evidence that the preceding hypothesis helps improve similarity sufficiently so that SM-1 is achieved. I noted, however, that the shapes of the profiles (RISL and wet-lab) were dissimilar, and so I began to explore time-variant mechanisms.

4.5.5. Implementing mechanistic deterioration improved similarity

Results following RISL implementation, refinement, and parameter tuning of four time-variant changes achieved SM-2 are presented in Fig. 4.8. Time-variant RISLs differ in four ways from the time-invariant RISLs used to generate the simulation results in Fig. 4.7. A METABOLITE-specific TRANSPORTER was added to S2 to increase the rate of return at early times of METABOLITE to PERFUSATE. The accessibility of S2 to mobile objects in the PERFUSATE decreased with time as specified in section 4.3.5 and Tables 4.7–4.9, decreasing the rate of both passive and active HEPATIC permeation by all mobile objects (the three DRUGS). The efficiency of CYP decreased with time as graphed in Fig. 4.9B. Loss of efficiency suggests loss of viability or diminished perfusion or some combination. In addition, a portion of those immobile objects became inactive (Fig. 4.9C) each cycle. All other parameters were readjusted to achieve SM-2.

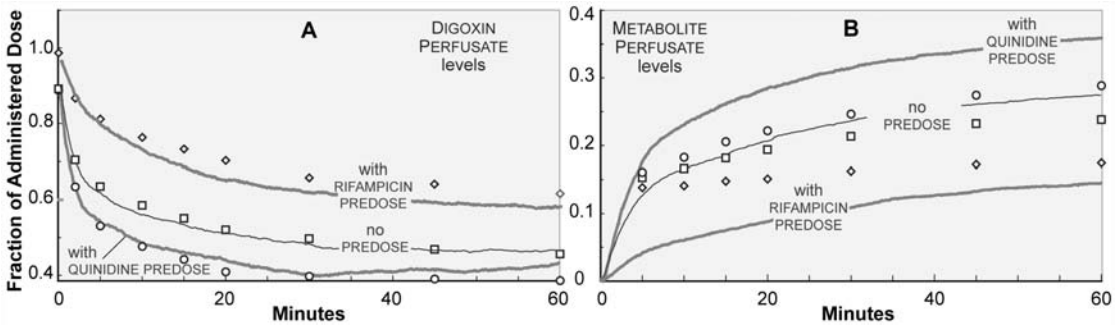


Figure 4.8. Simulated results that achieved SM-2, but not SM-3, using RISLs with the time-variant mechanisms graphed in Fig. 4.9. The symbols are the mean wet-lab data. The curves are mean (n = 6) simulated values. **A**: Shown is the mean DIGOXIN PERFUSATE levels for each of the three indicated treatments. In all three cases, simulated values achieved SM-2; most were within 10% of referent values. For the QUINIDINE PREDOSE case, DIGOXIN PERFUSATE levels, as fraction of DOSE, reached a minimum of about 0.4 at 30 minutes, and then began increasing slightly. That was caused by DIGOXIN returning primarily from S3 because of eroding TRANSPORT and METABOLIC functions. **B**: Shown are the mean METABOLITE PERFUSATE levels corresponding to the three treatments in **A**.

The parameter values are presented in Tables 4.7, 4.8 and 4.9.

I achieved an early decline in DIGOXIN levels along with rapid rise in METABOLITE level, both of which were not achieved using the time-invariant RISL (Fig. 4.7). Even though METABOLITE levels for both the RIFAMPICIN and QUINIDINE PREDOSE cases are far from target values, the shapes of the METABOLITE time course profiles were similar to the shapes of the referent profiles. Collectively, the six sets of simulated data in Fig 4.8 achieved SM-2. Further improvement was infeasible without making the time-variant mechanism more complicated, which was done to achieve the RISL results in Fig. 4.3. A change in parameterization that would bring the RIFAMPICIN-PREDOSE curve closer to referent values (open diamonds) would cause each of the other five profiles (and possibly the RIFAMPICIN alone and the QUINIDINE alone profiles) to also shift. The net result of such an effort can be failure to achieve SM-2.

Table 4.7. Compound specific parameters values for the RISL data in Fig. 4.8

COMPOUND Parameters	Remarks		Values			
			DIGOXIN [a]	RIFAMPICIN [a]	QUINIDINE [a]	METABOLITE [b]
MW [c]	molecular weight of the solute		781	822	324	601
logP [c]	logP - common logarithm of octanol/water partition coefficient		1.14	3.60	2.53	1.60
pKa [c]	pKa closest to *insilicoPH*		13.5	7.9	8.6	13.5
inihibitorType	uptake inihibitor? True/False		n/a	TRUE	FALSE	n/a
closeToInterface	Fraction sufficiently close to interface such that it is *transitional* [d]	S1	-- [e]	0.750	0.750	0.030
		S3	0.125	0.040	0.040	0.060
		S5	0.100	0.050	0.050	0.100
transitProb	Initial estimates of trans-membrane transit probabilities, after scaling by *timestepfactor*	S1→S2	0.11	0.11	0.16	0.13
		S2→S1	0.19	0.014	0.10	0.11
		S2→S3	0.19	0.012	0.14	0.11
		S3→S2	0.11	0.11	0.16	0.13
		S3→S4	0.11	0.11	0.16	0.13
		S4→S3	0.19	0.012	0.14	0.11
		S4→S5	0.19	0.012	0.16	0.11
		S5→S4	0.11	0.11	0.16	0.13

[a] Physicochemical properties reflect those of the referent drug.
[b] Physicochemical properties reflect those of digoxigenin bis-digitoxoside.
[c] Molecular weight, logP, and pKa are used in calculating the initial, passive transit probabilities, as described earlier. For *pKa*, only the value closest to 7.4 is listed.
[d] The *closeToInterface* value for all solutes in S2 and S4 was always 1. This reflects the fact that the referent barriers are sufficiently thin so that *all* solutes within them at the start of a 10 second interval (a simulation cycle) will have had an opportunity to transition by the end of that interval.
[e] *CloseToInterface* is a time-variant parameter. It, together with *internalProb*, specifies the *fraction transitional* (Table 4.9).

Table 4.8. Parameters for compound-protein interactions for the RISL data in Fig. 4.8

Compound-Protein Interaction Parameters		**Values**			
		Digoxin	**Rifampicin**	**Quinidine**	**Metabolite**
isaSubstrate substrate of ? TRUE / FALSE	Mt	FALSE	FALSE	FALSE	TRUE
	Pgp	TRUE	FALSE	FALSE	TRUE
	Oat	TRUE	TRUE	FALSE	TRUE
	Cyp	TRUE	FALSE	FALSE	FALSE
	Binder	FALSE	FALSE	FALSE	FALSE
assocProb reflects binding affinity to protein	Mt	0	0	0	0.9
	Pgp	0.4	0	0.99	0.4
	Oat	0.5	0.8	0	0.05
	Cyp	0.5	0	0	0
	Binder	0.6	0	0	0.24
releaseProb reflects dissociation from protein	Mt	1.0[a]	1.0[a]	1.0[a]	0.9
	Pgp	0.7	1.0[a]	0.05	0.7
	Oat	0.7	0.075	1.0[a]	0.9
	Cyp	0.2	1.0[a]	1.0[a]	1.0
	Binder	0.3	1.0[a]	1.0[a]	0.04

Table 4.9. Time-variant parameters and their values, corresponding to Fig. 4.9

Time-variant parameter			**Initial value**	**Terminal value**
Effective surface area changes	*Fraction transitional*[a] in S1		0.35	0.096
	sitesN (neighborhood)	Cyp	15	2
		Pgp	9	4
		Oat	20	4
		Mt	35	4
Metabolic and transport activity changes	*efficiencyProb*[a]	Cyp	0.99	0.05
		Pgp	1	1
		Oat	0.6	0.6
		Mt	1	1
	defunctProb	Cyp	0.0025	0.023
		Pgp	0.00028	0.0026
		Oat	0.0035	0.032
		Mt	0.00028	0.0026

[a] The method for changing *fraction transitional* and *efficiencyProb* (Fig. 4.9A and 4.9B) is detailed in chapter 3, under Parameter scaling.

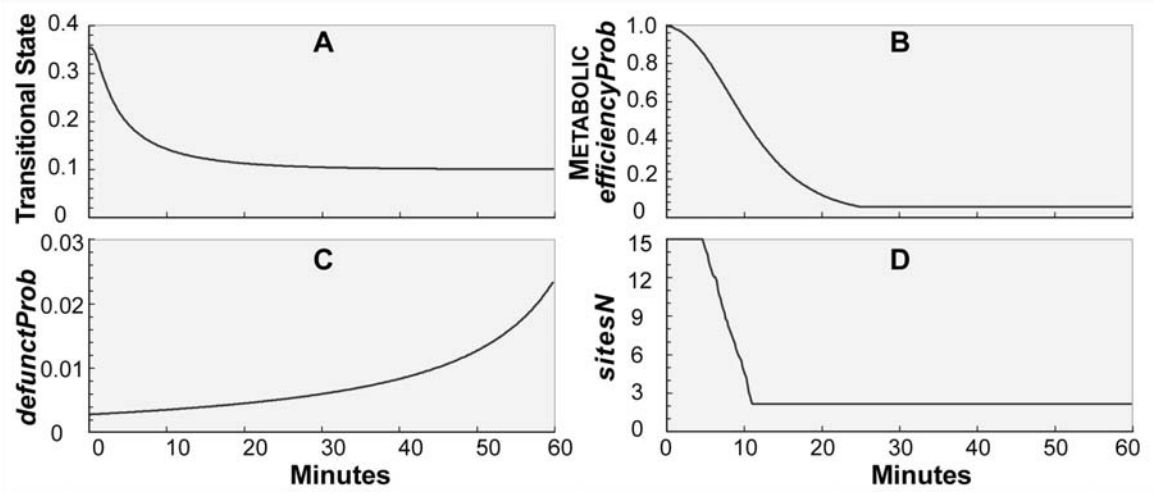


Figure 4.9. Values of time-variant parameters for the RISL data in Fig. 4.8. Each graph shows time-dependent parameter values used to generate the experimental results shown in Fig. 4.8. **A**: The graph shows the time-variant values of the fraction of COMPOUNDS in S1 that were in the transitional state. COMPOUNDS in S1 were in one of the three states discussed in the text: 1) *external* (to the LIVER), 2) *internal* (within EXTRACELLULAR HEPATIC SPACES but not given the option to transition within a simulation cycle, or 3) *transitional* (close enough to CELL MEMBRANES—S2—so that transition was an option within a simulation cycle). The fraction external needed to be constant to simulate a constant perfusate flow rate. During tuning, an algorithm was used to adjust the final value and the rate of decline. The values for fraction internal were automatically adjusted. The initial value of fraction transitional was 0.35; the value at 60 minutes was 0.096. **B**: The graph shows the time-dependent values of (METABOLIC) *efficiencyProb*. The frequency at which METABOLISM occurs for a COMPOUND bound to CYP is controlled by the parameter (METABOLIC) *efficiencyProb*. The graph shows the decrease in its value (initially 0.99) to a minimum value of 0.05. During tuning, an algorithm was used to adjust the final value along with the early and subsequent exponential rate of decline. *EfficiencyProb* = 1 for PGP, and it did not decrease with time; *efficiencyProb* = 0.6 for OAT, and it too remained constant. **C**: The graph shows the time-dependent values of *defunctProb* for CYP; the pattern of change was the same for PGP and OAT. During tuning, I used the following quadratic relationship. The initial values used were *defunctProb* = 0.0025 (CYP), 0.00028 (PGP), and 0.0035 (OAT). In each case, the *defunctProb* value at 60 minutes was about 10 times that of the initial value. Quadratic relationship: $defunctProb(t) = initial\_defunctProb \bullet (defunctRate + t)/(defunctRate - t)$, for $t \leq 60$ MINUTES, $defunctRate = 500$. **D**: The graph shows the mean time-dependent values of *sitesN* for CYP averaged over 10 CYP; the pattern of change was the same for PGP and OAT. Factors such as accessible hepatocyte surface area and effective cellular volume were simulated as described in section 4.3.5. The initial value each object's (OAT, PGP, CYP, and MT) *sitesN* was decremented randomly. The average decrement was one ELEMENT every three simulation cycles. The initial and minimum *sitesN* values were 15 and 2 for CYP, 9 and 4 for PGP, and 20 and 4 for OAT. Table 4.9 shows the values for each of the time-variant parameters.

4.5.6. Consequences of selectively turning off mechanistic features

I focus on four conditions, corresponding to the answer to four questions. **E**) **E**nzymes: are ENZYMES in S1 beginning at $t = 0$? **S**) **S**urface area: are COMPOUND accessible HEPATIC surface areas and access to CYP, PGP, and OAT shrinking as in Fig. 4.4 A and B? **D**) **D**ysfunction: are ENZYMES and TRANSPORTERS becoming less efficient and defunct as in Fig. 4.4C and D? **B**) **B**inders: are BINDERS present? The key function of BINDERS is to retain METABOLITE within CELLS (within S3), a process that is inhibited by RIFAMPICIN in the RISL.

In Figs. 4.10 and 4.11, the consequences of eight of 16 possible combinations are shown. The combinations are coded as follows. When the answer to all four questions in the above order (ESDB) is yes, then that RISL is identified as 1111. That corresponds to the RISL used for Fig. 4.3. The "healthy" RISL in Fig. 4.5 is identified as 0001. All RISLs in Figs. 4.10 and 4.11 should be compared to these extremes as well as to each other.

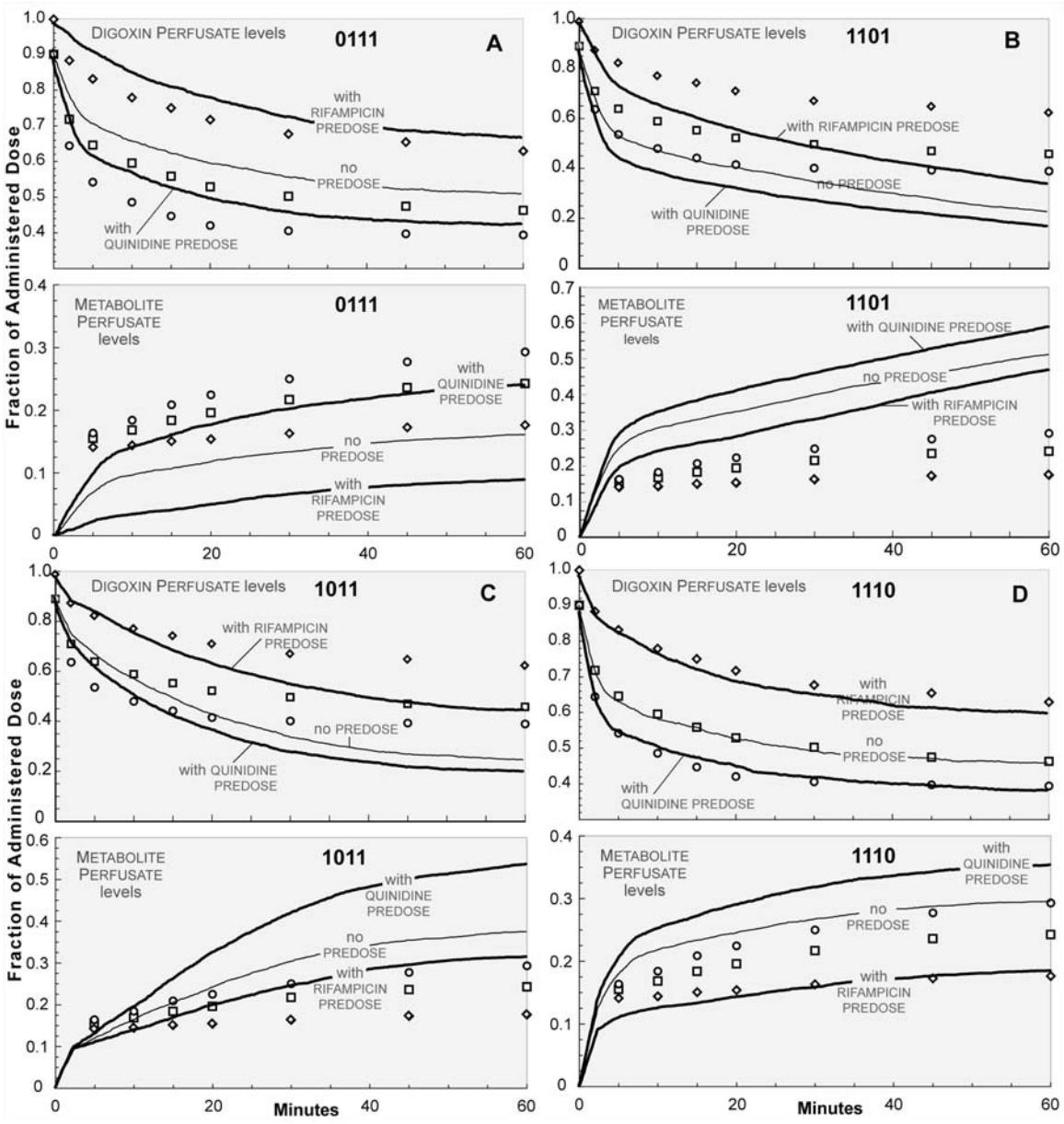


Figure 4.10. Perfusate profiles for four different RISLs. **A:** ESDB = 0111. Shown are the consequences of not having ENZYMES in S1. There was always too much parent DRUG, and too little METABOLITE. ENZYMES in S1 convert some DIGOXIN to METABOLITE early on. **B:** ESDB = 1101. Shown are the consequences of keeping S1 and S3 ENZYMES active for the duration of the experiment. METABOLITE formation does not level off, because ENZYMES are not dying out. However, their neighborhood (*siteN*) still shrinks. **C:** ESDB = 1011. Shown are the consequences of ENZYMES and TRANSPORTERS becoming less efficient and defunct, but their neighborhoods do not shrink. The initial drop of DIGOXIN become much less drastic, and correspondingly fewer METABOLITE return to perfusate at early time **D:** ESDB = 1110. Shown are the consequences of having no BINDERS. DIGOXIN PERFUSATE levels were very much the same, but without BINDERS, more METABOLITE entered PERFUSATE in DIGOXIN-only and QUINIDINE-only groups.

