Prepublication Author Copy

Moving Beyond in Silico Tools to in Silico Science in Support of Drug Development Research

C. Anthony Hunt* and Glen E.P. Ropella

Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco; and Tempus Dictum, Inc.

*__Correspondence:__
Prof. C. Anthony Hunt
513 Parnassus Ave., S-822
San Francisco, CA 94143-0912, USA
Phone: 415-476-2455
FAX: 415-514-2008
E-mail: a.hunt@ucsf.edu

Running Title: In Silico Science in Support of Drug Development Research

Key words: mechanistic; modeling; simulation; agent-based; predict; translation

List of abbreviations: ISL: In Silico Liver; SS: Sinusoidal Segment; VS: visionary scenario

Abstract

Exploitation of concretized mechanistic models and simulation methods enables acquiring a competitive advantage through deeper, easily shared, mechanistic insight into the disease and/or health phenomena that are the focus of the R&D organization. The models are analogues of the biological wet-lab models used to support that R&D. An analogue is an explanatory and evolving hypotheses about the mechanistic consequences of xenobiotic or biologic interventions. As such they are fundamentally different from the familiar inductive, equation based, pharmacokinetic, pharmacodynamic, and related models. Analogues are designed for experimentation and to be useful in the face of incomplete data and multiple uncertainties. They use interchangeable components and require iterative refinement. They enable linking coarse-grained systemic phenomena to fine-grained molecular details, including molecular targets. To simplify and focus discussion we describe one example of the new class of models, In Silico Livers. We present a vision of how the biological wet-lab side of the R&D process might function when these models and methods are fully implemented within a common computational framework. Accumulated mechanistic knowledge is easily measured and visualized in action, and therefore it can be easily challenged. Components within analogues that have been validated for many compounds can use programmed “intelligence” to automatically parameterize for, and respond to, a new, not previously seen compound based on its physicochemical properties. Each analogue can be tuned to reflect differences in experimental conditions and individuals, making translational research more concrete while moving us closer to personalized medicine.

Introduction

Acquiring deeper, more useful and exploitable mechanistic insight into disease and/or health phenomena of interest will improve research productivity and provide a competitive advantage. Because the phenomena emerge from complex systems, we must rely on models to provide mechanistic insight. Mechanistic models are hypotheses about how we think phenomena are generated. To acquire deeper insight, we need more explanatory mechanistic models. When dealing with biology, having explanatory mechanistic models necessarily precedes having predictive mechanistic models. A truly useful explanatory mechanistic model is one where we can observe putative cause-effect events at several layers as they unfold. Unfortunately, the vast majority of current mechanistic biological models are conceptual. Typically, diagrams support a prosaic description of a hypothesized, plausible mechanism. In many cases, equations are used to make predictions about patterns in data. The equations often describe hypothetical measures of features of the conceptual mechanism, such as the concentration time-course of drug at some idealized target site, if some string of assumptions are all true. A problem with conceptual mechanistic models is that they are difficult to falsify [Hunt et al., 2009]. Conceptual models have the benefit of making intuitive and analytic sense, trading that for concrete, personalized detail. On the other hand, synthetic models have the benefit of particular, specific detail capable of representing the most precise phenomena, but to do so they trade off some generality. Often, the only way to challenge a conceptual mechanistic hypothesis is to design and conduct wet-lab experiments. For many mechanistic hypotheses the experiments are unethical, too costly, or currently infeasible. Nevertheless, to acquire deeper insight, it is essential to have means of mechanism falsification. Mechanism validation provides no new knowledge. Consequently, insight is not improved. Mechanism falsification, on the other hand, provides new knowledge and improved insight. Hypothesis falsification is thus the primary source of new scientific knowledge.

An objective of this report is to present relatively new modeling and simulation methods that have been designed specifically to facilitate discovery of plausible multilayer, multiscale, multi-attribute mechanisms that survive repeated attempts of falsification or achieve degrees of validation using in silico experimentation. We begin with a vision of how the biological wet-lab side of the R&D process might function when these methods are fully implemented to stimulate drug (new therapeutic entity) discovery, accelerate and streamline development, and lower wet-lab costs. To simplify and focus discussion we describe one example of the new class of models, In Silico Livers [Yan et al., 2008a,b; Park et al., 2009; Park et al., 2010]. We work backwards from the vision below to the enabling essentials by describing the characteristics of the model types needed.

Note that hereafter, when describing features and components of an in silico analogue, we use SMALL CAPS. So doing avoids confusion. It makes clear that they are not the same as the in vitro/in vivo biological counterparts to which they map.

An Achievable, Future R&D Process

Imagine an R&D organization that uses 30 in vitro (including several high throughput cell culture screens) and 15 animal models (all mammalian) to support its drug (or biologic) discovery and development activities in two disease areas. The envisioned R&D strategy (Fig. 1) calls for having evolving, in silico analogues of each wet-lab model that have the capabilities listed in Table 1. The analogues are constructed and undergo validation using the synthetic method of modeling and simulation [Hunt et al., 2009]. Many of their components are interchangeable. They exist within a common computational framework. For each wet-lab model there corresponds several hierarchical, multi-attribute analogues, each exhibiting essentially indistinguishable phenotypes when dosed with in silico counterparts to any one of a common set of compounds. Having multiple analogues of the same referent is an acknowledgement and a representation of biological and pharmacological uncertainty that can shrink, but not vanish.

Within the framework, a cadre of quasi-autonomous agents, represented as software constructs, (Experiment Agents) manage in silico experiments, translations, and exchanges between analogues. A software agent is an autonomous object that schedules its own events and interacts with other agents and objects in its environment. The protocols and mechanistic interventions of wet-lab experiments all have in silico counterparts. For most compounds studied previously, results from in silico experiments are indistinguishable from corresponding wet-lab results. Most of the biomimetic components within an

analogue are able to distinguish (and adjust their responses to) different COMPOUNDS by reading the attached list of the physiochemical properties for the referent compound. Most analogue components are "intelligent": they are programmed with the results of many earlier validation and falsification experiments and can arrive at a customized parameterization that determines how they interact with a new compound. Experimental results for new (not previously seen) COMPOUNDS can be obtained from all in silico analogues. When a new compound's structure is somewhat similar to those of several compounds already studied in one or more wet-lab models, response predictions from all analogues may be acceptably reliable. The framework also allows embedding such "intelligence" into the DRUG object, promoting it from passive object to an active component of the analogue. So doing would also allow one to scale the model to characterize families of compounds as well as particular molecules. As it matures, the framework can be treated as an in silico domain expert. When a new COMPOUND is quite different from any previously studied COMPOUNDS, response predictions may be equivalent to the best available expert, ballpark predictions.