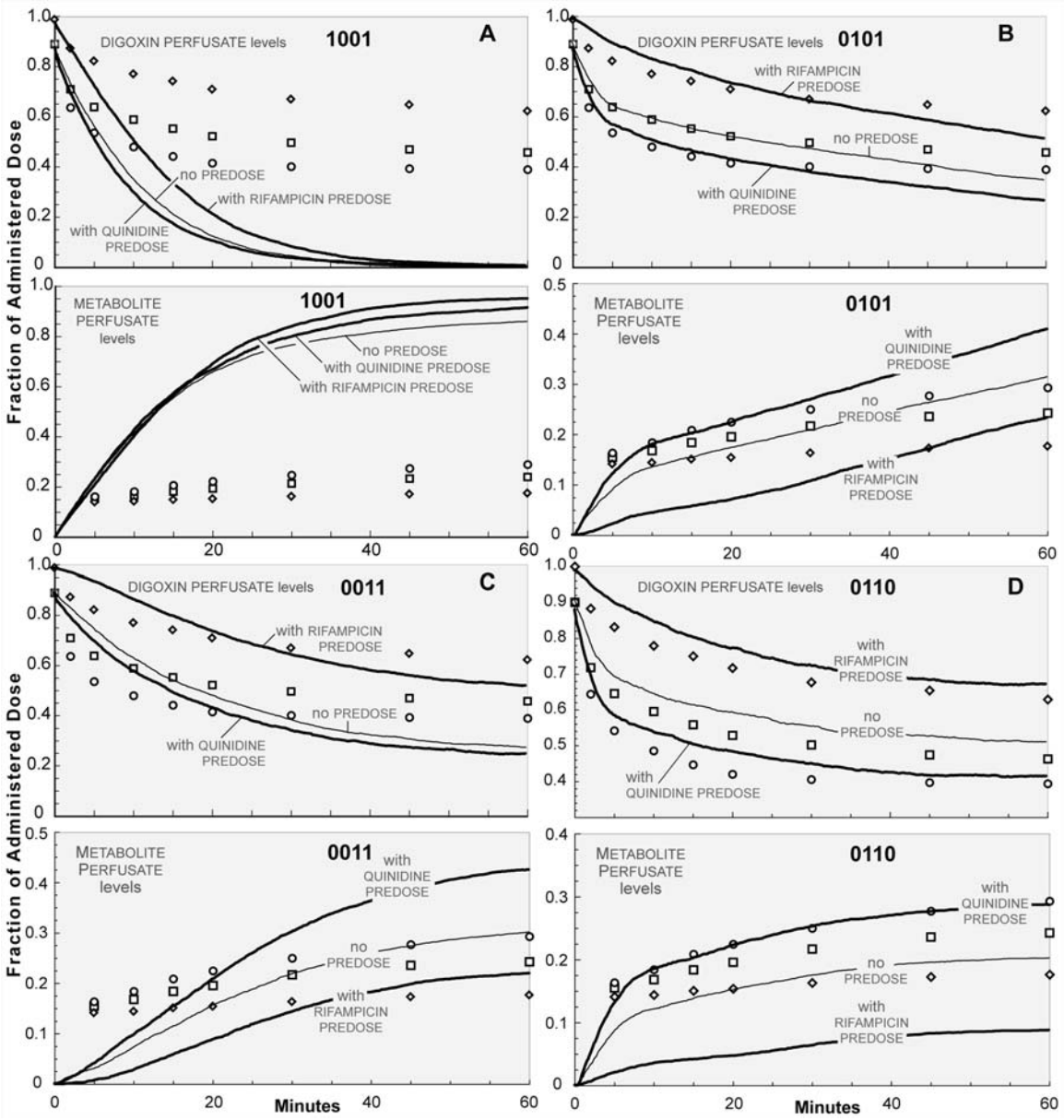


Figure 4.11. Perfusate profiles for another four different RISLs. **A:** ESDB = 1001. The most dramatic change was when ENZYMES and TRANSPORTERS stayed active and their access along with accessible HEPATIC surface areas does not shrink. Note that BINDERS were present, so the RIFAMPICIN treatment group had the highest PERFUSATE METABOLITE levels. **B:** ESDB = 0101. Shown are the consequences of keeping ENZYME and TRANSPORTER activities and numbers constant. **C:** ESDB = 0011. Shown are the consequences of keeping constant COMPOUND accessible HEPATIC surface areas and access to CYP, PGP, and OAT. **D:** ESDB = 0110. Shown are the consequences of removing the S1 enzymes and BINDERS.

5. Mechanistic Insight from In Silico Pharmacokinetic Experiments: Roles of P-glycoprotein, Cyp3A4 Enzymes, and Microenvironments

5.1 Introduction

Mouly et al. (2004) reported paradoxical observations following a vectorial study of saquinavir transport across a monolayer of modified, p-glycoprotein (P-gp) and Cyp3A4-expressing Caco-2 cells: there were higher intracellular levels of saquinavir, yet less metabolite formation after apical compared to basal dosing. The data clearly indicated that the intracellular mechanisms during transport were more complicated than anticipated. The authors suggested mechanistic explanations, but no established experimental wet-lab methods were available to test them.

Recent advances in discrete event modeling and simulation (M&S) of complex systems (Hunt et al., 2009) enables implementing in silico methods to test the plausibility of mechanistic hypotheses. I used the synthetic modeling method (Liu and Hunt 2006; Hunt et al., 2006; Garmire et al., 2007; Yan et al., 2008a; Lam and Hunt., 2009; Park et al, 2009) illustrated in Fig. 5.1 to implement an in silico analogue of the wet-lab system, and then used it to explore the

plausibility of the above explanations. The approach was straightforward: drawing on prior efforts and reusing validated components, I built an in silico analogue of the confluent, asymmetric, cell monolayer system used by Mouly et al. It is designed for experimentation and hypothesis generation and falsification (Hunt et al., 2008; Lam and Hunt, 2009). I call it ISECC for In Silico Experimental Caco-2 (cell monolayer) Culture. Concrete, working versions of hypothesized spatial mechanisms were implemented within. In silico experiments used the same design as the referent wet-lab experiments. Simulation following component tuning tested each mechanistic hypothesis. Similarity Measures (SMs) with respect to referent experimental observations were used to set expectations. When results did not meet expectations, the mechanistic hypothesis was falsified; when results did meet expectations, a degree of validation was achieved.

The simple mechanisms hypothesized by Mouly et al. (2004) as implemented were falsified: when instantiated in the ISECC they failed to validate. That was unexpected. Although the proposed mechanisms alone were insufficient, it turned out that they were necessary parts of a more complicated explanation. To discover an ISECC mechanism that would validate, I implemented and followed the iterative ISECC refinement protocol in Fig. 5.2. It targeted wet-lab results from an expanding subset of 16 experimental conditions. I progressed through four stages. By discovering a set of mechanisms that would validate (achieve the prespecified SMs) for two of the 16 conditions, I achieved Stage 1 (Lam and Hunt, 2008). Stage 4 targeted results of all 16 wet-lab conditions coupled with more stringent SMs. In between were many falsification, re-validation processes. The following illustrates. A specific ISECC mechanism is discovered that validates for six targeted conditions using a prespecified SM. The targeted attribute list is then expanded to include transport results for two additional conditions. The current ISECC is falsified because, for the specified SM, experimental results using it over the eight conditions fail to match referent results. I then readjust the current ISECC and/or add in

new components such that the new ISECC validates again. Following the parsimony guideline, validation is achieved by discovering a plausible, marginally more complicated mechanistic hypothesis that survives the experimental challenge. The iterative refinement protocol enabled and facilitated discovery of plausible, new mechanistic details through exercise of abductive scientific reasoning, a primary means of knowledge creation and creative cognition (Hunt et al., 2009). The M&S method also facilitated that discovery by making it relatively easy to implement and explore multiple mechanistic hypotheses.

The ISECC that survived the most stringent SM challenge at Stage 4 produced transport measures that were statistically indistinguishable from referent wet-lab observations. It required a 7:1 ratio of apical transporters to metabolizing enzymes, a 97% reduction of efflux activity by an inhibitor, a biased distribution of metabolizing enzymes, heterogeneous intracellular spaces, and restrictions on drug movement within some of those intracellular spaces.

A purpose of conducting wet-lab experiments like those cited is to gain new knowledge regarding mechanistic details of directional transport and metabolism. As done in the cited work, hypotheses about those details are induced from the data (Rescigno 2004; Hunt et al., 2008). To date, all hypothesized mechanistic explanations for the data have been conceptual and thus difficult to falsify. Designing and conducting new wet-lab experiments has been the only practicable means to experimentally falsify those hypothesized, conceptual mechanisms. Even when inductive mathematical models have been fit to the data, the mechanisms remained conceptual. Experimenting on synthetic analogues like ISECC provides a heretofore-unavailable means of discovering new mechanistic details and testing their plausibility. This approach provides a powerful new expansion of the scientific method: an independent, scientific means to challenge, explore, better understand, and improve any inductive mechanism and, importantly, the assumptions on which it rests (Hunt et al., 2009).

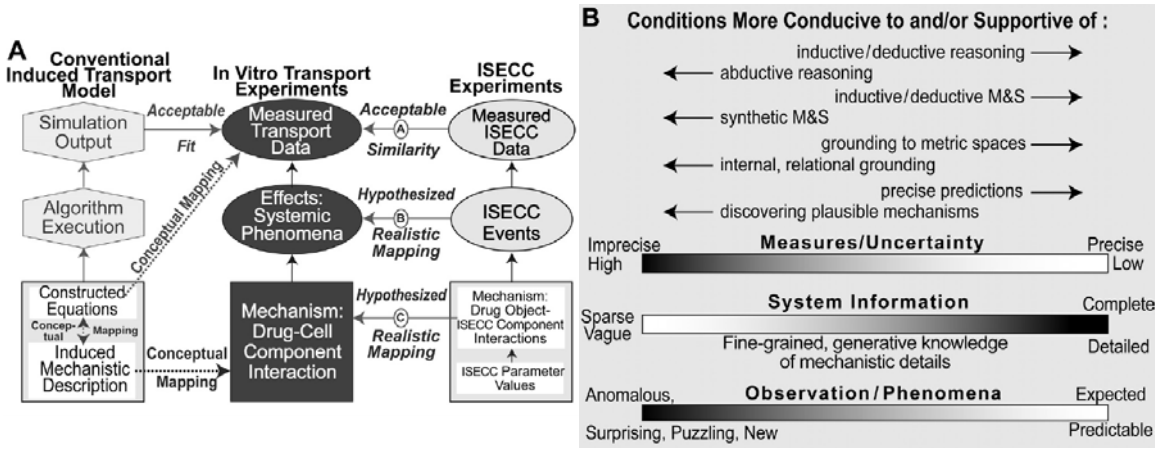


Figure 5.1. Relationships between in vitro, transport experiments, conventional induced transport models, and ISECC experiments. **A**, center: cell monolayers in an experimental context such as a Transwell device are the referent wet-lab systems. During experiments, cellular components interact with transiting drug molecules to cause changes in amount transported and metabolized within specified intervals. Influential mechanistic details are reflected in the collected data. Left: the researcher identifies patterns in the wet-lab transport data. From those and prior transport knowledge, a mechanistic description of what is thought to have occurred is induced, thus establishing an abstract, conceptual mappings from that description to transport mechanisms. In some cases, the researcher goes further and offers a set of transport equations believed capable of describing the data. An algorithmic representation the equations in software is constructed and executed to simulate parameterized equation output. Metrics specify the goodness of fit of the simulated output to the data. Right: the abstract mechanistic description may be different from that on the left side. Software components are designed, specified, coded, verified, and assembled and connected guided by that mechanistic description. The product of the process is a collection of abstract mechanisms rendered in software. A clear mapping—C—is intended to exist between ISECC components and how they plug together, and cell monolayer and intracellular details. Relative similarity is controlled in part by parameterizations. Importantly, mapping C can be concretized iteratively. Compilation and source code execution gives rise to a working analogue. Its dynamics are intended to represent abstractly (mapping B) corresponding dynamics (believed to occur) within the monolayer cultures during an experiment. Mapping B can also be concretized iteratively. Measures of simulated dynamics provide time series DATA that are intended to mimic corresponding measures of wet-lab transport experiments. Achieving increasingly stringent similarity measures enable mapping A to be made concrete. **B**, Conditions supportive of all three reasoning methods are sketched (see Appendix for brief descriptions). Obviously, scientists engaged in drug transport research would like knowledge about all cell monolayer and subcellular characteristics to be rich and detailed, and for uncertainties to be limited. Such conditions (toward the far right side), which are common in non-biological, physical systems, favor developing inductive models that are increasingly precise and predictive. However, for the system and experiments described in Mouly et al. (2004) we are on the left side, where frequent abduction is needed and synthetic M&S methods can be most useful.

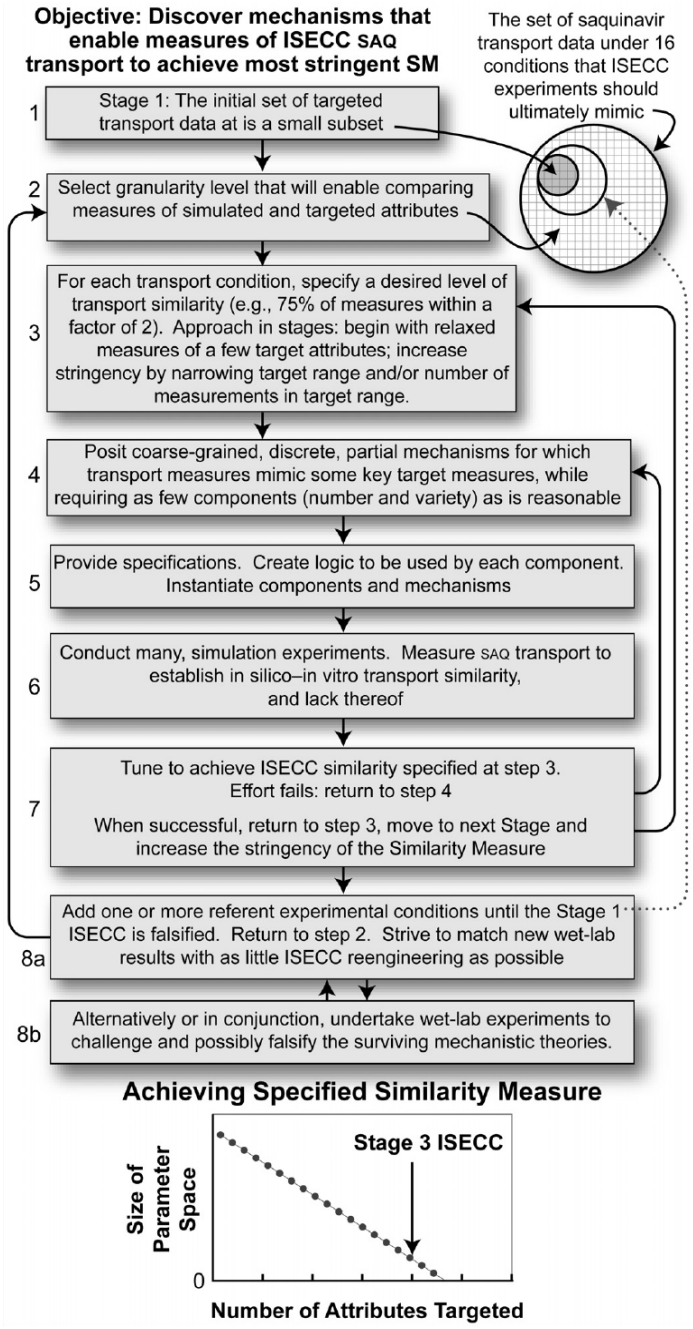


Figure 5.2. The iterative protocol used to refine and improve the ISECC. At the beginning, the set of attributes to be targeted are specified. In this case, that set included results from sixteen different experimental conditions. Similarity goals to be achieved are specified. Abductive reasoning may be required at steps 4-8. Induction and deduction occur during steps 5-7. The final ISECC described herein was preceded by seventeen less complicated ISECC (described in the Fig 5.9). Each in sequence was falsified at step eight. **Bottom**: a graph illustrating the relationship between the parameter space for a specific ISECC that validates and the number of attributes targeted. Indicated is an ISECC that validates by achieving the Stage 3 but not Stage 4 SMs. Its parameter space is small. However, because parameter influences are all networked and the variability in outcomes between simulation runs is non-trivial, there are many similar mechanism and parameter vectors within that space that can produce essentially the same measured simulation phenomena.

If the number of attributes targeted is reduced, then for the same ISECC the size of the space that contains parameter vectors that can validate increases. However, if the number of attributes targeted is increased at step eight in the protocol, that same ISECC is falsified because the parameter space that enables validation shrinks to zero. Adding a component or feature to that ISECC makes a new, more complicated ISECC that has its own curve; it may be shifted up so that, again, there is a region of parameter space that contains mechanisms and parameter vectors that can validate. The relationship is shown linear for illustration purposes.

5.2 Methods

In silico experiments were conducted beginning with a basic structure of the ISECC similar to that detailed in Liu and Hunt (2006) and Garmire et al. (2007). The second generation ISECCs used herein map to a monolayer of epithelial cells separating apical from basal media- or fluid-filled compartments. The 3D nature of the in vitro system is simulated using a stack of five two-dimension toroidal grids. To distinguish clearly in silico components and processes from corresponding in vitro tissue culture counterparts, I use SMALL CAPS when referring to the former; *italics* denote ISECC parameters, variables and internal states.

In contrast to inductive models, which often focus on precise prediction, the ISECC and the synthetic modeling and simulation (M&S) methods used are designed for discovering and testing plausible, mechanistic explanations (Hunt et al., 2009) of the referent drug transport data (Mouly et al., 2004). The approach is ideal for discovery and understanding of transport phenomena produced by a system of interacting components. A specific ISECC instantiates (is represented by a concrete instance of) a mechanistic hypothesis (Fisher and Henzinger 2007; Hunt et al., 2008; Hunt et al., 2009). Execution and comparison of results to referent data tests the hypothesis. A protocol was followed that facilitated generating multiple mechanistic hypotheses. I eliminated the least plausible using in silico experimentation.