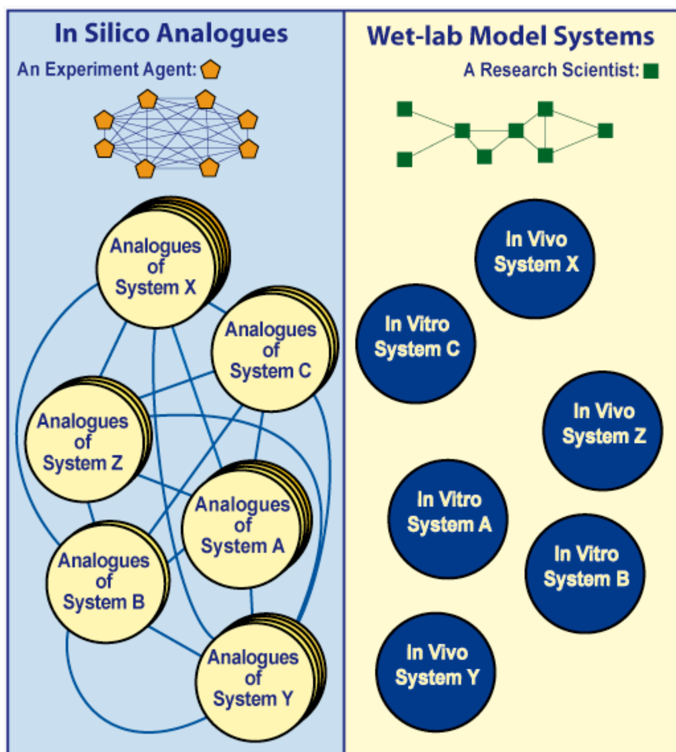

Fig. 1. Networked in silico counterparts to wet-lab models. Future pharmaceutical R&D companies are envisioned having in silico counterparts to each wet-lab (in vitro and animal) model used to support pipeline development. Collectively, the system on the left is a concretized, executable, observable, interactive knowledge embodiment. Analogue mechanisms represent everything the organization knows or thinks it knows about its wet-lab models and the disease areas of focus. Each analogue on the left, like the ISL in Fig. 2, is constructed and validated using synthetic methods of modeling and simulation. Components are exchangeable, object-oriented, and agent based. For each wet-lab model there are several analogues that are somewhat different mechanistically. Differences map to mechanistic and biological uncertainties. Each analogue is an abstract, multiscale and hierarchical system capable of mimicking multiple referent attributes in the presence and absence of different compounds and other types of interventions. All analogues "live" within a common framework. An Experiment Agent is an in silico counterpart to a research scientist member of the wet-lab research team. Wet-lab scientists interact and exchange information using an often-inefficient social network (top). Experiment Agents manage analogue operations, conduct in silico experiments, analyze results, and manage and coordinate the direct exchange of information between analogues (connecting lines between analogues), parameterizations, and components. Direct interaction among Experiment Agents is directed and specified by the framework.

Results of large numbers of low cost in silico experiments can be used to more thoroughly explore far more of the expected response and phenotype landscapes of wet-lab models than is currently economically possible. The insight gained from such explorations can guide design of a critical few wet-lab experiments, in a region of high uncertainty for specific doses. When experimental results align sufficiently with analogue predictions, decision makers may increase the trust placed in the analogue's coverage of the wet-lab model's phenotype, and decide that no further testing is needed: the available knowledge and uncertainty are sufficient for decision making. Discrepancies between wet-lab observations and in silico behaviors for one analogue can be used by the Experiment Agents to revise and improve the micromechanisms (specific event types) and algorithms that make analogue components more "intelligent". Those technical improvements along with the new knowledge represented can be used by all other analogues and their components to make adjustments in light of the new knowledge. When additional experiments are conducted using a new COMPOUND, some outcomes may be different from those from predecessor analogues. Those differences may suggest new lines of wet-lab study or additional analogues or mechanisms that merit exploration.

In the above visionary scenario (VS), the combined operation of the framework, the

analogues, their "intelligent" components, etc.—everything on the left side of Fig. 1—is an embodiment of most of what the organization knows about the two disease areas and the compounds that they have studied to date. The framework and everything within instantiates what has been learned by experimentation on the wet-lab models coupled with knowledge and insights gleaned from the literature. The framework and its content are an up-to-date instantiation of all accumulated, new, and proprietary mechanistic knowledge in an accessible, easily understood, observable, and interactive form. As a consequence, the company has reduced dependency on the corresponding, difficult to challenge, conceptual mechanistic mental models of current and past domain experts. Today, much of the mechanistic insight emerging from R&D resides within mental models where differences, similarities, and inconsistencies are difficult—often impossible—to ascertain. The valuable insight is in a form that can easily dissipate and is difficult to retain and control.

The analogues within the framework of the VS will not be limited to wet-lab models. There will be abstract yet individualizable patient analogues along with models of translation [Hunt et al., 2008]. Park et al. [2010] provide an example of the latter. Translation models show, for example, how networked mechanisms, components, and aspects of an analogue can be morphed into features of a generic human analogue. Another translational model might show how micromechanistic details are morphed between analogues of in vitro rat and human hepatocyte cultures. The morphing process shows what must be added and what is lost in translation. Such morphings will provide easily understood, mechanistic interpretation of how cause-effect relationships resulting from an experimental intervention in a wet-lab model are believed to manifest (or not) in the human analogue. The expectation is that those relationships have wet-lab and human counterparts.