5.2.1 Summary of wet-lab methods

Using modified, Cyp3A4 and P-glycoprotein (P-gp) expressing Caco-2 monolayers (Schmiedlin-Ren et al., 1997) cultured in Transwell™ devices, Mouly et al. (2004) studied the role of Cyp3A4-mediated metabolism (Fitzsimmons et al., 1997; Eagling et al., 2002; Parker and Houston 2008) and P-gp-mediated efflux (Kim et al., 1998; Wacher et al., 1998) in saquinavir metabolism and disposition (Su et al., 2004). In control conditions, saquinavir was added to either apical or basal compartment at the start of the experiment to achieve concentrations of 5, 10, 20 or 40 μM. For the P-gp inhibited conditions, an apical dose of the P-gp inhibitor

LY335979 (zosuquidar trihydrochloride; simply inhibitor hereafter) (0.5 μM) (Dantzig et al., 1999) was co-administered with saquinavir. The cells were incubated at 37º C for up to four hours, at which time the apical and basal media, along with the cell monolayer were collected and analyzed for saquinavir and its major metabolite M7.

5.2.2 Objective and approach

My objective was to discover a single set of components and micromechanisms, which when parameterized would provide a cohesive, concrete, plausible explanation for the results of sixteen different, experimental conditions presented in Fig. 1 of Mouly et al. (2004), including the paradoxical observations. I was not seeking a traditional, differential equation explanation of transport. Rather, I sought an actual, working mechanism—*an analogue*—comprised of quasi-autonomous biomimetic processes and parts, which when measured during execution would give TRANSPORT DATA similar to the referent wet-lab data. There are likely many equally plausible mechanisms that differ in some details, yet give rise to essentially the same phenomena. For this project, the goal was simply to discover one. I used the synthetic modeling and simulation method (Hunt et al., 2006; Lam and Hunt, 2009) illustrated in Fig. 5.1. I started with the ISECC from Lam and Hunt (2008), which plugged together validated, quasi-autonomous software components to form an abstract yet mechanistically realistic analogue of transport through a monolayer of cells into which one could add, alone or in combination, objects representing different compounds.

My experimental approach followed the iterative refinement protocol in Fig. 5.2: cycles of ISECC synthesis, testing and evaluation, validation or falsification, assessment, cogitation, and system revision until one satisfied predetermined similarity criteria. Other researchers and I have used the protocol successfully (Hunt et al., 2006; Tang et al., 2007; Engelberg et al., 2008; Kim et al., 2009; Lam and Hunt, 2009). It strives to adhere to the guideline of parsimony, which is important when building agent-oriented analogues that are expected to become increasingly

complex. I began by instantiating the conceptual mechanism described in Mouly et al. (2004). Even after multiple rounds of iterative refinement, it failed to mimic wet-lab results. Failure of that best-at-the-time mechanistic explanation demonstrated that conceptual mechanistic descriptions can be flawed in ways that are not readily apparent. Flaws, when they exist, begin becoming obvious after I implement and begin testing the mechanism synthetically. (Lam and Hunt, 2009; Hunt et al., 2009)

5.2.3 ISECC structure and components

I specified that ISECC components and their assembly be consistent with a Transwell device having a confluent monolayer of polarized epithelial cells exhibiting the following characteristics. There are at least five distinct spaces: two dosing compartments (apical and basal), and an intracellular space between a pair of asymmetric membrane barriers (apical and basal). Added compounds can move within and between spaces consistent with their physicochemical properties (PCPs). Efflux transporters (P-glycoprotein, etc.) located only on the apical membrane enable facilitated translocation of some compounds across that membrane. An added compound in the intracellular space, upon encountering a subcellular component (a cytochrome P450, for example) can be metabolized; however, those subcellular components are not uniformly distributed within the intracellular space. There are regions of intracellular heterogeneity. Examples include, lysosomes, the nucleus, mitochondria, and endoplasmic reticulum (Khoo et al., 2002; Vernochet et al., 2005).

Cells are complicated 3-dimemsional (3D) structures. For the sake of visualization, it is tempting to use a 3D grid to represent cell spaces and place different objects within that map to subcellular features. However, the available detailed knowledge is insufficient to validate. Even with considerably more knowledge, to complete a 3D representation, I would accumulate a long list of weak assumptions. The approach taken strives to avoid unnecessary assumptions while specifically taking into consideration uncertainty and ignorance.

An ISECC uses five identically sized spaces, S1–S5. The mappings are as follows: S1 → apical compartment; S2 → apical membrane; S3 → intracellular space; S4 → basal membrane; and S5 → basal compartment. S1–S5 are two-dimension grids. In this study, each space is arbitrarily subdivided using a 50x50 square grid. Using more coarse-grained spaces increases the variance of measures requiring averaging more simulation runs. Using more fine-grained spaces will decrease variances but run times are increased. Objects called ELEMENTS are placed at each grid location. The S2–S4 elements are containers for CELLULAR components and for COMPOUNDS moving around within and between spaces. S1 and S5 elements are simply containers for COMPOUNDS. Each S3 element, for example, maps to a small fraction of total intracellular space from apical to basal membranes. A small fraction of S1 and S5 elements map to spaces between cells including tight junctions (TJ) (Liu and Hunt 2006; Garmire et al., 2007). Elements can have different properties relative to mobile COMPOUNDS. For example, the SOLUBILITY of a COMPOUND in one element type can be specified to be different than in another element. Elements are given properties so that the COMPOUND's entry and exit from a particular element can be differentiated based on those properties and the COMPOUND's PCPs. For example, a LIPID-LIKE element can restrict entry of highly POLAR or CHARGED COMPOUNDS, whereas highly HYDROPHOBIC COMPOUNDS can accumulate (Khoo et al., 2002; Vernochet et al., 2005). An ELEMENT's internal logic is a placeholder for more fine-grained mechanistic detail that can be added when that is needed and the information is available to do so. Within S1, S2, S4, and S5, the elements at each grid location are identical.

It was apparent from the original wet-lab data (Mouly et al., 2004) and from my group's early work (Yan et al., 2008b; Hunt et al., 2008; Lam and Hunt, 2009) that an ISECC would need to simulate some intracellular heterogeneity. I did so by specifying that the space within a S3 element can be heterogeneous. Because an element is the limit of ISECC resolution, the details of within-element heterogeneity are left unresolved. To facilitate conceptualization, I describe an

S3 element as being subdivided VERTICALLY into SUBCELLULAR microenvironments called *zones* (Lam and Hunt, 2008). Technically, I simply specify that objects within an ELEMENT can be in different states, and behavior is state-dependent. Elements in S1, S2, S4 and S5 behave as if the contents are homogeneous and well mixed. To iteratively improve ISECC validation, it was necessary to increase the number and alter the properties of S3 *zones*. The following paragraph provides a generalized description applicable to all ISECCs.

The *zones* of each S3 ELEMENTS are designated Z*i*, $i = 0$ to n. The parameter *maxZ* specifies n, and for the various ISECCs studied, $2 \leq n \leq 6$. Zones map to intracellular gradients and microenvironments. A zone could, for example, map to a compound transport pathway (Weisiger 1996; Weisiger 2002; Weisiger 2007), nucleus (Tran et al., 2003), mitochondria (Khoo et al., 2002), a portion of the endoplasmic reticulum, etc. Being below the level of resolution, their shape is indeterminate. Z0 is always adjacent to S2 and Z*n* is always adjacent to S4. All other zones can be conceptualized as being layered between those two. I encountered a need for some zones, specifically those "close" to a MEMBRANE—Z0 and Z*n*—to behave as if they were less aqueous, more hydrophobic (Khoo et al., 2002; Vernochet et al., 2005). Consequently, a parameter (*HMspace*) controlled fraction of Z0 and Z*n* zones was designated a HYDROPHOBIC MICROENVIROMENT (HME). The fraction was tuned for each ISECC. For the final ISECC described herein, Z0 in 15% of elements were specified as being HME. In addition, in one-fifth of the elements containing a HME Z0, their Z6 were also specified as being HME.

5.2.4 Mobile and stationary objects

A COMPOUND (or DRUG) is a mobile object that maps to an unspecified number of xenobiotic molecules (Liu and Hunt 2006; Garmire et al., 2007; Yan et al., 2008a; Lam and Hunt, 2009). I use three types: SAQ, M7 and M1. They map to saquinavir and its two metabolites (Eagling et al., 2002). Each COMPOUND is assigned a set of PCPs; they map to physicochemical

properties of the referent compound. Different, stationery objects are assigned and confined to subsets of S2 and S3 elements. I use two types: CYP and PGP. They are discussed below.

During each simulation cycle, each COMPOUND, selected pseudo-randomly (randomly hereafter), has one opportunity to 1) move laterally to a neighboring element, 2) transit to a neighboring space, and 3) if in S3, to relocate to a different zone in the same element. COMPOUND transition within and between spaces is parameter specified and uses validated algorithms described in chapter 3, (Lam and Hunt, 2008; Lam and Hunt, 2009), with one exception: movement into and out of HME zones. Given a lateral movement opportunity, a COMPOUND selects randomly one of its eight neighboring elements and moves there, except special rules apply when the COMPOUND is attempting to move into or out of an element containing HME zone(s). In the final ISECC, the option to move into and out of a HME zone is governed by the value of an in silico distribution constant (value 689.5), which is calculated as described in the chapter 3 using the DRUG's *logP*, *pKa* and *insilicopH,* scaled by relative time-step duration. Because SAQ has a large *logP*, once it is in a HME zone, the probability of each simulation cycle favors staying there: a SAQ in an HME has 1/(1 + 689.5) chance to move from HME Z0 to a neighboring non-HME element.

For a COMPOUND to transit between spaces, it must be sufficiently close to a MEMBRANE interface to transit into that space. The probability of being so located is governed by *closeToInterface.* Because S2 and S4 elements map to volumes that are very small relative to S1, S3, and S5, and because a simulation cycle maps to minutes of wet-lab time, *closeToInterface* for COMPOUNDS in S2 and S4 is always 1. From S3, in addition to being sufficiently close to interface, a COMPOUND must be in Z0 to transit to S2; it must be in *Zn* to transit to S4. All COMPOUNDS transitioning from S2 to S3 are placed in Z0. Similarly, all COMPOUNDS transitioning from S4 to S3 are placed in *Zn*. Depending on PCPs as in Garmire et al. (2007),

brief transcellular transport around TJs from S1 to S5 and vice versa, within a single simulation cycle, does occur, but with very low frequency.

COMPOUND probabilistic movement between zones within a S3 element follows rules. Those rules are different for a COMPOUND that is currently in a HME zone and one that is in some other zone. Relocation for a COMPOUND not in a HME zone is controlled by *disperseProb* in Table 5.1; the values are tuned. The number of zones and movements between zones evolved as part of the iterative refinement process (Fig. 5.2). There can be up to eight different inter-zone movements, depending on project stage (discussed below) and ISECC specifications. Three general guidelines emerged during ISECC evolution. To enable sufficient COMPOUND to quickly reach S1 or S5, there must be a direct route from Z*n* to Z0. There is a bias towards Z0. Relocation between S3 zones in LY-treated-CULTURES (hereafter called LY-CULTURES) is slower than in CONTROL-CULTURES. Additional detailed observations about COMPOUND movement are in the following paragraph.

A COMPOUND in Z0 or Z*n* is in one of two states: it either is or is not currently in a HME zone. During a simulation cycle, a COMPOUND that is currently not in a HME zone will be given an opportunity to transition to another zone. In CONTROL CULTURES a COMPOUND may move to one of four nearest zones: Z(*i*+1), Z(*i*+2), Z(*i*–1), or Z(*i*–2), where $i+2 \leq n$ and $i-2 \geq 0$. Corresponding COMPOUND relocation options in ISECC that map to LY-CULTURES are to move to one of two nearest zones: Z(*i*+1), or Z(*i*–1), where $i+1 \leq n$ and $i-1 \geq 0$.

An element can contain one (Z0) or two (Z0 and Z6) HME zones. In the first case a COMPOUND currently in the HME zone is not given an option to change zones within that element; it stays in Z0. It may, however, like all COMPOUNDS in non-HME zones, relocate to a corresponding zone in an adjacent element, following the special rules above. In the second case a COMPOUND in either HME zone is given an option switch between HME zones: it moves to Z0 with a probability of 0.8 or to Z*n* with a probability of 0.2.

I also found it necessary to modulate SAQ mobility between non-HME zones within the same element. This trait could map to nonspecific binding, or exceeding a solubility limit, etc. I elected to name the parameter controlling this trait *solubility*. During each simulation cycle, a SAQ "queries" the total number of SAQS in S3. It will have an opportunity to move to a new zone only if a pseudo-random number (PRN) is less than [*solubility*/(SAQS in S3·(1 ± 0.2·*error*)], where *error* is randomly generated and has value between 0 and 1. *Solubility* (limit) is tuned; the value used for Fig. 5.3 is 250. This would be the first ISECC property revisited if I were to add a more stringent SM (SM-4) that required a higher degree of ISECC similarity with wet-lab results at the low and high doses.

5.2.5 Enzymes, transporters

The logic used by each CYP and PGP along with their function verification was reported in (Lam and Hunt, 2008). I have used, reused, verified, and validated several variations of CYP and PGP (Garmire and Hunt 2008; Lam and Hunt, 2009). Different model uses and different referent data can call for different capabilities. The ones used here are the simplest. Their logic diagrams are presented in chapter 3, Fig 3.2. A stationary CYP maps to an unspecified number of the cytochrome P450 enzymes that metabolize saquinavir and its primary metabolite. A stationary PGP, assigned randomly only to S2 elements, maps to an unspecified number of apical membrane components and processes, including P-gp transporters, that are responsible for facilitated efflux of saquinavir and its metabolite from intracellular space to the apical dosing compartment. Stationary means that the object is assigned randomly to and remains in one element for the duration of the simulation. A PGP can export COMPOUNDS from Z0 (either normal or HME) to S1. Based on results from Stage 1 ISECC experiments (discussed below), I specified that SAQ can be METABOLIZED by CYP to M7, which in turn can be METABOLIZED by CYP to M1. Further, M7 and M1 are both substrates of CYP and thus are competitive inhibitors of SAQ metabolism. I specified that SAQ, M7 and M1 are substrates of PGP. I also specified that M7 and M1 are less hydrophobic

than SAQ. BINDING to and RELEASE from CYP and PGP are governed by *assocProb* and *releaseProb*. CYP has an additional step: METABOLIZE, which is governed by *efficiencyProb* (*efficiencyProb* of PGP is 1). For this study, the specifications and internal logic of CYP and PGP validated in chapter 3 and in Lam and Hunt (2008) were expanded as follows. The CYP and PGP randomly selected their binding neighborhoods at the start of their assigned logic in each simulation cycle. CYPs are assigned only to two specific zones. For the finalized Stage 4 ISECC (Fig. 5.3), the Z5/Z4 assignment ratio is 3/2.

5.2.6 Probabilistic parameters and event scheduling

Most events, such as binding to CYP, are probabilistic and have probability parameter values in the 0 to 1 range. When an event option arises, a participating component draws a pseudo-random number (PRN) from the designated range. Its value is compared with that of a parameter to decide what action to take. For example, if the PRN is less than the value assigned to *assocProb*, then the event occurs. Otherwise, it does not. By using a uniform distribution with variable frequency parameters throughout the simulation, I added uncertainty that maps to our ignorance about many of the details that are below my level of resolution. Assumptions about frequency values were separated from, and external to the ISECC.

Upon initiation of a simulation, spaces and element objects are created, zones are specified, and all objects (COMPOUND, CYP, and PGP) are assigned to elements and added randomly to a list. During each simulation each object in the order listed is given an opportunity to execute its assigned logic. Execution order is shuffled randomly at the start of each simulation cycle. The number of COMPOUNDS in each space is recorded every five simulation cycles and written to external files for later analysis.

Table 5.1 Parameters and values for the ISECC in Fig. 5.3 that validated for Stage 4.

Parameter Group	Parameters	Remarks		CONTROL CULTURE	LY-CULTURE
ISECC System Parameters	*systemSize*	Size of SPACE S1–S5		50x50	50x50
	maxZ	Number of intracellular ZONES		7	7
	TJspace	Fraction of ELEMENTS that are TJ		0.15	0.15
	HMspace	Fraction of ELEMENTS that are HME		0.15	0.15
	numPgps ††	Number of PGP ††		500	15
	numCyps	Number of CYP		70	70
	insilicoPH	In silico pH		7.4	7.4
DRUG† parameters	*MW*	In silico molecular weight		670	670
	logP	In silico logP		3.025	3.025
	pKa	In silico pKa closest to in silico pH		7.13	7.13
	solubility	In silico intracellular solubility		250	250
	closeToInterface	Fraction sufficiently close to membrane interface	S1	0.8	0.8
			S5	0.04	0.04
	transitProb	Trans-membrane transit probability between SPACES For DRUGS in S3, DRUGS must also be in the respective ZONE to transit; For DRUGS in S1 and S5, DRUGS must be sufficiently close to membrane interface to transit	S1→S2	0.218	0.218
			S1→S5	0.002	0.002
			S2→S1	0.476	0.476
			S2→S3	0.476	0.476
			S3→S2	0.086	0.086
			S3→S4	0.086	0.086
			S4→S3	0.476	0.476
			S4→S5	0.476	0.476
			S5→S1	0.333	0.333
			S5→S4	0.019	0.019
	disperseProb ††	Dispersion probability across intracellular ZONES for DRUGS not in HME	Z*i*→Z0 ††	0.235	0.222
			Z*i*→Z(*i*–2) ††	0.059	0
			Z*i*→Z(*i*–1) ††	0.176	0.222
			Z*i*→Z(*i*+1) ††	0.176	0.111
			Z*i*→Z(*i*+2) ††	0.059	0
		for DRUGS in HME present in both Z0 and Z*n*	Z*i*→Z*n* ††	0.118	0.111
			Z0→Z*n*	0.2	0.2
			Z*n*→Z0	0.8	0.8

DRUG† parameters	*isaSubstrate*	Substrate of protein	CYP	true	true
			PGP	true	true
	assocProb	Binding probability to protein	CYP	0.35	0.35
			PGP	0.975	0.975
	releaseProb	Dissociation probability from protein	CYP	0.1	0.1
			PGP	1.0	1.0
ENZYME and TRANSPORTER parameters	*sitesN*	Neighborhood size	CYP	30	30
			PGP	60	60
	maxSites	Maximum capacity	CYP	1	1
			PGP	10	10
	efficiencyProb	efficiency parameter for CYP and PGP	CYP	0.7	0.7
			PGP	1.0	1.0
Experiment parameters	*randomizeOrder*	Randomize order of activation		true	
	a2bDirection	Dosing compartment: S1 = apical, S5 = basal		S1 or S5	
	numSolute	Dose: number of SAQS (*1000)		1, 2, 4, 8	
	simulationStep	Number of cycles to run		120	

† Parameters for M7 and M1 are the same as those of SAQ except for the following:
Lower *logP*, lower *MW*, lower *assocProb* to CYP and PGP, and lower *releaseProb* from CYP and PGP

†† Parameter values different between control and inhibitor-treated experiments

5.2.7 Experimental condition

I simulated 16 different experimental conditions: four treatments at four dosing conditions. The four treatments were: APICAL (S1) SAQ dosing of CONTROL and LY-CULTURES, and BASAL (S5) SAQ dosing of CONTROL and LY-CULTURES. Four SAQ amounts were used: 1,000, 2,000, 4,000 and 8,000. They mapped to 7.5, 15, 30 and 60 nanomoles and 5, 10, 20, 40 μM, respectively. The number of CYP used ranged from 20 to 70, depending on Stage (described below). The number of PGP used for experiments CONTROL CULTURES was one of the following: 250, 400, 500, or 600. The number of PGP used for experiments LY-CULTURES was one of the following: 10, 15, 20, or 25 (early ISECCs that used 0 were falsified). To complete one experiment, I averaged results from ten (selected for convenience) repeated simulations. All simulation ran for 120 simulation cycles.

5.2.8 Groundings

The units, dimensions, and/or objects to which a variable or model constituent refers establish groundings. Inductive ordinary differential equation models are typically grounded to metric spaces. So doing provides simple, interpretive mappings between output and parameter values and DATA. However, metric grounding creates issues that must be addressed each time one needs to expand the model to include additional phenomena. Adding a term to an equation, for example, requires defining its variables *and premises* to be quantitatively commensurate with everything else in the model. Such expansions can be challenging and even infeasible when knowledge is limited and uncertainty is high, which is the situation that I faced. Iterative model refinement as described in Fig. 5.2 becomes slow and complex when the model is grounded to metric space. To discover plausible concrete mechanistic explanations for the targeted data, I needed the ability to simultaneously explore different regions of plausible mechanism space at different levels of detail, and relate results to wet-lab observations. To facilitate that process it must be easy to change mechanistic or component details at any level without having to invest significant time in analogue reengineering. I have discovered that the best way to achieve those objectives is to remove metric grounding from the ISECC and confine it to quantitative feature-to-feature and phenomena-to-phenomena mappings (Lam et al., 2009), and that is what I did.

The mechanisms responsible for generation of the TRANSPORT DATA do not interact according to any external measurement methods. They are independent of any measures used by an outside observer. From that fact, I inferred that the ISECCs needed to achieve my objectives must employ similar internal organization, which in modeling terms, means each component is grounded to other components rather than to a metric imposed by an outside observer, and that is the course I have followed.

During iterative ISECC refinement several different groundings were used. For the final ISECC and the experimental results reported herein, the quantitative groundings were as follows:

1,000 SAQ → 7.5 nanomoles, 75 simulation cycles → three hours (during Stage 1 it was 60 simulation cycles → three hours), and one ISECC experiment → one wet-lab experiment. The SAQ grounding took into consideration the lower limit of analytical detection in referent experiments (Mouly et al., 2004), which was < 0.012 nanomoles, equivalent to 1.6 SAQ objects.

5.2.9 Similarity measures

The ultimate goal was to have ISECCs such that simulation results intended to match a specific TRANSPORT experiment would be statistically indistinguishable from a repeat of that targeted wet-lab experiment. Taking into consideration the variability between wet-lab experiments within and between conditions, I specified that my goal would be met by achieving the Stage 4 objectives stated below. The most stringent Similarity Measure (SM), SM-3, was that no more than one of a selected subset of conditions is outside the range of 66.7%–150% of referent values. Meeting SM-3 would imply statistical indistinguishability to wet-lab observation. To enable adhering to the parsimony guideline and progress through many ISECC refinement cycles, two less stringent SM targets were used. For SM-1 (least stringent), simulation results of no more than three of the targeted conditions are outside 50%–200% of referent values. For SM-2, simulation results of no more than three of the targeted conditions are outside target range of 66.7%–150% of referent value, or no more than one is outside range of 50%–200% of referent values.

5.2.10 Achieving targeted attributes

The goal of discovering one ISECC with one set of plausible micromechanisms was approached in stages following the protocol in Fig. 5.2. Early in the process, results from only two of 16 experimental conditions were targeted using SM-1. At Stage 4 the targeted attribute list was expanded to all 16 experimental conditions using a combination of SM-3 and SM-1.

Robustness to minor parameter change (e.g., ~10% for one parameter and ~5% for 2–4 parameters) was an additional targeted attribute. I illustrate with an example. Having an ISECC that achieved Stage 2, I sought component changes and additions that would enable achieving Stage 3. Several changes that helped, but failed, I kept. Consider a predecessor to the ISECC that achieved Stage 3. It is possible that there is a small region (maybe more than one) of that ISECC's component and parameter space that could have achieved Stage 3, but I failed to locate it. However, had I found it, I would have observed that small changes in several parameters caused the ISECC to fall short of the SM target. I encountered several such instances in route to the ISECC in Table 5.1. I decided that in order to validate, an ISECC must tolerate small changes (in the 5-10% range) in parameters while still achieving targeted SMs. This specification is currently arbitrary and qualitative. However, it can be made more precise and quantitative when there is a need to do so.

Stage 0: meets design specifications

Stage 1: target one dose, under control (no inhibitor) condition only

Qualitative: paradoxical observation: apical dosing produces higher intracellular SAQ levels, but less M7 metabolite (Mouly et al., 2004)

Stage 2: target the two middle doses, control conditions only

Qualitative: large difference in receiving compartment after apical versus basal dosing

Quantitative: simulation results meet SM-1

Stage 3: target the two middle doses for control and inhibitor treatment

Qualitative: for inhibitor treatment relative to control: more SAQ in cells, more total M7

Quantitative: results from CONTROL CULTURES meet SM-2; those from LY-CULTURES meet SM-1

Stage 4: target all four doses for control and inhibitor treatment

Qualitative: evidence of metabolism saturation at larger doses (Parker and Houston 2008); similar M7 per intracellular SAQ ratio

Quantitative: results for two middle doses meet SM-3 whereas those from lowest and highest dose meet SM-1

While progressing from early Stage 1 to Stage 4 validation, I followed somewhat standardized parameter sweeping methods to discover a specific ISECC parameterization that would enable achieving the targeted attributes and SM for that ISECC. I explored four types ISECC characteristic change. 1) Change spatial properties: I changed S3 element properties by subdividing the space within into different numbers and types (HME) of zones. 2) Change mechanistic components: for a given set of elements and properties, I changed the ratios and numbers of CYP and PGP. 3) Change logic used by components. 4) For specific valuations of the three preceding sets of characteristics, I tuned the values of remaining parameters. I began by randomly probing regions of property and parameter space and running experiments. I selected the two or three locations that moved me closest to validation, and randomly sampled nearby parameterizations. That process was repeated until a decision was made to fix the first three ISECC characteristic types. Focus then shifted to tuning just a few of the remaining parameters to improve results even further.

5.2.11 Software and simulation time

I reused validated components from previous projects (Liu and Hunt 2006; Garmire et al., 2007; Lam and Hunt, 2008; Lam and Hunt, 2009). ISECCs were assembled within the Swarm platform using its libraries (http://swarm.org). I coded in Java Swarm. Most experiments used a single processor and ran under Microsoft Windows XP (Redmond, WA) with Java Software Development Kit and Java 2 Runtime Environment installed (Sun Microsystems, Santa Clara, CA). Source code was compiled with Java 2 Software Development Kit SE version 1.4.2_13 and

executed with Java 2 Runtime Environment version 1.4.2_13 (www.java.com). Output DATA files were processed, graphed, and analyzed using Microsoft Excel. A complete set of ISECC for all 16 conditions averaged about one hour when using one processor; analysis of results often took longer.

5.3 Results

5.3.1 Summary of wet-lab results

Mouly et al. (2004) reported that saquinavir's apparent permeability is always higher for basal-to-apical transport, compared to that of apical-to-basal. Because saquinavir is a P-gp substrate, that result was expected. As predicted, inhibition of P-gp by inhibitor increased apical-to-basal while decreasing measured basal-to-apical permeability. In addition, inhibitor treatment significantly increased the intracellular level of saquinavir, and formation of saquinavir metabolite at the end of the three-hour experiments. However, despite the lower intracellular saquinavir level after basal compared to apical dosing, basal dosing consistently, and paradoxically, produced more metabolite M7 in both control and inhibitor-treated cultures, and for the four doses studied. Such a paradox had not been reported previously, and its cause could not be isolated based on experimental observations. The investigators offered two potential explanations: 1) a saquinavir concentration gradient exists intracellularly, and the parent drug presented at a higher concentration at the enzymes' microenvironment after basal dosing; and 2) secondary metabolism of M7 was greater after apical dosing. Variances in results for repeat experiments were typical of *in vitro* transport studies. Coefficients of variation of the various measures ranged from 16 to 80% and averaged 35% for experiments on five matched cultures.

5.3.2 Summary of ISECC experiments

The ISECC, which validated against the Stage 1 attributes and SM, was described in (Lam and Hunt, 2008). From there, I cycled through the iterative refinement protocol several

hundred times before discovering an ISECC that validated by successfully achieving the attributes and SM specified by the Stage 4 criteria. All had the same five spaces and all included

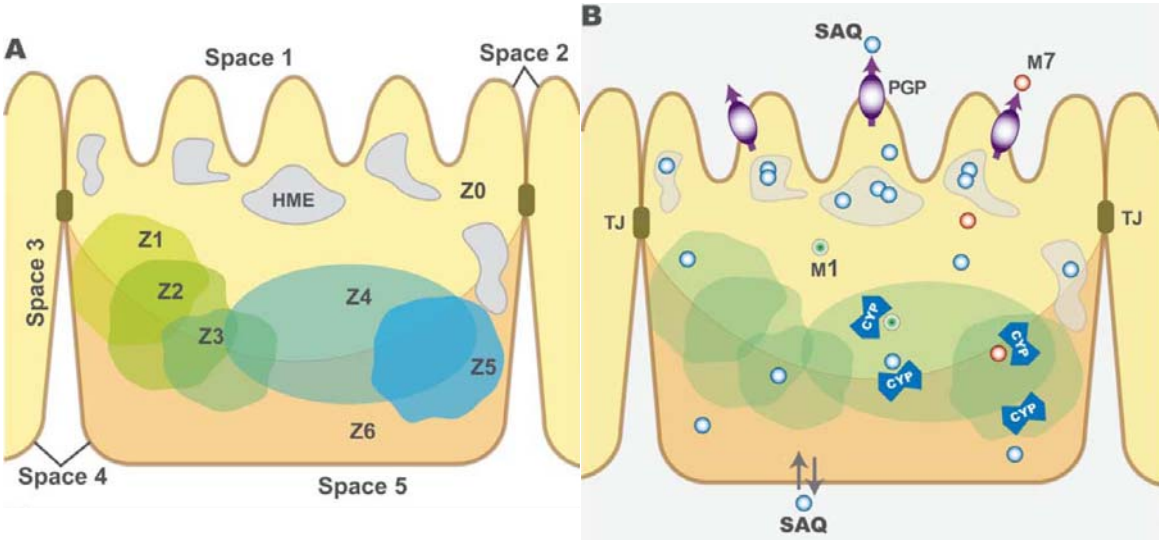


Figure 5.3. Illustration of spatial heterogeneity for the ISECC that achieved Stage 4 validation. Illustrated is the ISECC resulting from Table 5.1 parameterizations that gave the DATA in Figs. 5.4–5.6. **A:** Five identically sized spaces are used: S1–S5, as described in the text. The element at the interface of two CELLS (which maps to tight junction spaces between cells, and other cell-cell attachments) is called a TJ. Each of the 2,500 elements comprising S3 is subdivided into seven zones, Z0–Z6. Because an element is the lowest level of spatial resolution, the actual shapes and locations of Z0–Z6 are not specified, although aspects of their relative arrangement are specified. Z0 and Z6 map to large portions of intracellular space adjacent to each of the membranes. Z1–Z5 are located more centrally, they represent a series of heterogeneous subcellular microenvironments of indefinite shape and volume; they may map to subcellular structures such as nucleus, endoplasmic reticulum or mitochondria. The Z0 in 15% of S3 elements are specified to map to subcellular, lipid-like, hydrophobic microenvironments called HME (a SAQ has a high affinity for these spaces). Some (20%) S3 elements that contain a HME-Z0 also contain a HME-Z6. **B:** Illustrated are the three types of mobile objects (SAQ, M7, and M1) and the two types of immobile objects (PGP and CYP). At the initiation of a simulation, objects mapping to Cyp3A4 enzymes (CYP) and P-gp transporters (PGP) are created and placed randomly in S3 (zones Z4 and Z5 only) and S2, respectively. Objects representing saquinavir (SAQ) are placed either in the APICAL (S1) or BASAL compartment (S5). Those SAQ map to the apical dose and basal dose, respectively. In each simulation cycle, each mobile object gets an opportunity to move between elements in adjacent spaces, and between elements within the same space. It can also move between zones within a S3 element. PGP and CYP are confined to assigned elements. CYP probabilistically METABOLIZE SAQ in Z4 and Z5 to M7, and further METABOLIZE M7 to M1. PGP is responsible for the active EFFLUX of SAQ and its METABOLITES from Z0 to S1. Other properties of mobile and immobile objects are as detailed in the text.

CYP in S3 and PGP in S2. The results presented in Figs. 5.4–5.8 summarize TRANSPORT results for the validated ISECC described in Fig. 5.3. Figures 5.4–5.6 correspond to Fig. 1A–C in Mouly et al. (2004). An overview of seventeen ISECCs that were thoroughly explored and eventually falsified before achieving the targeted Stage 4 validation criteria is provided as Fig. 5.9.

5.3.3 Plausible explanations for paradoxical results

The two potential mechanisms offered by Mouly et al. (2004) as plausible explanations proved necessary but not sufficient to cause comparable results within the ISECC context. They suggested that an intracellular gradient might exist from dosing to receiving compartment. I implemented that mechanism in a Stage 1 ISECC. I achieved SM-1 for two doses for both control and inhibitor-treated cultures but were unable to achieve SM-2. The accumulation of relatively large amounts of SAQ in the receiving compartment could not be achieved while also keeping the amount of (all) metabolite formed small. The results demonstrated that within the ISECC a gradient effect alone was not sufficient. Thereafter, I discarded the specification that the INTRACELLULAR space behaves as a well-stirred space. Later experiments demonstrated that allowing S3 to be heterogeneous and contain pockets (HME) having different SAQ affinities was necessary to achieve SM-2 and later SM-3.

Mouly et al. (2004) also suggested secondary metabolism (to M1) as an explanatory mechanism, and that formation of M1 from M7 was greater after apical dosing. That mechanism would require M7 disposition to be different following apical and basal dosing. However, within the validated ISECC, secondary METABOLISM caused competitive inhibition of SAQ METABOLISM, and that accounted for why a large fraction of the SAQ dose passed through the CELL while only a tiny fraction was present as M7. Furthermore, METABOLISM saturation began influencing results at higher doses.

5.3.4 P-gp activity, inhibitor-treated cultures, and temporal mapping

All parameterization decisions were interconnected. I narrowed options by specification early of the number of SAQ to be used per simulation: 1,000–8,000. Exploratory simulations showed that an appropriate number of PGP for that range of SAQ would be 250–600. For simplicity, I simulated inhibition by deactivating all but a few PGP. Results from exploratory simulations indicated that the number of fully active PGP in LY-CULTURES needed to be > 0 and

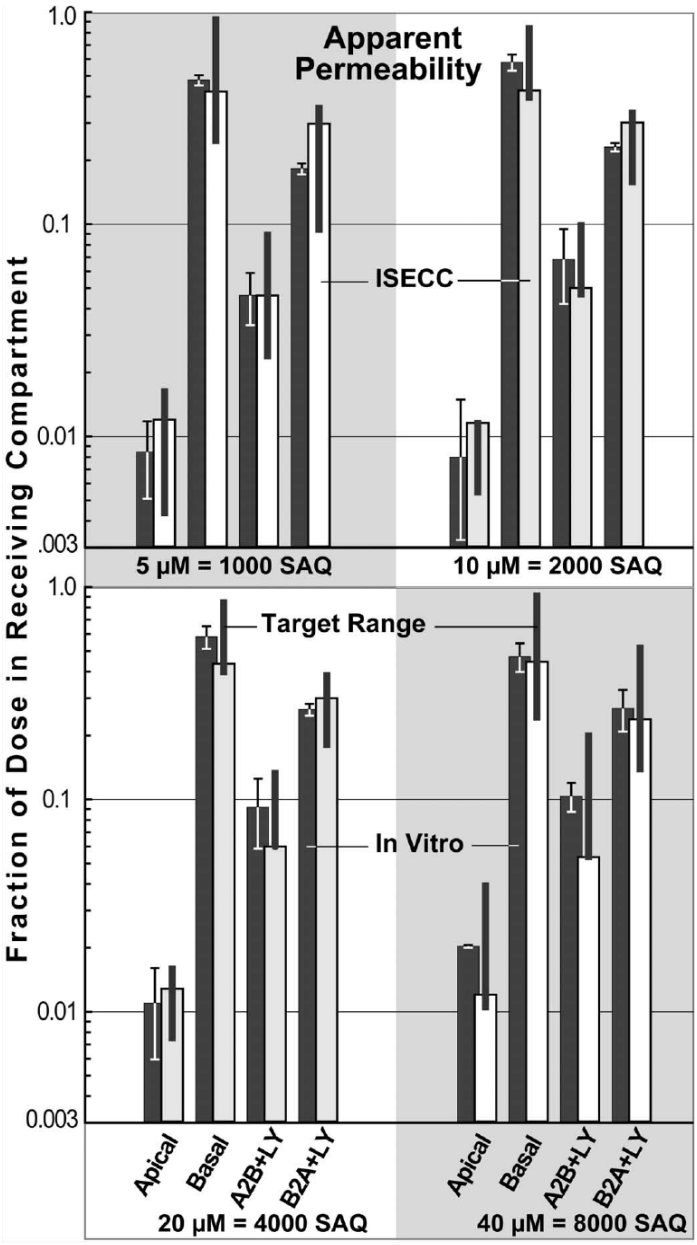


Figure 5.4. Apparent permeabilities. Graphed are apparent permeabilities for saquinavir and SAQ. The former values correspond to Fig. 1A in Mouly et al. (2004); the vertical bars show ± 1 SD. The heavy vertical bars with ISECC DATA show Stage 4 target ranges. For the ISECC experiments, apparent PERMEABILITY = (amount of SAQ in receiving compartment after 75 simulation cycles [maps to three hours]) ÷ (SAQ dose). The results are means for ten ISECC simulations parameterized using Table 5.1 values.

that an appropriate range might be 10–25 PGP. Thereafter, when seeking a parameterization that would validate, experiments of LY-CULTURES used 10, 15, 20, or 25 PGP to see which was best. The validated ISECC in Table 5.1 used 500 PGP for CONTROL and 15 PGP for LY-CULTURES for the four dose conditions. A quantitative mapping of the results in Figs. 5.4–5.6 to wet-lab means that inhibitor treatment reduced P-gp activity by about 97%.