Additional exploitable, efficiency-enhancing uses can be described for the VS. However, the ideas presented are sufficient to make the case that there are excellent reasons to initiate a step-wise transformation of the current R&D reality toward that of the VS. How technically challenging will it be to realize the VS? Are the requirements so far beyond the resources of a large pharmaceutical company that it will require a massive undertaking by an agency such as the National Institutes of Health in the USA? No. The task can be undertaken and completed in stages by a small or medium sized organization. Hunt et al. [2009] have argued that realization of the VS requires computational scientists to evolve and maintain the framework along with five essential in silico technologies and several new skills. 1) It needs a computational framework that supports the concurrent operation of multiple models (and model types), some of which will contain thousands of quasi-autonomous agents. Several frameworks with that capability exist, including Mason [Luke et al., 2005], RePast [North et al., 2006], and Swarm [Minar et al., 1996]. 2) It needs a method to build, iteratively revise, and keep current mechanistic analogues of most or all wet-lab experimental systems used to support a organization's research. The synthetic method of modeling and simulation [Hunt et al., 2009] is such a method; although to date there are only a few examples of validated, biomimetic, in silico analogues. 3) The iterative analogue refinement process described below needs to be automatable. Research is needed to achieve that goal. 4) When the mechanisms of one analogue are revised after being challenged by the results of new wet-lab experiments, that new mechanistic insight needs to be transferred (in time, automatically) to the Experiment Agents that manage other analogues with similar micromechanisms as that being revised. Below we discuss a strategy for moving in that direction with HEPATOCYTE CULTURES and the In Silico Liver described below. 5) Methods are needed to make analogue components sufficiently "intelligent" so that during simulations the generative consequences of their self-parameterizations, upon reading the new COMPOUND'S physiochemical properties are consistently within a ballpark that is sufficiently small, and biomimetically constrained to support critical decision-making. A rationale and strategy for achieving HEPATOCYTE "intelligence" in responding to COMPOUNDS is presented below.

An In Silico Liver Designed for Use Within the Visionary Scenario

No forward or reverse engineering path has been described to achieve the VS while relying exclusively n inductive computational and systems biology methods. That is primarily because those methods are not intended for the types and variety of uses within the VS. Although the methods have been used to build several virtual tissues and organs (for examples, see http://nsr.bioeng.washington.edu/), they were not built for uses of the type described above. The first consideration in any modeling effort is to state why models are being created and how they will be used. The VS is such a statement. Given the VS, we can state specifications needed for its achievement. Those specifications are embodied within the syn-

thetic method of modeling and simulation and the analogues constructed.

The In Silico Liver (ISL) is a product of the synthetic method of modeling and simulation. It is an example of an analogue that has been designed for use in an environment like that illustrated in Fig. 1. It has evolved through several refinement cycles [Yan et al., 2008a,b; Park et al., 2009; Park et al., 2010]. See [Sheikh-Bahaei, et al., 2010; Tang and Hunt 2010; Lam and Hunt 2010; Kim et al., 2010; Kim et al., 2009a,bc; Engelberg et. al., 2009] for examples of other products of the synthetic method. An ISL is a simulation framework: an in silico counterpart to an entire wet-lab experimental system (analytical instrumentation and all) built to study the hepatic disposition of xenobiotics in perfused rat livers. An ISL comprises an experiment agent, a data management module, a statistical observer module (used to analyze data), a parameter manager, plus data from the perfusion experiments (and interpolations), *RefModel*, and LOBULE. *RefModel* is the parameterized, classical pharmacokinetic model used by Hung et al. [2002]. Concurrent execution of it and an ISL enables judging similarities or lack thereof. A LOBULE is one Monte Carlo variant of the complete analogue system illustrated in Fig. 2. Park et al. [2010] reported three LOBULE variants (NORMAL and two DISEASED). For simplicity, anatomical and physiological characteristics of all hepatic lobules within a specific liver, normal and diseased, were presumed to be somewhat similar [Hunt et al, 2006; Yan et al., 2008a]. Pooled and averaged results from executions of multiple Monte Carlo variants of a single, parameterized LOBULE represents a single wet-lab outflow profile; those results comprise one ISL experiment.

Figure 2 shows that LOBULES are abstractions designed for their context: the degree of accuracy in concrete detail is explicitly chosen according to the purposes of the intended experiments. They are not intended to contain any accurate, though irrelevant detail. ISL components can be modified and plugged together in different ways as needed to represent different lobular properties. ISL experiments help reduce uncertainties about hepatic mechanisms by enabling the formulation and testing of fine-grained mechanistic hypotheses about plausible details that may be occurring during drug disposition. Software objects represent spatiotemporal aspects of hepatic organization and function. Events occurring during execution are measured and studied simultaneously, analogous to how wet-lab experiments are conducted. An ISL—NORMAL or DISEASED—achieves a degree of validation when the similarity between its outflow profile and a referent perfusion profile is judged adequate based on a quantitative comparison referred to as a Similarity Measure. We have learned to adhere to a strong parsimony guideline [Hunt et al., 2009]. Neither an analogue nor one of its mechanisms should be significantly more complicated than is needed to achieve the targeted Similarity Measure.