Each event occurring within a simulation cycle during an ISECC simulation contributes in a small way to overall outcomes. That means that it is possible for a small change in one or a

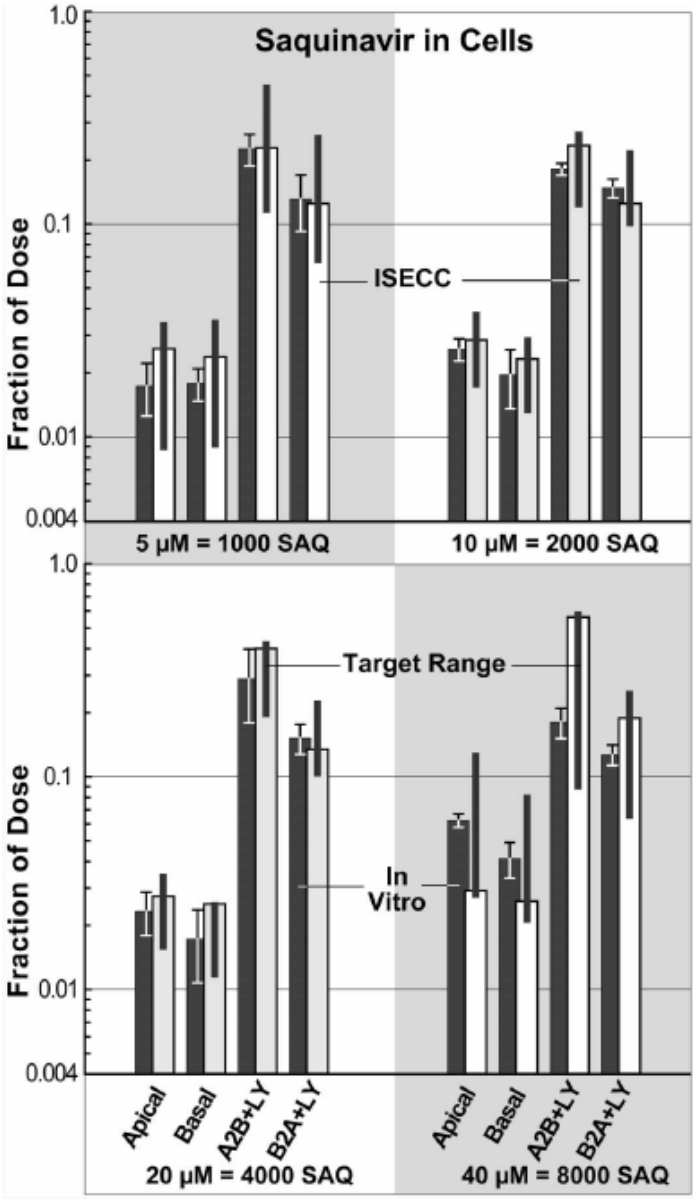


Figure 5.5. *In vitro* and simulated intracellular sequinavir accumulation. The results are for the same experiments as in Fig. 5.4. Graphed are the 1) amounts of saquinavir that were intracellular after three hours; they correspond to Fig. 1B in Mouly et al. (2004), and 2) amounts of SAQ that were INTRACELLULAR (S2 + S3 + S4) after 75 simulation cycles along with target ranges (heavy bars).

few micro-mechanisms to be offset by other changes, such as changing how a simulation cycle maps to wet-lab time. With that realization in mind, during each iterative refinement cycle alternative ISECC-to-wet-lab temporal mappings were explored to see if an adjustment in that mapping could improve SM outcomes. Once the amount ranges for dose, PGP, and CYP were fixed, it quickly became apparent that ISECC TRANSPORT results at 60–90 simulation cycles mapped best to the wet-lab three-hour transport data. The choice does not influence any of the ISECC premises. For the ISECC represented by Table 5.1, the ISECC TRANSPORT results at 75 simulation cycles provided the best match to referent data. For that mapping, one simulation cycle maps to ~2.4 minutes.

5.3.5 Components and features of the ISECC that validated

Each ISECC for which simulation results matched a prespecified set of targeted attributes and a prespecified SM achieved a degree of validation. Improving mapping A in Fig. 5.1 while also increasing the attributes targeted, increases confidence that there is some validity to mapping B in Fig. 5.1. The reported measure of transport for each experimental condition is a phenotypic attribute of the in vitro system. If the five-space representation is acceptable for a monolayer of essentially identical cells, then those 16 attributes, taken together, place serious constraints on the space of mechanisms that will validate, especially when the variety and properties of stationary component types allowed is limited to two: CYP and PGP. I argue that an ISECC that validates at Stage 4 should be taken more seriously (as having plausible biological counterparts) than an ISECC that can only validate at Stage 2. Further, any ISECC that validates at or above Stage 1 should be taken more seriously than any descriptive conceptual mechanism. What follows is a description of the features of the ISECC specified by Table 5.1.

The INTRACELLULAR environment (S3) is subdivided into seven ZONES, with Z0 adjacent to S2, and Z6 adjacent to S4. Z0 in 15% of elements are specified as being HME. In addition, in one-fifth of the elements containing a HME Z0, the Z6 in that element is also specified as being HME. COMPOUND movement into and out of HME is governed by an in silico distribution. SAQ is highly hydrophobic. Once it enters a HME it preferentially stays in HME. The net result is that a large portion of INTRACELLULAR SAQ gets sequestered in HME and is unable to reach METABOLIC ENZYMES. Further, that portion is larger after APICAL dosing. I suggest S3 zones in ISECC map to heterogeneous microenvironment within the cell monolayer.

During INTRACELLULAR COMPOUND movement, there is a bias towards Z0. A

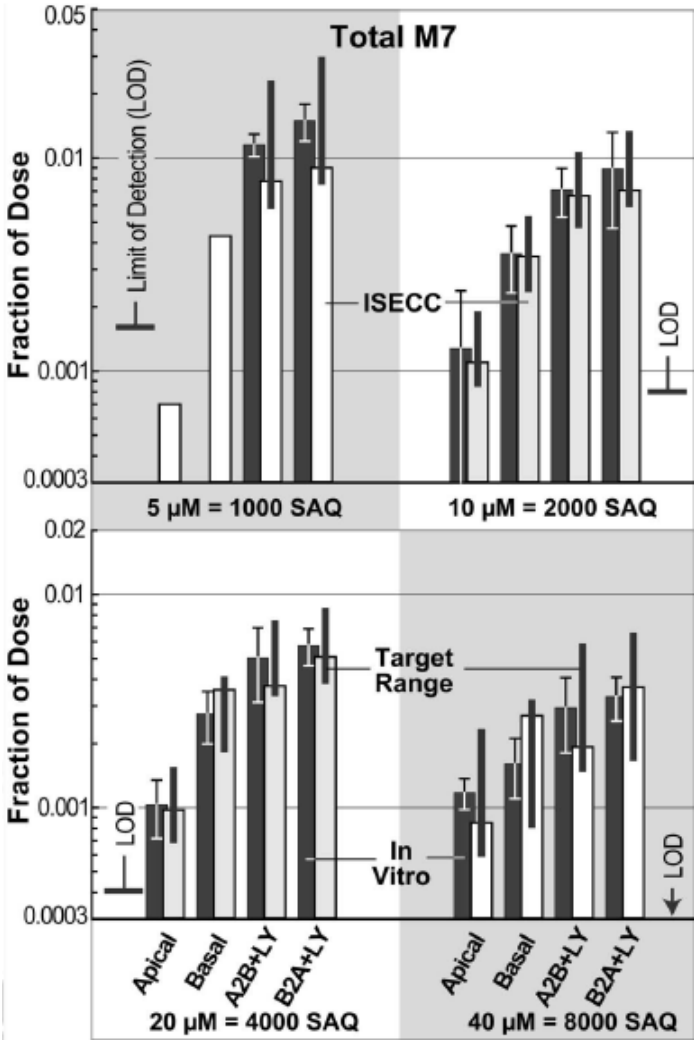


Figure 5.6. *In vitro* and simulated total metabolite. The results are for the same experiments as in Figs. 5.4 and 5.5. Graphed are 1) the dose fraction present as M7 after three hours; the values correspond to Fig. 1C in Mouly et al. (2004), and 2) the fraction of DOSE present as M7 within all ISECC spaces after 75 simulation cycles along with target ranges (heavy bars). For each dose, the lower limit of M7 detection is indicated.

COMPOUND within an element may move to the APICAL (Z0) and BASAL (Z6) zones, or to one of the four nearest adjacent zones (Z*i* + 1, Z*i* + 2, Z*i* – 1, or Z*i* – 2). In LY-CULTURES, COMPOUND movement is slowed (less displacement over time). A COMPOUND within LY-CULTURE S3 element may move to the APICAL (Z0) and BASAL (Z6) zones, or to one of the two nearest adjacent zones (Z*i* + 1, or Z*i* – 1). These movements, including the effective shortcut from Z0 to Z6, may map to any of a variety of yet unidentified micromechanisms. Plausible examples include saquinavir binding (somewhat preferentially) to a saturable component of a transcellular transport system, for example, fatty acid intracellular transport system described by Weisiger (1996; 2002; 2007), or becoming associated (somewhat preferentially) with a nonequilibrium, apically directed intracellular flux of the type described by Kurakin (2009).

Assume that the above saturable, transcellular saquinavir transport hypothesis is valid. In LY-CULTURES, INTRACELLULAR SAQ movement is slowed. That may map to the inhibitor also competitively inhibiting saquinavir movement. Assume that too is valid. I should then see decreasing apparent permeability (Papp) when the saquinavir dose increases in control cultures, and I should see increasing Papp when saquinavir dose is increased in inhibitor-treated cultures. Both predictions are consistent with wet-lab observations. However, because I had no evidence-based insight into actual mechanisms, the ISECC used abstract SAQ movement rules as placeholders for all plausible, fine-grained mechanisms. By so doing, I preserved ignorance. No component based carrier micromechanism was actually implemented. Consequently, there were no ISECC apparent permeability observations analogous to the above-cited wet-lab data. Exploring that scenario is an option for a future ISECC study.

To achieve Stage 4 validation I found it effective to impose a limit on the amount of SAQ that could be within S3 during a simulation cycle. I did that by specifying SOLUBILITY in S3 to be 250. I suggest that there may be an effective counterpart within the Transwell cells.

The TRANSPORT properties of each explored ISECC type moved closer to validation when CYPS were sequestered within a portion of S3. For the final ISECC, CYPS were assigned 40% to Z4 and 60% to Z5; a 50/50 assignment resulted in failure to achieve the Stage 4 SMs. Those assignments meant that a SAQ was more likely to encounter a CYP when coming from S5 (BASAL dosing) rather than S1 (APICAL dosing). These CYP assignments within the ISECC map to a non-uniform, intracellular distribution of saquinavir metabolizing enzymes. Because

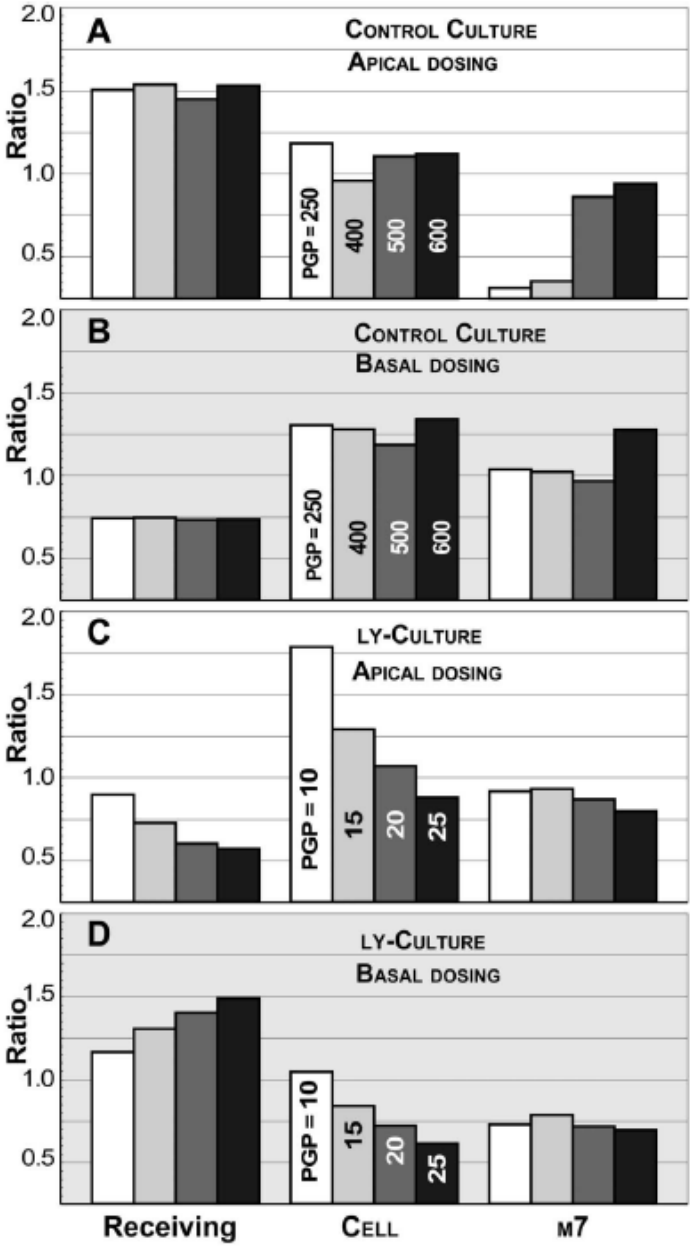


Figure 5.7. Robustness of ISECC to changes in PGP numbers. Separate sets of experiments using the ISECC in Table 5.1 were completed using the indicated number of PGP. Everything else was unchanged. Bar heights are ratios of ISECC-to-wet-lab results of the type in Figs. 5.4–5.6 at matched intervals (75 simulation cycles and three hours) after APICAL (**A**, **C**) and BASAL (**B**, **D**) dosing. **A** and **B**: control conditions: the results are for PGP = 250–600; results for 500 and 600 meet SM-3; all results meet SM-2. **C** and **D**: Inhibitor treatments: results are for 10–25 active PGP remaining after inhibitor treatment; the DATA for 10 and 15 PGP meet SM-3; all results meet SM-2. Note that coefficients of variation of the wet-lab measures ranged from 16 to 80% and averaged 35%. Comparable adjustments of other parameters and components caused the same gradual change in ISECC TRANSPORT DATA.

movement within S3 was slower in LY-CULTURES, proportionally less SAQ reached Z4 and Z5, and that caused less M7 formation. I suggest that these events may have counterparts within Transwell cells treated with inhibitor.

Because SAQ was assigned a large MW (670) there was very little PARACELLULAR TRANSPORT. If I lower MW, PARACELLULAR TRANSPORT increases. That is because 15% of S1 and S5 elements are marked TJ. The TJ map to all cell-cell attachments and to any spaces between cells. The probability to transit to S1 from S5 when in a TJ element (0.25) was specified to be much greater than probability to transit to S5 from S1 via a TJ element (0.001).

5.3.6 ISECC robustness to parameter change

Upon achieving Stage 4 (or an earlier stage) goals, ISECC robustness to parameter change for a variety of parameters was measured. Observing that a small change (5–15%) in one parameter caused the ISECC to invalidate was considered abiotic; a search for a different region of parameter space was initiated. To illustrate the process for the ISECC in Table 5.1 for both CONTROL and LY-CULTURES, the effects of parameter change on the three TRANSPORT measures are provided in Figs. 5.7 and 5.8 for the influential parameter PGP levels and for choice of temporal mapping. The values graphed in both figures are the ratio of ISECC-to-wet-lab transport measures. The number of PGP in Table 5.1 was 500 for control and 15 for LY-CULTURES. Figure 5.7 presents TRANSPORT results for 250–600 for control and 10–25 for LY-CULTURES. The temporal mapping for Figs. 5.4–5.6 is 75 simulation cycles $\rightarrow$ three hours. Fig. 5.8 presents TRANSPORT results for eleven additional temporal mappings: 30–120 simulation cycles $\rightarrow$ three hours. Results of these and other robustness explorations (not shown) showed that many other parameter vectors close to the one in Table 5.1 can also produce ISECC that validate.

5.4 Discussion

An implication of the ISECC validation evidence is that the mechanisms depicted in Fig. 5.3 have counterparts during saquinavir transport through Caco-2 cells. The simulation results provide concrete, scientific evidence that interactions with heterogeneous intracellular microenvironments, coupled with intricate, intracellular saquinavir movements, some possibly carrier-mediated, provide a plausible explanation of the cited paradoxical observations. All compounds will encounter such heterogeneity, but the influence may not be evident from traditional assessments of transport data. It seems unlikely that such microenvironment heterogeneity will be confined to Caco-2 cells. Microenvironment differences within and between cell types as a function of health and disease may influence therapeutic availability to target sites. Future therapeutics may exploit heterogeneous intracellular disposition to enhance efficacy and minimize toxicity. The heterogeneities may also contribute to intra- and interindividual variability in disposition and response. Wet-lab technologies and experiments designed to detect the influence of microenvironment heterogeneities are needed to help determine their importance.

The degree of similarity between the ISECC and wet-lab data in Figs. 5.4–5.6 gives strength to the hypothesis that the mechanisms depicted in Fig. 5.3 have counterparts during saquinavir transport through Caco-2 cells. However, as abstract models, ISECC mechanisms are flawed. Nevertheless, I seek to minimize discrepancies by exploring multiple mechanistic explanations using the iterative refinements protocol in Fig. 5.2, and only exploring those that are not abiotic.

Even though the final ISECC mechanisms are far less complex than what I already know about epithelial cells, a logical initial reaction upon reading the Fig. 5.3 explanation would be to muse about simpler mechanistic explanations: are there no simpler explanations? There may be, but I have not yet found one. The iterative refinement protocol in Fig. 5.2, coupled with the

parsimony guideline, has proven to be an effective tool in resisting making ISECCs unnecessarily complicated. Following that protocol created a mechanism exploration path (shown in Fig. 5.9) for which the next added mechanistic detail (or complication) depended on predecessor mechanisms. Ideally, I would prefer to explore many branches of many parallel paths. I cannot rule out the existence of alternative paths that lead to somewhat simpler explanations. Some have not been fully explored, in part because my objective was to find one plausible set of micromechanisms to achieve a prespecified SM for all 16 experimental conditions. Below, I present four of several mechanistic exploration paths that may merit future attention, wet-lab as well as in silico.

Path 1: for the path followed, Path 0, I specified that an ISECC represents a monolayer comprised of essentially identical cells. Reality may be more complicated. There may be differences in saquinavir transport within and through different subsets of cells. For example, cells may be classifiable into at least two types for which the intracellular micromechanisms are simpler than those in Fig. 5.3, but together they achieve the same Stage 4 SMs. I explored this option but failed to come close to the targeted SMs. Wet-lab experiments, possibly using new imaging technologies, could help confirm or rule out such differences and thus indicate the relative importance of exploring this path.

Path 2: for Path 0, I specified that INTRACELLULAR micromechanisms be invariant. That was not the case in Lam and Hunt (2009). Interactions of cells with saquinavir (and the inhibitor) may initiate time-dependent mechanism changes. I did not explore this path to any appreciable extent, but based on my earlier work, I can be confident that this path could lead to mechanisms that achieve the Stage 4 SMs. However, it would be debatable whether or not successful time-variant mechanisms would be *simpler*; but they would be different. Again, coupling in silico with wet-lab experiments could indicate the relative importance of exploring this path.

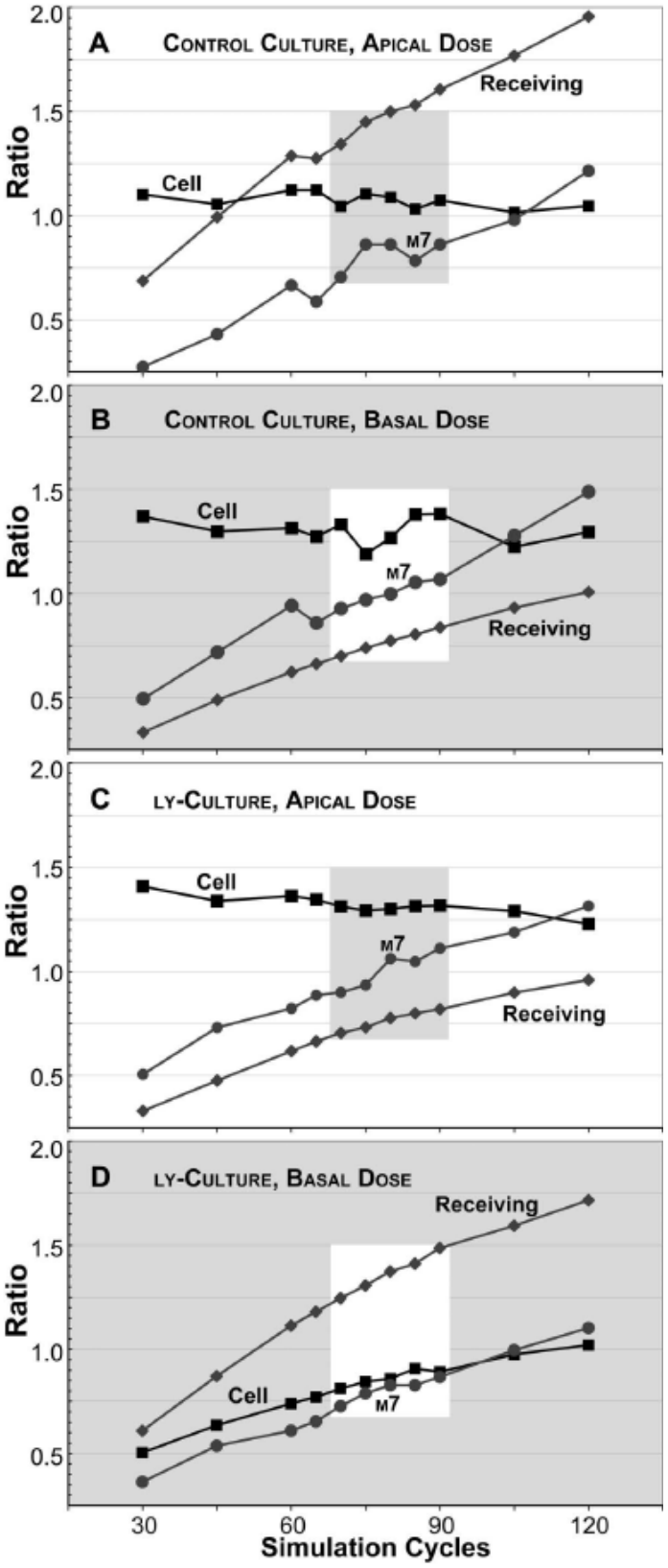


Figure 5.8. Robustness of ISECC to changes in temporal mapping. Temporal mapping is established by selecting the number of simulation cycles that map to three hours of wet-lab time. Y-values are ratios (as in Fig. 5.7) of ISECC-to-wet-lab results of the type in Figs. 5.4–5.6. X-values are time in simulation cycles. **A** and **C**: results from apical dosing. **B** and **D**: results from basal dosing. **A** and **B**: control conditions; **C** and **D**: Inhibitor treatments. Diamonds ◆: amount of SAQ in receiving compartment; squares ■: amount of SAQ in cell; circles ●: total amount of M7. All results for time = 60-120 meet SM-2. The box shows the target range for SM-3. Results within the box for time = 70-90 also meet SM-3. Comparable adjustments of other parameters and components caused the same gradual change in ISECC transport data for temporal mappings.

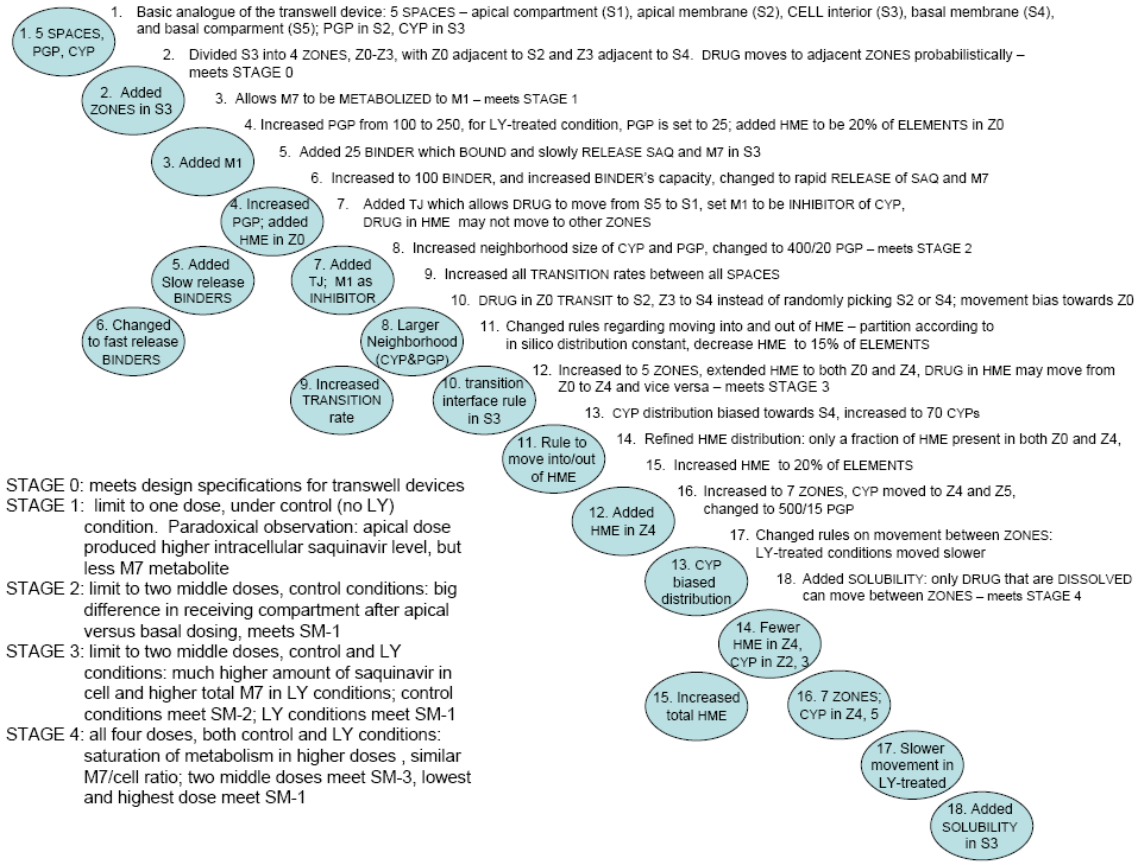


Figure 5.9. Overview of eighteen ISECCs that were thoroughly explored and eventually falsified before achieving the targeted Stage 4 validation criteria. A series of 18 ISECCs were assembled and tested. Analogues meeting the specification of the Stages are noted. At each step, the parameter space was extensively explored until an observable improvement in similarity was achieved. Key parameter adjustments, rule changes and/or component added are noted for each steps. Branching occurred when changes initially improved similarity but later failed to validate with additional modifications.

Path 3: Kurakin (2009) argues persuasively for abandoning the conventional biological paradigm rooted in classical mechanics and equilibrium thermodynamics in favor of one in which all cell subsystems are viewed as dynamic, adaptive, nonequilibrium systems that are part of their environment. The synthetic M&S method enables building such analogues, but it is not clear that they can be mechanistically simpler than the one in Fig. 5.3.

Path 4 (not yet explored): given the abstract simplicity of ISECC components, it may be unrealistic to insist on a linear mapping between amount of SAQ used and wet-lab dose.

Because all ISECCs are simplifications of reality, they are mechanistically depleted and thus flawed in specific ways. Nevertheless, some can be very useful. There are many referent system attributes such that when one is added to the targeted attribute list, the Fig. 5.3 ISECC will be falsified. I had no evidence indicating that any of the above four mechanistic exploration paths should be preferred over Path 0. Nevertheless, along Path 0 there may be mechanistic branches that did not occur to me yet may contain simpler systems.

One solution to the above mechanism generation and selection issues is, when feasible, to use simulations to identify phenomena that suggest new wet-lab experiments that in turn could rule in or out exploration of a mechanistic branch. Doing so would require timely coordination of wet-lab and simulation experiments at steps 8a and 8b in Fig. 5.2.

Another strategy for addressing the mechanism generation and selection issues is to seek ensembles of different explanatory analogues spanning an acceptable variety of mechanistic options, and allow them to compete in offering plausible explanations of an increasingly rich set of targeted phenotypic attributes. The cost of implementing such a solution can be kept reasonable by developing automated modeling methods capable of discovering mechanistically different analogues. The information provided in Fig. 5.9 documents that I built, tested, and falsified many ISECCs before discovering the one that validated. The process requires extensive human effort and time. The technology can be developed to use software agents within the existing computational framework to manage the process of changing a falsified ISECC by adjusting the parameter vector and/or the mix of components with the objective of achieving a prespecified SM.

There are several features of the ISECC in Fig. 5.3 that merit discussion. Saquinavir is known to preferentially bind to mitochondria and the nucleus (Khoo et al., 2002). The HME may map in part to these subcellular structures. I suggest that the movement logic used by SAQ to change zones, especially movement between Z0 and Z6, may map to saquinavir binding to

mobile components of one or more intracellular transport system, such as the fatty acid transport system (Weisiger 2007). If future evidence supports that mapping, then the ISECC TRANSPORT results can be taken as predictions that the carrier mechanism(s) will be saturated at higher saquinavir doses. The mechanistic changes in LY-CULTURES may map to the inhibitor reducing saquinavir interaction with the carrier system(s), possibly through competitive binding. The preceding conjecture is consistent with wet-lab (but not ISECC) observations that apparent saquinavir permeability is decreased with increasing saquinavir dose in control cultures but is increased with increasing saquinavir dose in inhibitor-treated cultures. Comparable ISECC results were not observed because, as explained in Results, SAQ movement rules functioned as placeholders for all plausible, fine-grained, componentized mechanisms. Well-designed wet-lab experiments may help clarify the biological implications of this conjecture.

Introducing the idea of a SAQ SOLUBILITY limit within S3 elements enabled improving high and low DOSE ISECC results sufficiently to achieve Stage 4 validation. However, should I set a new Stage 5 goal that includes a more stringent SM for high- and low-DOSES, the current ISECC would be falsified (invalidated). To achieve the new Stage 5 I would need to explore alternative micromechanistic features including improvements to the current SOLUBILITY algorithm. Given those considerations, it is premature to seek Caco-2 counterparts to which the current SAQ SOLUBILITY micromechanism may map.

Early during the iterative ISECC refinement process, it became clear that the pace of refinement and mechanistic exploration was accelerated by specifying mappings, including temporal mapping (simulation cycles to wet-lab minutes), the final adjustments in the process. So doing seemed counterintuitive initially, because most inductive transport models nail down temporal mappings at the start by equating some parameter of the induced model with some independent wet-lab measure with which it is identified. Doing the same in ISECCs limits one to a small set of refinement paths, while making mechanistic refinement increasingly difficult. I

noted that different PGP numbers and temporal mappings were equally acceptable. Because methods of measurement are external to both wet-lab models and ISECCs, I now argue that when feasible, establishing interpretive mappings should be among the last adjustments made within a refinement protocol cycle (Lam et al., 2009). It is noteworthy that moving mapping finalization to the end of the protocol takes advantage of the ignorance- and uncertainty-preserving characteristics of the synthetic method (Hunt et al., 2009).

The extent of P-gp inhibition by the inhibitor amounts used during wet-lab experiments was unknown and so needed to be factored into grounding decisions. Consequently, for every cycle of the protocol, I used four PGP levels in each CONTROL and LY-CULTURE. I then selected the pair (one for each) that moved me closer to validation.

Figure 5.2B illustrates the following. An ISECC that is marginally just complicated enough to validate will be fragile to changes in targeted attributes (its phenotype): add any one of a number of wet-lab attributes to the targeted set and the ISECC is at risk of being falsified. The advantage of such an ISECC is that the portion of mechanism and parameterization space that enables validation is small. However, by being fragile to attribute change, it is less biomimetic. I argue that a scientifically more interesting ISECC will be one that has some degree of robustness to changes in phenotype, and that requires having a larger mechanism and parameterization space that validates: many somewhat different mechanisms, each having a number of satisfactory parameterization vectors, a process called multi-modeling (Hunt et al., 2009). For multi-modeling to be successful, it must become methodologically scientific. Scientific M&S will accelerate scientific progress by facilitating fast-paced cycles of hypothesis (about mechanisms) generation, selection, and falsification. Each cycle requires synthetic M&S coupled with inductive and deductive methods; during such a cycle, abduction drives the creation of mechanistic hypotheses. Those mechanistic hypotheses that meet criteria are selected for in silico experimentation designed to ensure that only those with explanatory, heuristic value survive

falsification. The cyclic process exercises, leverages, and enriches the mental models of domain experts in new ways. Multi-modeling that can be semi-automated is the M&S frontier. Nevertheless, during analogue refinement, parsimony remains an important guideline.

A purpose of conducting wet-lab experiments like those cited is to gain new knowledge regarding mechanistic details. Experimenting on synthetic analogues like ISECCs, provides a heretofore unavailable means of discovering new, plausible mechanistic details, especially when relevant wet-lab experiments are highly complicated, costly or impossible. I have demonstrated a new means of achieving deeper insight into the generative mechanisms responsible for phenotype. The approach extends the scientific method to M&S while enabling achieving deeper understandings of pharmacological and therapeutic causal linkages.

6. Conclusion and perspectives

6.1. It is not real

All models are wrong; some are useful.

With that in mind, undoubtedly any and all in silico analogues are flawed, incomplete and, in short, not real. Skeptics are correct that no matter how detailed, thorough, or complicated the mechanisms can be, they are abstractions rendered in software, which may or may not correlate with reality. An unvalidated, ill-informed or poorly constructed analogue could be perfect in performance, but makes poor predictions in reality. In silico simulations, no matter how sophisticated, can never replace wet-lab experiments using real biological parts. No computational model can fully represent the complexity inherent within biological systems and, thus, models cannot be entirely correct.