When ISL components, and their arrangements, along with behaviors generated during execution are judged acceptable, the detailed micromechanisms causing traceable cause-effect relationships

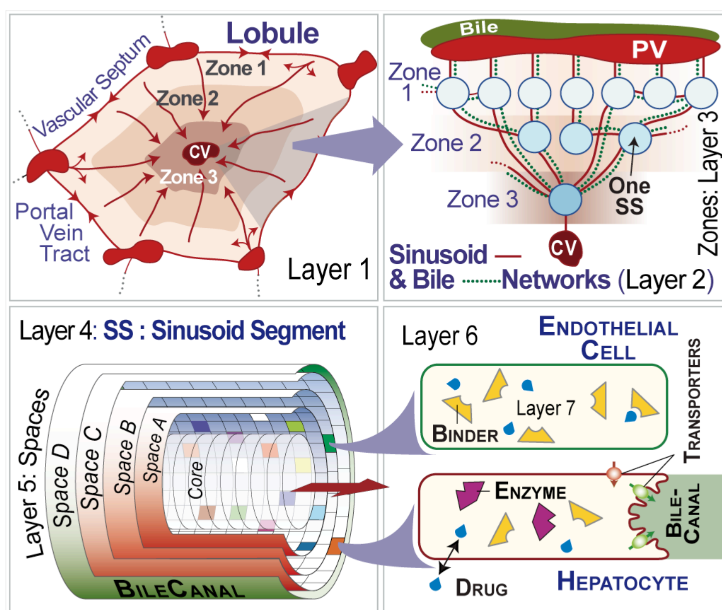

Fig. 2. Multilayer structure of an In Silico Liver (ISL). A LOBULE maps to the primary functional unit of the liver. It consists of a PORTAL VEIN, a CENTRAL VEIN, and interconnected sinusoids. Flow is PORTAL VEIN to CENTRAL VEIN. There are three zones. A LOBULE network is specified using a directed graph where nodes are organized into three zones. Intra-zonal connections are possible. Three types of interzone connections are used: Zone 1 → Zone 2, Zone 1 → Zone 3, and Zone 2 → Zone 3. A Sinusoidal Segment (SS) maps to a unit sinusoid structure and function. One SS is placed at each graph node. It contains a Core and three two-dimensional toroidal grid spaces. Space A maps to the interface between blood flow and cells. Space B contains ENDOTHELIAL CELLS and EXTRACELLULAR spaces; ENDOTHELIAL CELLS contain BINDERS. Space C contains HEPATOCYTES and EXTRACELLULAR spaces; HEPATOCYTES contain BINDERS that also function as XENOBIOTIC METABOLIZING ENZYMES. Thousands of objects representing XENOBIOTICS are dosed into the PORTAL VEIN and then move within and between spaces, as well as into and out of CELLS consistent with physiochemical properties. Space D maps to bile flow.

may correspond to the hepatic micromechanisms. At that stage, the traced dynamics of administered COMPOUND within and between an ISL provide heretofore-unavailable insight into plausible drug disposition details.

A LOBULE is a network of sinusoidal segments. The relative arrangement of hepatic function and blood flow is represented at the LOBULE layer using a directed graph called a SINUSOID network, which is a means of achieving an adequate variety of PORTAL VEIN to CENTRAL VEIN flow paths. Each Monte Carlo variant maps to a distinct arrangement of flow paths from PORTAL VEIN to CENTRAL VEIN within a portion of an acinus. A SINUSOID network is subdivided into three zones; it is easy to increase the number when that is required. Zonation enables mimicking quantitative and functional differences between periportal and perivenous lobular regions. A graph edge specifies a flow connection from a Sinusoidal Segment (SS) exit to a downstream SS entrance. Assigning edges pseudo-randomly with a particular distribution at the start of each trial mimics lobular variability within and between livers.

A SS is a software agent that represents all needed aspects of sinusoid function that can influence drug disposition. One SS is assigned to each graph node. Each is somewhat different and the stochastic differences are parameter controlled. SSs per zone are Zone 1 > Zone 2 > Zone 3. Typically, 70 or so are needed to have sufficient PORTAL VEIN to CENTRAL VEIN path variety to reduce fluctuations within outflow profiles. The PORTAL VEIN is connected to all Zone 1 SS and all Zone 3 SS are connected to the CENTRAL VEIN.

An SS consists of a Core and three identically sized, layered toroidal spaces as illustrated in Fig. 2. Spaces A–C within a SS are identical, but SS sizes are Monte Carlo specified. The spaces are subdivided into a specified number of square grid spaces. Parameters allow the resolution of the spaces to be changed. The Core maps to blood flow. It provides a direct PORTAL VEIN to CENTRAL VEIN path through which COMPOUNDS (mobile objects) can traverse. Space D maps to bile flow. Space A maps to the interface between vascular flow and the endothelial layer. Space B maps to easily accessible spaces and cells presumed to be primarily endothelial cells. Space C maps to less accessible spaces and cells, primarily the space of Disse, hepatocytes, and bile canaliculi. CELLS in Space B are called ENDOTHELIAL CELLS; those in Space C are called HEPATOCYTES.

CELLS contain whatever objects are needed to represent *required* intracellular processes, such as drug binding, metabolism, transport, and sequestration. Because in situ perfusions typically have short durations, we specified that cell biology and biochemistry were relatively constant. Consequently, details not needed are abstracted away, but can be added easily when needed [Hunt et al., 2006; Yan et al., 2008a; Park et al., 2009]. Some xenobiotics such as diltiazem are sequestered in organelles such as lysosomes and mitochondria, as well as being bound to intracellular material. However, motivated by parsimony, sequestration and binding are not resolved: everything within a cell that can bind or sequester a xenobiotic is conflated and represented by some number of identical binding objects (hereafter, simply BINDERS). Within hepatocytes, we do not currently resolve binding to metabolic enzymes, such as the CYP450 isozymes, and binding to or sequestration by other cell components. BINDERS called ENZYMES handle BINDING inside HEPATOCYTES. They use a parameter to determine which of their BINDING events ends with release of METABOLITE. When needed, several different objects that produce the same net event sequence can replace BINDERS or ENZYMES.

Testing and Exploring Mechanistic Hypotheses

It is important to understand that an ISL simulation has fundamentally different objectives from execution of a classical, inductive, mathematical data model. An ISL simulation is exactly analogous to conducting a wet-lab experiment. It is very particular and can contain specific details to enable mimicking precise phenomena. It typically has one of two objectives. 1) Test a hypothesis: execution of this ISL, configured and parameterized in this particular way, when dosed in a specified way with this particular COMPOUND, will produce phenomena, which when measured are quantitatively similar to what will be observed or has been observed in analogous wet-lab experiments. In the latter case, execution provides data needed to test the hypothesis: the targeted Similarity Measure is or is not achieved. When it is, a degree of validation has been achieved. When the targeted Similarity Measure is not achieved, the mechanism is falsified; new knowledge is achieved. We then look inside during execution to see how and where failure occurred, and with that new knowledge, we return to the Iterative Refinement Protocol, described in Fig. 3, and select and improve mechanisms until they survive.

2) Alternatively, an experiment may be designed to explore the mechanistic and phenomenal consequences of a simulated hepatic intervention such as administration of a new compound or disruption of a particular event within a cause-effect network or a change in lobular architecture. In such cases, simulation results can reasonably anticipate (predict) the cascade of causal events that generate the disposition or pharmacological end points in the referent wet-lab model. These two experimental objectives complement and leverage current inductive modeling and simulation methods [Hunt et al., 2009]. For the second objective, there will be an advantage to having “intelligent”, active compounds because that would open the door to scaling the analogue to characterize families of compounds as well as particular molecules in a disease, response specific, or target specific fashion.

Building and Iteratively Revising Analogues

In silico experiments intended for mechanism discovery and refinement must follow protocols, analogous to wet-lab experiments. When discovering or testing a mechanistic hypothesis, we follow the Iterative Refinement Protocol (IR Protocol) in Fig. 3: cycles of analogue and component synthesis, testing and evaluation, validation or falsification, assessment, cogitation, and system revision. The process continues until a prespecified Similarity Measure is satisfied. Similarity Measures typically begin weak and then are strengthened as done by Lam and Hunt [Lam and Hunt, 2010]. We have used the protocol successfully for different types of analogues [Yan et al., 2008b; Engelberg et al., 2008; Kim et al., 2010; Tang et al., 2010; Lam and Hunt, 2010]. For ISLs, step 1 is to assemble a set of hepatic and dispositional attributes, static and dynamic, which the ISL and its components will need to closely mimic in order to begin being scientifically useful. The current granularity is that illustrated in Fig. 2.