However, it should be noted that a model – any model – is always an abstraction, a representation, or a description of the referent. In most cases, a model is a reduction from the referent. Only the (small) part of the referent, relevant to answering the scientific question, is preserved, whereas the remained are abstracted away, and replaced by, placeholders or

experimental apparatus. Thus, the intent of using a model is not to replicate the referent or to make perfect predictions or even to be correct. Models are used for some specific purpose, i.e., to answer specific research questions, to test hypotheses, etc. It follows that the model's value is not to be measured in terms of realism, or truism, but rather in its usefulness: whether it serves the purpose, or purposes, of the defined specific use. The advance of science depends on discovering better and more useful models, not "correct" ones.

In this dissertation I present a novel, synthetic approach to model and simulate pharmacokinetic processes. The goal is to provide mechanistic insights into the referent biological system. Specifically, the focus is on improving understanding into biological systems and pharmacokinetic properties, and not on providing precise prediction of pharmacokinetic quantities. I believe that the presented analogues meet my specific aims.

6.2.Limitations

As with all scientific endeavors, the analogues and the methods presented invariably have limitations. In this section, I outline some of them and offer my comments.

The analogues were validated by a small subset of experimental observations. It is entirely possible that the analogues may not validate on similarly conducted experiments. Because the presented analogues were optimized for the reported observations from a specific set of experiments, it is likely that, when tested on different observations from somewhat different experiments, simulated results would present significant discrepancies. How those discrepancies that are to be resolved remain open to questioning, and may require re-parameterizations, specific modifications of rules and assembly, or even complete reengineering. Undeniably, if the analogues were constructed with respect to a larger data set from standardized experiments and protocols, the confidence about the validity and robustness of the model, its rules, assembly, and structure would be much higher. As a result, the conclusion could be made stronger had such a larger data set been available.

The analogues incorporated new, unexpected yet biomimetic in silico mechanisms, and hence I asserted that these in silico mechanisms may have their biological counterparts. Although peripheral evidences have been cited to support the hypothesized mechanisms, those insights remain to be tested, in the future, with carefully designed experiments. Modeling and simulation can never create reality; wet-lab experimentations are needed to confirm the findings. Unfortunately, it was not possible to identify an equally plausible competing mechanism for each of the analogues.

When the analogues are used as an experimental apparatus for new investigations, the simulated results stand as qualitative and quantitative predictions of experimental results. For example, with regard to RISL, I presented predictions of digoxin's pharmacokinetic and metabolic profile when both uptake and efflux inhibitors are pre-administered. However, I did not present validation of those predictions. While I am confident that the simulated results are qualitatively accurate, I do not have the opportunity to quantitatively assess how good the predictions were. Doing so would require replicating the original experiment, including, but not limited to, using the same batch of reagent, the same group of rats, the same apparatus, the same bioanalytical methods, and the same personnel with their surgery and manipulating techniques. I found it to be implausible that one could carry out that replicating experiment. Similarly, from ISECC, I hypothesized that heterogeneous intracellular microenvironments existed and sequestered saquinavir inside Caco-2 cells, making the drug unavailable for metabolism. I did not have direct, wet-lab experimental evidence to support that hypothesis. In order to obtain direct evidence, advanced intracellular imaging technology may be required. Unfortunately, I lacked such technology and its related resources.

While the truthfulness of the conclusions inferred from the analogues cannot be ascertained from simulations alone and while the hypothesized mechanisms may, or may not, play a role in the emergence of the observed phenomena and those predictions may, or may not,

be accurate, one thing is certain: experimenting on synthetic analogues, like the ones presented, provides a heretofore unavailable means of discovering new, plausible mechanistic details. A validated, hypothetical generator in the form of biomimetic mechanisms constructed from object-oriented software components would replace vague, unverified concepts with a concrete instance of a plausible mechanism, although we are just as ignorant of the wet-lab system, now, as we were before. The collection of mechanisms, rules, assembly and interactions of components can be subjected to testing and falsification, and, in the absence of other competing theories, stand as the current best explanation for the phenomena.

6.3.Knowledge discovery and embodiment

The value of the analogues does not solely lie in the (correctness of) the predictions (predicted mechanisms); rather, its value lies in the ability to quickly generate mechanistic hypotheses, and test their plausibility with respect to available data, and in the potential to identify competing mechanistic explanations. Given simulation results, one can prioritize which explanation may merit further targeted experimental exploration. As we have demonstrated, simulations bring mechanism to life: from abstract subjective intangible concepts to concrete, objective, extant instances. Simulations make hypotheses testing, selection, and falsification possible. These cannot be done with conceptual mechanisms.

Often, discrepancies between model expectations and experimental realities are seen as failures of the model to generate useful predictions. The inaccurate predictions lead to discrediting the model and its methods, and to loss of interests in, and value, of the model. The notion that the sole purpose of modeling is to produce useful (accurate) predictions is a result of a simplistic, myopic, and narrow-minded view of the use of modeling in biomedical research. Models are tremendously useful above and beyond generating predictions and testing hypotheses.

As I have demonstrated, building a model about a referent biological system is a primary means to achieving a deeper understanding of the referent. Although unsettling, discrepancies

between model expectations and experimental realities present important opportunities to identify gaps in current knowledge. As scientists, we deal more with what we do not know or do not understand than with what we do know or understand. Science does not advance when experiments merely confirm the current theory. Our knowledge evolves more swiftly when our model predictions are proven inadequate, when our system information is proven incomplete, and when we are surprised by unexpected experimental observations. When a well-constructed, verified model is (unexpectedly) invalidated by experimental findings, it is evidence that the current best knowledge is lacking. Scientists may then explore why the current best model failed, and how it could be improved. When a revised model with new mechanistic details is validated, it represents hypotheses that are previously overlooked or newly discovered. As such, a synthetic analogue concretely organizes our ignorance, if not knowledge, about the referent system: it tells us what we (think we) know and what we do not know. Science is about capturing, refining, and ultimately reducing our ignorance of a given system, or in other words, building better models. One has to know where the model is lacking before building better ones. Hence, a model that reveals our ignorance is as equally valuable, scientifically, as models that just describe what we already know.

In synthetic analogues, such as the RISLs and ISECCs, components and their interactions represent micro-mechanistic features, including anatomical, physiological, and molecular details at different levels during execution. Because of such multi-level similarities, following rounds of improvement, testing, and validation, descendant analogues of this class have the potential to evolve into executable biological knowledge embodiments. While such embodiments are needed, they are beyond the scope of current pharmacokinetic, pharmacodynamic, and related modeling methods. Knowledge embodiment is made feasible because synthetic analogues provide concrete instances of that knowledge, rather than computational descriptions of conceptual representations.

When an analogue is executed, it demonstrates the consequences of our instantiated mechanistic hypotheses, and whether or not they match with details of the referent system.

The envisioned synthetic analogues can facilitate the merger of knowledge and expertise, contributed across organizational domains into executable and, therefore, observable and falsifiable systems of plausible mechanisms and hypotheses. Together, they will represent the current, best theory for aspects of system function. It will be possible to observe different aspects of knowledge in action and to do so from different perspectives, as we do with wet-lab systems. As these analogues become more mature and are validated over a larger range of referent experiments, adjusting (tuning) an analogue to represent heterogeneous systems, and systems under different conditions, will be feasible. For example, one may use an analogue to represent a normal rat liver in one in silico experiment, and use an appropriately adjusted analogue, or a set of heterogeneous analogues, for diseased rat liver in another experiment. One may modify a rat liver analogue to represent a mouse liver or a human liver. These modifications and adjustments are possible because uncertainty can be preserved and cross-validation of component functions can specify which features to change. It will be possible to take copies of the same analogue and tune each separately to reflect differences in measured, patient-specific attributes. The collective knowledge, coupled with collective uncertainty, can be made specific for groups of patients and even for individual patients.

6.4.Scientific modeling and simulation

A vision motivating research on synthetic analogues is to account for the causal basis of the observed data and to shed light on prevailing mechanistic hypotheses about drug dynamics, specifically to gain new knowledge regarding mechanistic details of disposition and metabolism.

To systematically leverage synthetic analogues to achieve that vision, we must follow the scientific method for investigation, the objective of which is knowledge discovery (or questioning and integrating prior knowledge). The method begins with pharmacokinetic phenomena in need

of explanation or investigation. We pose hypotheses and then strive to falsify their predictions through experimentation. One may conduct wet-lab experiments; alternatively, as demonstrated in the previous chapters, we conducted in silico experiments.

Scientific research involving computational analogues — scientific modeling and simulation — is characterized by testing multiple, similarly plausible models, just as abduction requires testing multiple hypotheses and induction requires multiple observations. Note that abduction and induction occur at a level above the computational analogues. The described framework for the scientific use of computational models requires designing and conducting multiple experiments on the analogue, and constructing multiple analogues that merit being objects of experimentation.

The stages in scientific modeling and simulation are illustrated on the right side of Fig 2.1. The assembly of micro-mechanisms in RISLs and ISECCs were hypotheses. Each execution was an in silico experiment. Measures of phenomena during execution provided data. When that data failed to achieve a pre-specified measure of similarity with referent wet-lab data, the mechanism was rejected as a plausible representation of its wet-lab counterpart. The collection of initial hypotheses was then refined iteratively through rational analysis, including experimentation and deduction, in both the minds of the researchers as well as in computer simulations. The iterative model refinement protocol guided cycles of abductive, mechanism-focused, exploratory modeling. Many mechanisms were tested and rejected en route to the mechanisms discussed in the conclusions. Multiple rounds of iterative refinement, followed by mechanistic failure, illustrated the fact that complex conceptual mechanisms can be flawed in ways that were not readily apparent to the researchers. The flaws only became obvious after we actually invested in the effort to implement and test the mechanism synthetically. Those that survived can be further refined in the face of data from newly conducted wet-lab experiments or new inductive models.

At the end of this iterative process, the most robust explanatory and predictive hypotheses can be integrated into larger bodies of theory.

New knowledge comes about by seeking and confronting contrast, anomaly, and surprising or unexpected observations. Studying and explaining the discrepancies between simulation and reality requires abductive reasoning, which is primarily how new theories and hypotheses are first conceived. As I have demonstrated, synthetic modeling and simulation method, coupled with the iterative refinement protocol, facilitate this abduction step in the scientific cycle.

Synthetic analogues are ideal for discovering plausible mechanisms, relations between components, and mechanism-phenotype relationships. They are good at explanation. Experimenting on synthetic analogues provides a powerful new means of discovering and testing the plausibility of mechanistic details. They provide an independent, scientific means to challenge, explore, better understand, and improve any inductive mechanism and, importantly, the assumptions on which it rests.

6.5.Perspective: towards a virtual patient

Few would disagree that medicine should be personalized. To practice the ideal personalized medicine is to answer this question: what is the optimized treatment for a patient who has never before been treated? Currently there are two possible approaches. One is based on domain experts making subjective, empirical recommendations taking into account the unique characteristics of the current patient. Another is using the proven optimal treatment for the group of patients that is sufficiently similar, in measurable attributes, to the current patient, such as age, sex, and genetic markers. An exciting, scientifically rigorous third approach is possible by using a validated virtual patient analogue, capable of exploring the predicted outcomes from available drug treatment options, given the current patient's individual demographics and disease progression, and by selecting the ones that most likely will provide the desired therapeutic

outcome, with the highest probability of success coupled with reduced, characterized risk of adverse events. Advances in proteomics, metabolomics and pharmacogenomics are moving the current practice of medicine to stratified evidence-based medicine. Results from genomic-wide association studies, or epidemiological studies with large cohort, may correlate gene mutations (or markers) to clinically significant outcomes. Predictions about treatment efficacy and toxicity may be made based on results from genetic tests. Patients may be categorized into strata based on genetic makeup and demographics. Treatment options for a defined stratum may be clinically compared, evaluated, and selected.

However, the advances in the -omics technologies generated massive amounts of data of low information value, but little actionable interpretations of any heuristic value, not to mention mechanistic understanding. Translation from genomic information to personalized prescriptions requires concrete, explorable, and traceable genotype-phenotype linkage, and an understanding of how pharmacological and toxicological response are generated from interactions of genes and environment through the hierarchy of mechanisms.

The massive amount of data, especially data in genetic mutations and biomarkers expression, can be used to help develop synthetic, biomimetic tissue, organ and even patient analogues. Using the synthetic modeling and simulation methods, domain experts may organize their theory and uncertainty by building computational analogues. These analogues are validated against the available data. Those that survive falsifications are representation of current best knowledge. New studies can be conducted to fill the gaps in knowledge. The capabilities of the synthetic analogues, listed in Table 2.1, guarantee relatively easy revision, modification, accumulation, and integration of knowledge. Components – knowledge – can be shared between analogues. Validated analogues can be linked and organized into models for biological pathways, tissues, and organs and stand as dynamic, mechanistic models of physiology and disease progression. Semi-automation can greatly increase the efficiency of this process. Finally, the

analogues can be tuned to reflect relevant characteristics of a particular patient, and, in fact, become the *personalized* analogue of the unique patient. Treatments can be given to this unique patient analogue and the outcomes can be compared. Needless to say, these virtual patients are too simple and not real. However, they would be useful as an informative laboratory to support clinical decision- making.

That's the future. The synthetic modeling and simulation method, and the idea of iterative, scientific modeling and simulation are two enabling steps on the path towards developing virtual patient analogues.

The question is not *if* a virtual patient is going to be developed, but *when*, and *how*.

7. Appendix

7.1. Induction

Conditions supportive of all three reasoning methods are sketched in Fig. 5.1B. Induction is arrival at a conjecture (universal conclusion) based on a pattern observed in many particular cases. Induction begins with the measurement of a collection of objects. It is fundamentally and foremost a method for studying phenomena, not mechanism. The data from which a pattern is induced act as a statement *solely* about the phenomena for which the measure was designed. Hence, every data set, and subsequently every inductive model, has embedded in it the aspects or usage protocols plus premises commensurate with the measures used to obtain the data. In general, a model induced from large sets of the same type of data will be more precise, more specific, and more suitable for prediction. On the other hand, a model induced from lots of variant data (measures of different phenotypic attributes) will be less precise but more resilient and more general; its predictions will reflect greater uncertainty. In the former context, one has to worry most about over-fitting data, which makes the model too specific to a single data set to be useful. At the same time, one has to worry about inducing a model that is too general and misses crucial patterns in the data.

Often, the modeler is also interested in the generative mechanisms responsible for the data. In those cases, features of idealized, conceptual constructs (a feedback circuit, a two-compartment model, etc.) can provide mechanistic insight. This is particularly true when the mathematical description of an idealized measure of a construct feature (given a set of premises) is known to have a form that is the same (or nearly the same) as the induced mathematical description (a sum of exponentials, for example). How well the parameterized features of the conceptual construct map to components and features of the referent, along with the acceptability of the premises, are separate issues outside the current scope. The initially induced mathematical description is prosaically expanded to include the conceptual features, but remains a hypothesis about patterns in the data.

7.2.Deduction

Deduction is automatic and/or mechanical transformation of a set of statements. It is the purely mechanical (syntactic) transformation of the premises to a conclusion. As such, no meaning (semantics) need exist for deductive systems. All executing computer programs (absent human or real-world interactions) are deductive systems. Likewise, mathematical transformations are also deductive systems. The most fundamental element of deduction is the engine that actually makes the transformation. In the case of a computer program, the instruction pointer provides the impetus for transformation. In mathematics, the engine is the human manipulating the symbols. A simulation is an operating, deductive system designed to mimic the behavior of some referent. Its alphabet and grammar are specified in part by the language in which the program is written and in part by the constructs the programmer creates. The premises are statements about the initial conditions of the program. The conclusions are statements about the final conditions of the program. The conclusions (of interest) become the outputs or "behaviors" of the simulation, but no new knowledge can be created. Any meaning applied to the

premises, grammar, or conclusions are inferred by those examining the program and simulation, and so remain conceptual and hypothetical.

7.3. Abduction

Abduction is arrival at a conjecture based on a pattern observed in one or a few particular cases. Abduction is conceptualizing (multiple) mechanisms: explanations that, if true, could account for or generate a similar anomalous, interesting, or surprising observation. Abductive inference involves hypothesis generation and selection, and is an important occurrence during wet-lab research. Abduction is most likely and appropriate for ambiguous systems (the left side of Fig. 5.1B). The task of resolving the ambiguity is best approached through multiple aspects and with multiple mechanistic hypotheses measured with multiple measures.

To illustrate abductive inference, consider the following situation: measurements of an experimental treatment group exhibit unexpected values when compared to those of control groups and data from past experiments. Further, the results do not fit well with known categories of similar phenomena. In such a situation, researchers offer many speculative, candidate explanations: were this condition or circumstance true, it could explain the anomalous or new observations. Some explanations may focus on material used in the experiment (a possible bad batch of reagent, etc.). Other explanations may focus on the conduct of the experiment (the samples may have been mishandled, etc.). Others are ideas about mechanistic explanations. Generation of varied explanatory hypotheses, some highly speculative, following the observation of the anomalous behavior is part of abduction. The next phase involves a process that narrows the competing ideas to those deemed most plausible. Following abduction, the consequences of these hypotheses are logically or experimentally deduced, and then evaluated using induction. After testing, when the set of plausible hypotheses is dramatically reduced, those remaining represent the current best explanation(s) until some new observation falsifies one or more of them. At that stage, the entire scientific reasoning cycle may repeat itself.

The above cycle (Fig. 5.2) occurs frequently when engineering and refining a synthetic analogue, like ISECCs. The behaviors of the first implementation often fall short, frequently far short of expectations, even though it is the modeler's best hypothesis about how components should plug together to obtain the targeted phenomena. That shortfall falsifies the best hypothesis. The modeler has learned that the selected region of mechanism space is too abiotic. The modeler must rethink plausible micro-mechanisms. The solution is to jump to another region of mechanism space and experiment to determine if the new micro-mechanisms produce phenomena that are more similar to targeted phenomena. So doing exercises creativity. When improvements are seen, the modeler can conjecture that the new micro-mechanisms are more biotic. Each failed cycle exercises thinking creatively about plausible mechanisms. Each improvement in the similarity of the analogue's behavior to that of the referent adds new knowledge and improves insight into referent mechanisms. Both failed and fruitful cycles are often characterized by abductive reasoning.