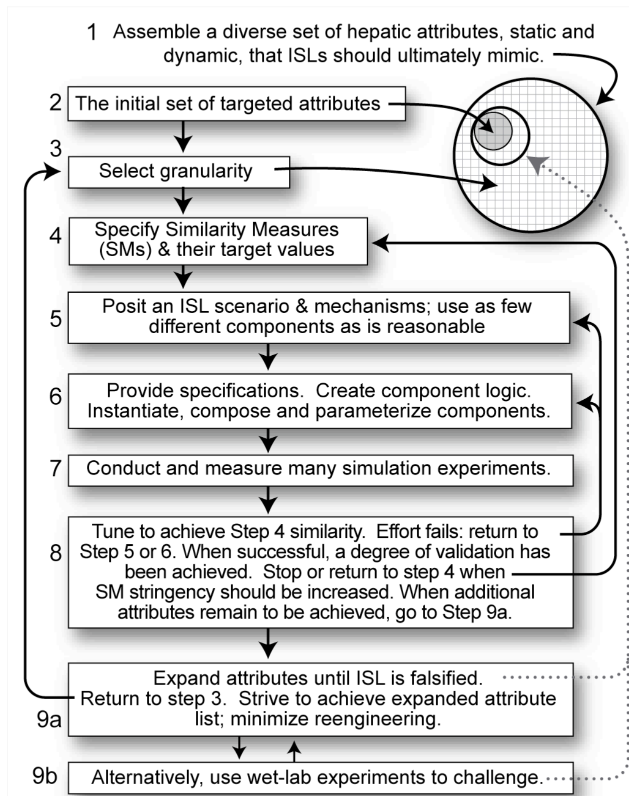

Fig. 3. The Iterative Refinement Protocol. The objective is to discover spatiotemporal mechanisms that will make components and phenomena measured during ISL experiments or experiments using other analogues increasingly biomimetic by iteratively validating and then challenging it until it is falsified.

During each IR Protocol cycle, we pursue a minimum satisfactory condition to achieve that cycle's goal. ISL mechanisms, for example, will be sufficiently complicated to sequentially achieve the targeted attributes, but no more so. Because an ISL is an extensible, modular device, we know that we can add additional detail when needed, and doing so will be relatively straightforward. The parsimony guideline encourages resisting making an ISL any more complicated than needed to achieve the current, targeted Similarity Measures, while leaving unspecified the many other mechanisms that might generate the targeted phenomena. We keep ISL components simple by conflating fine-grained biological details that are not currently needed and representing them collectively using abstract objects and/or agents having relatively simple operating principles. For example, in [Park et al., 2010] everything within a CELL that could bind a xenobiotic was conflated and represented by ENZYMES and BINDERS (Fig. 2). For the attributes targeted in that study, no more detail was needed. When additional detail is needed for other experimental contexts, old and new mechanisms coexist in order to provide multi-aspect insight into the map between generators and phenomena; we call this parallax modeling.

Drilling Down: Linking Coarse-grained Systemic Phenomena to Fine-grained Molecular Details, Including Molecular Targets

The components in Fig. 2 are more abstract and coarse-grained than a signaling pathway or gene network. They need to be only as complicated and specific as required to achieve the attributes targeted (Fig. 3). They can be made more fine-grained and specific iteratively, as the set of attributes targeted expands, as done by Lam and Hunt [2010]. Once a degree of validation has been achieved, the behaviors of current components during simulation can be used for cross-model validation during refinement to somewhat more fine-grained (greater mechanistic detail) counterparts.

ISL micromechanisms are specific events such as illustrated in Fig. 4. They map to a conflation of all fine-grained processes in lobules that influence the disposition of the drugs studied to date. When we do not have specific evidence on how some specific hepatic pathway is contributing, then there is no scientific value in simply implementing some representation simply for the sake of including it. However, to achieve the VS, it is essential to be able to achieve such linkage. So, how can that be done? Seek experimental data that falsifies the current ISL *and* demonstrates a role being played by the pathway. For example, experiments are conducted on normal and genetically altered mice in which a component of the pathway is missing (has been knocked out). A result from the latter, but not the former, falsifies the current ISL. That evidence forces us to posit one or more new, more fine-grained micromechanisms that incorporate one or more features containing the knocked-out component. Such an approach was used by Tang and Hunt [2010] to falsify a coarse-grained micromechanism and replace it with one that was more fine-grained in which components mapped directly to individual macromolecules.