Like induction, abduction starts with a measure selected by the researcher. As such, the hypothetical mechanisms inferred (current beliefs) are inherently and irrevocably dependent upon the measure through which the phenomena are defined and revealed. They are aspect and perspective dependent. Change aspect and/or perspective and the hypothetical mechanisms inferred may change. Unlike induction, however, abduction does not necessarily produce (overly) precise or (overly) general mechanistic explanations. More often, as with ISECCs, they are specific to the observation or experiment and its context. That is because the focus of abduction is on the current few interesting cases. The hypotheses (explanatory models) must be elaborated through deduction followed by validation through induction in order to learn how precise or general each hypothetical mechanism actually is. For that reason, abduction preserves ignorance, in contrast to the truth preservation of deduction. The researcher is just as ignorant after abducing an explanation as before.

Abduction references: Yu CH (1994) Abduction? Deduction? Induction? Is there a logic of exploratory data analysis? http://www.creative-wisdom.com/pub/Peirce/Logic_of_EDA.html (accessed 8/19/09). Magnani L (2000) *Abduction, reason and science – processes of discovery and explanation*. Kluwer Academic/Plenum Publishers, NY, NY. Gabbay DM and Woods J (2005) *A practical logic of cognitive systems, Volume 2: The reach of abduction: insight and trial*. Elsevier.

8. Reference

Arab HA, Cheung K, Hickman PE, Potter JM, Kadkhodaee M, and Roberts MS (2007) Effects of hypoxia/reperfusion injury on drug disposition in the rat isolated perfused liver. *Clin Exp Pharmacol Physiol* **34(4):**332-338.

Artursson P and Borchardt RT (1997) Intestinal drug absorption and metabolism in cell cultures: Caco-2 and beyond. *Pharm. Res* **14:**1655-1658.

Artursson P and Karlsson J (1991) Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco-2) cells. *Biochem Biophys Res Comm* **175:**880-885.

Bailey SM and Reinke LA (2000) Effect of low flow ischemia-reperfusion injury on liver function. *Life Sci* **66:** 1033-1044.

Benet LZ (1972) General treatment of linear mammillary models with elimination from any compartment as used in pharmacokinetics. J *Pharm Sci* **61(4):**536-541.

Bessems M, 't Hart NA, Tolba R, Doorschodt BM, Leuvenink HG, Ploeg RJ, Minor T, and van Gulik TM (2006) The isolated perfused rat liver: Standardization of a time-honoured model. *Lab Anim* **40(3):**236-246.

Borchardt RT (1995) The application of cell culture systems in drug discovery and development. *J.Drug Targeting* **3:**179-182.

Booth CL, Brouwer KR, and Brouwer KL (1998) Effect of multidrug resistance modulators on the hepatobiliary disposition of doxorubicin in the isolated perfused rat liver. *Cancer Res* **58:** 3641-3648.

Booth CL, Pollack GM, and Brouwer KL (1996) Hepatobiliary disposition of valproic acid and valproate glucuronide: use of a pharmacokinetic model to examine the rate-limiting steps and potential sites of drug interactions. *Hepatology* **23:** 771-780.

Dantzig AH, Shepard RL, Law KL, Tabas L, Pratt S, Gillespie JS, Binkley SN, Kuhfeld MT, Starling JJ, and Wrighton SA (1999) Selectivity of the multidrug resistance modulator, LY335979, for P-glycoprotein and effect on cytochrome P-450 activities. *J Pharmacol Exp Ther* **290**:854–862.

de Lannoy IA and Silverman M (1992) The MDR1 gene product, P-glycoprotein, mediates the transport of the cardiac glycoside, digoxin. *Biochem Biophys Res Commun* **189:** 551-557.

DiStefano JJ 3rd. (1982). Noncompartmental vs. compartmental analysis: Some bases for choice. *Am J Physiol* **243(1):**R1-R6.

DiStefano JJ 3rd and Landaw EM (1984) Multiexponential, multicompartmental, and noncompartmental modeling. I. Methodological limitations and physiological interpretations. Am *J Physiol* **246(5 Pt 2):**R651-R664.

Eagling VA, Wiltshire H, Whitcombe IW, and Back DJ (2002) CYP3A4-mediated hepatic metabolism of the HIV-1 protease inhibi-tor saquinavir in vitro. *Xenobiotica* **32**:1-17.

Engelberg JA, Ropella GE, and Hunt CA (2008) Essential operating principles for tumor spheroid growth. *BMC Syst Biol* **2**:110.

Fisher J and Henzinger TA (2007) Executable cell biology. *Nat Biotechnol* **25:** 1239-1249.

Fitzsimmons ME and Collins JM (1997) Selective biotransformation of the human immunodeficiency virus protease inhibitor saquinavir by human small-intestinal cytochrome P4503A4: potential contribution to high first-pass metabolism. *Drug Metab Dispos* **25**:256–266.

Gao J, Hugger ED, Beck-Westermeyer MS, and Borchardt RT (2000) Estimating intestinal mucosal permeation of compounds using caco-2 cell monolayers. *Curr. Protoc. Pharmacol* **S8:**7.2.1-7.2.23.

Garmire LX, Garmire DG, and Hunt CA (2007) An in silico transwell device for the study of drug transport and drug-drug interactions. *Pharm Res* **24:** 2171-2186.

Garmire LX and Hunt CA (2008) In silico methods for unraveling the mechanistic complexities of intestinal absorption: metabolism-efflux transport interactions. *Drug Metab Dispos* **36**:1414-1424.

Gores GJ, Kost LJ, and LaRusso NF (1986) The isolated perfused rat liver: Conceptual and practical considerations. *Hepatology* **6(3):**511-517.

Hidalgo IJ (2001) Assessing the absorption of new pharmaceuticals. *Curr Top Med Chem* **1(5):**385-401.

Hidalgo IJ, Raub TJ, and Borchardt RT (1989) Characterization of the human colon carcinoma cell line (Caco-2) as a model system for intestinal epithelial permeability. *Gastroenterology* **96(3):**736-749.

Hung DY, Siebert GA, Chang P, and Roberts MS (2005) Hepatic pharmacokinetics of taurocholate in the normal and cholestatic rat liver. *Br J Pharmacol* **145(1):**57-65.

Hunt CA, Ropella GE, Lam TN, Tang J, Kim SH, Engelberg J, and Sheikh-Bahaei (2009) At the Biological Modeling and Simulation Frontier. *Pharm Res* **26**:2369-2400.

Hunt CA, Ropella GE, Park S, and Engelberg J (2008) Dichotomies between computational and mathematical models. *Nat Biotechnol* **26**:737-738.

Hunt CA, Ropella GE, Yan L, Hung DY, and Roberts MS (2006) Physiologically based synthetic models of hepatic disposition. *J Pharmacokinet Pharmacodyn* **33**:737-772.

Jang GR, Harris RZ, and Lau DT (2001) Pharmacokinetics and its role in small molecule drug discovery research. *Med Res Rev* **21(5):**382-396.

Kato Y, Tanaka J, and Koyama K (2001) Intralobular heterogeneity of oxidative stress and cell death in ischemia-reperfused rat liver. *J Surg Res* **95(2):**99-106.

Khoo SH, Hoggard PG, Williams I, Meaden ER, Newton P, Wilkins EG, Smith A, Tjia JF, Lloyd J, Jones K, Beeching N, Carey P, Peters B, and Back DJ (2002) Intracellular accumulation of human immunodeficiency virus protease inhibitors. *Antimicrob Agents Chemother* **46**:3228-3235.

Kim AE, Dintaman JM, Waddell DS, and Silverman JA (1998) Saquinavir, an HIV protease inhibitor, is transported by P-glycoprotein. *J Pharmacol Exp Ther* **286**:1439-1445.

Kim SH, Park S, Mostov K, Debnath J, and Hunt CA (2009) Computational investigation of epithelial cell dynamic phenotype in vitro. *Theor Biol Med Model* **6**:8.

Kim SH, Yu W, Mostov K, Matthay MA, and Hunt CA (2009) A computational approach to understand in vitro alveolar morphogenesis. *PLoS One* **4**(3):e4819.

Kurakin A (2009) Scale-free flow of life: on the biology, economics, and physics of the cell. *Theor Biol Med Model* **6**:6.

Lam TN and Hunt CA (2008) Mechanistic simulations explain paradoxical saquinavir metabolism during in vitro vectorial transport study. *Conf Proc IEEE Eng Med Biol Soc* **2008**:5462-5465.

Lam TN and Hunt CA (2009) Discovering plausible mechanistic details of hepatic drug interactions. *Drug Metab Dispos* **37**:237-246.

Lam TN, Ropella GE, and Hunt CA (2009) Focus on discovering mechanisms: a relativistic, agent-based, recirculating liver, in *Proceedings of the Spring Simulation Multiconference 2009, Agent-Directed Simulation Symposium* (ADS'09), The Society for Modeling and Simulation International, San Diego, CA.

Lau YY, Wu CY, Okochi H, and Benet LZ (2004) Ex situ inhibition of hepatic uptake and efflux significantly changes metabolism: hepatic enzyme-transporter interplay. *J Pharmacol Exp Ther* **308:** 1040-1045.

Liao J, Keiser JA, Scales WE, Kunkel SL, and Kluger MJ (1995) Role of epinephrine in TNF and IL-6 production from isolated perfused rat liver. *Am J Physiol* **268(4 Pt 2):**R896-901.

Liu L, Mak E, Tirona RG, Tan E, Novikoff PM, Wang P, Wolkoff AW, and Pang KS (2005) Vascular binding, blood flow, transporter, and enzyme interactions on the processing of digoxin in rat liver. *J Pharmacol Exp Ther* **315(1):**433-48.

Liu Y and Hunt CA (2005) Studies of intestinal drug transport using an in silico epithelio-mimetic device. *Biosystems* **82**:154-167.

Liu Y and Hunt CA (2006) Mechanistic study of the cellular interplay of transport and metabolism using the synthetic modeling method *Pharm Res* **23:** 493-505.

Liu Y, Weber SJ, and Onua ET (2004) Hepatic clearance and drug metabolism using isolated perfused rat liver. *Curr. Protoc. Pharmacol* **S26:**7.9.1-7.9.10.

Matsuura Y, Nishi S, Kariya N, Shimadzu K, and Asada A (2001) The effects of norepinephrine and prostaglandin E1 on pharmacokinetics of lidocaine in isolated perfused rat liver. *Life Sci* **68:** 2123-2129.

Meunier V, Bourrié M, Berger Y, and Fabre G (1995) The human intestinal epithelial cell line Caco-2; pharmacological and pharmacokinetic applications. *Cell Biol Toxicol* **11(3-4):**187-194.

Miller LL, Bly CG, Watson ML, and Bale WF (1951) The dominant role of the liver in plasma protein synthesis: Direct study of isolated perfused rat liver with aid of lysine-ν-C14. *J Exp Med* **94:**431-453.

Mouly SJ, Paine MF, and Watkins PB (2004) Contributions of CYP3A4, P-glycoprotein, and serum protein binding to the intestinal first-pass extraction of saquinavir. *J Pharmacol Exp Ther* **308**:941-948.

Nakashima E and Benet LZ (1989) An integrated approach to pharmacokinetic analysis for linear mammillary systems in which input and exit may occur in/from any compartment. *J Pharmacokinet Biopharm* **17(6):**673-686.

Nestorov I (2003) Whole body pharmacokinetic models. *Clin Pharmacokinet* **42(10):**883-908.

Noé B, Hagenbuch B, Stieger B, and Meier PJ (1997) Isolation of a multispecific organic anion and cardiac glycoside transporter from rat brain. *Proc Natl Acad Sci USA* **94:** 10346-10350.

Pang KS and Rowland M (1977) Hepatic clearance of drugs. I. Theoretical considerations of a "well-stirred" model and a "parallel tube" model. Influence of hepatic blood flow, plasma and blood cell binding, and the hepatocellular enzymatic activity on hepatic drug clearance. *J Pharmacokinet Biopharm* **5(6):**625-653.

Park S, Ropella GE, Kim SH, Roberts MS, and Hunt CA (2009) Computational strategies unravel and trace how liver disease changes hepatic drug disposition. *J Pharmacol Exp Ther* **328**:294-305.

Parker AJ and Houston JB (2008) Rate-limiting steps in hepatic drug clearance: comparison of hepatocellular uptake and metabolism with microsomal metabolism of saquinavir, nelfinavir, and ritonavir. *Drug Metab Dispos* **36**:1375-1384.

Press B and Di Grandi D (2008) Permeability for intestinal absorption: Caco-2 assay and related issues. *Curr Drug Metab* **9(9):**893-900.

Rescigno A (2010) Compartmental analysis and its manifold applications to pharmacokinetics. *AAPS J* **12(1):**61-72.

Rescigno A (2004) On the use of pharmacokinetic models. *Phys Med Biol* **49:** 4657–4676.

Rescigno A (2001) The rise and fall of compartmental analysis. *Pharmacol Res* **44(4):**337-342.

Rescigno A and Beck JS (1987) The use and abuse of models. *J Pharmacokinet Biopharm* **15(3):**327-344.

Schmiedlin-Ren P, Thummel KE, Fisher JM, Paine MF, Lown KS, and Watkins PB (1997) Expression of enzymatically active CYP3A4 by Caco-2 cells grown on extracellular matrix-coated permeable supports in the presence of 1,25-dihydroxyvitamin D3. *Mol Pharmacol* **51**:741–754.

Schwab AJ, Tao L, Kang M, Meng L, and Pang KS (2003) Moment analysis of metabolic heterogeneity: conjugation of benzoate with glycine in rat liver studied by multiple indicator dilution technique. *J Pharmacol Exp Ther* **305(1):**279-289.

Sheikh-Bahaei S and Hunt CA (2006) Prediction of in vitro hepatic biliary excretion using stochastic agent-based modeling and fuzzy clustering, in *Proceedings of the 37th Conference on Winter Simulation, Monterey, CA, Dec 03-06* (L. F. Perrone, et al., eds.) pp 1617-1624, INFORMS, Hanover, MD. Straatsburg IH and Frederiks WM (1997) In situ analysis of ischaemia/reperfusion injury in rat liver studied in three different models. *Int J Exp Pathol* **78:** 149-161.

Su Y, Zhang X, and Sinko PJ (2004) Human organic anion-transporting polypeptide OATP-A (SLC21A3) acts in concert with P-glycoprotein and multidrug resistance protein 2 in the vectorial transport of Saquinavir in Hep G2 cells. *Mol Pharm* **1**:49-56.

Sun H, Chow EC, Liu S, Du Y, and Pang KS (2008) The Caco-2 cell monolayer: usefulness and limitations. *Expert Opin Drug Metab Toxicol* **4(4):**395-411.

Tang J, Ley KF, and Hunt CA (2007) Dynamics of in silico leukocyte rolling, activation, and adhesion. *BMC Syst Biol* **1:** 14.

Tirona RG and Pang KS (1999) Bimolecular glutathione conjugation kinetics of ethacrynic acid in rat liver: in vitro and perfusion studies. *J Pharmacol Exp Ther* **290:** 1230-1241.

Tran H, Robinson S, Mikhailenko I, and Strickland DK (2003) Modulation of the LDL receptor and LRP levels by HIV protease inhibitors. *J Lipid Res* **44**:1859-1869.

Van Breemen RB and Li Y (2005) Caco-2 cell permeability assays to measure drug absorption. *Expert Opin Drug Metab Toxicol* **1(2):**175-185.

Vernochet C, Azoulay S, Duval D, Guedj R, Cottrez F, Vidal H, Ailhaud G, and Dani C (2005) Human immunodeficiency virus protease inhibitors accumulate into cultured human adipocytes and alter expression of adipocytokines. *J Biol Chem* **280**:2238-2243.

Wacher VJ, Silverman JA, Zhang Y, and Benet LZ (1998) Role of P-glycoprotein and cytochrome P450 3A in limiting oral absorption of peptides and peptidomimetics. *J Pharm Sci* **87**:1322-1330.

Wagner JG (1973). A Modern View of Pharmacokinetics. *J Pharmacokinet Biopharm* **1(5):**363-401.

Weisiger RA (1996) When is a carrier not a membrane carrier? The cytoplasmic transport of amphipathic molecules. *Hepatology* **24**:1288-1295.

Weisiger RA (2002) Cytosolic fatty acid binding proteins catalyze two distinct steps in intracellular transport of their ligands. *Mol Cell Biochem* **239**:35-43.

Weisiger RA (2007) Mechanisms of intracellular fatty acid transport: role of cytoplasmic-binding proteins. *J Mol Neurosci* **33**:42-44.

Wolkoff AW, Johansen KL, and Goeser T (1987) The isolated perfused rat liver: Preparation and application. *Anal Biochem* **167(1):**1-14.

Xiong H, Suzuki H, Sugiyama Y, Meier PJ, Pollack GM, and Brouwer KL (2002) Mechanisms of impaired biliary excretion of acetaminophen glucuronide after acute phenobarbital treatment or phenobarbital pretreatment. *Drug Metab Dispos* **30(9)**:962-969.

Yamada K, Hosokawa M, Fujimoto S, Fujiwara H, Fujita Y, Harada N, Yamada C, Fukushima M, Ueda N, Kaneko T, Matsuyama F, Yamada Y, Seino Y, and Inagaki N (2008) Effect of corosolic acid on gluconeogenesis in rat liver. *Diabetes Res Clin Pract* **80(1):**48-55.

Yan L, Ropella GE, Park S, Roberts MS, and Hunt CA (2008a) Modeling and simulation of hepatic drug disposition using a physiologically based, multi-agent in silico liver. *Pharm Res* **25:** 1023-1036.

Yan L, Sheihk-Bahaei S, Park S, Ropella GE, and Hunt CA (2008b) Predictions of hepatic disposition properties using a mechanistically realistic, physiologically based model. *Drug Metab Dispos* **36:** 759-768.

Yates JW, Jones RD, Walker M, and Cheung SY (2009) Structural identifiability and indistinguishability of compartmental models. *Expert Opin Drug Metab Toxicol* **5(3):**295-302.

Yoshihara S, Harada K, and Ohta S (2000) Metabolism of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) in perfused rat liver: involvement of hepatic aldehyde oxidase as a detoxification enzyme. *Drug Metab Dispos* **28(5):**538-543.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

9/6/2010

Date