A related, critical question is, how can one parse specific micromechanisms from the overall, typically coarse-grained systemic behaviors observed experimentally? An effective strategy made possible by the eight analogue capabilities in Table 1 is to expand the *variety* of attributes targeted. Addition of a new attribute to the targeted list can falsify a micromechanism. To revalidate, it may be necessary to replace a coarse-grained micromechanism with one that is somewhat more fine-grained, or to replace a rule with interacting components. Following that, a clearer mapping may exist between the more fine-grained (less abstract) micromechanism and the implicated pathway. Such a process was used by Lam and Hunt [2010] to move iteratively from coarser to more fine-grained micromechanistic hypotheses.

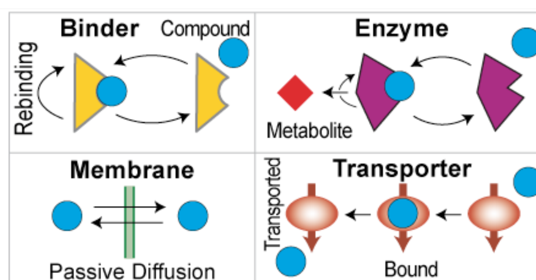

Fig. 4. HEPATOCYTE micromechanisms. These four micromechanisms are needed to mimic in vitro clearance. An arrow designates a probabilistic event that can occur within a simulation cycle, depending on object state. BINDERS, ENZYMES, MEMBRANES, and TRANSPORTERS are active objects; COMPOUNDS (BLUE) and METABOLITES (red) are passive objects. When needed, there can be two or more instances of each micromechanism.

Remove, Validate, and Return of Components to Advance ISL Capabilities

An important benefit, derived from the eight capabilities listed in Table 1, is that components, such as CELLS, can be isolated, experimented upon (and improved), and studied. In so doing, we are optimally positioned to draw on in vitro data to iteratively expand the mechanistic repertoire of, for example, HEPATOCYTES. The simulated in vitro cultures used earlier [Sheikh-Bahaei et al., 2006; Sheikh-Bahaei and Hunt, 2006; Sheikh-Bahaei and Hunt, 2007; Lam and Hunt, 2010; Sheikh-Bahaei et al., 2010] are being adapted to enable HEPATOCYTES to be removed from a validated ISL, placed in a CULTURE, and validated against fine-grained in vitro data. They can then be returned to the ISL. So doing enables the modeler to make increasing use of future in vitro data to improve both HEPATOCYTES and ISLs. The approach can be extended to other cell types used for in vitro screening and will generalize to drugs and biologics. As HEPATOCYTES and ISLs improve, these adjustments will provide methods to translate *mechanisms* (rather than

data) between species, as discussed by Park et al. [2010] for morphing normal to diseased livers. The same approach can be used when validating separately against data from humans or human cells.

To Adapt and Respond to New Drugs and Biologics, Analogue Components Need to be "Intelligent"

The expectation is that analogue improvement will need to rely increasingly on information from in vitro assays and in vivo biomarkers. Consider a new compound for which such information has become available. A likely question for the ISL framework could be this: when dosed in rodents or humans, is it likely to be hepatotoxic or to produce toxic metabolites, and if so, at what dose or exposure levels? The in vitro data will be used to parameterize those ISL components and events associated directly with toxicity pathways particular to the in vitro system. To address that question, however, it is essential that we have validated methods to predict plausible, compound-specific parameterizations for the many remaining ISL micromechanistic events for which no comparable counterpart exist in vitro. Among those for small molecules will likely be the events in Fig. 4 that contribute to hepatic clearance and metabolism.

An important, novel, and potentially transformative feature of this class of analogues is that they have been designed specifically to make solving this problem straightforward. Consider this future scenario: an ISL progeny has achieved degrees of validation for *m* xenobiotics. For each micromechanistic event (the arrows in Fig. 4 are examples), there are *m* different COMPOUND-specific parameterizations. The solution to the above problem is to make each ISL component sufficiently "intelligent" to enable it to predict its own event-specific parameterization upon encountering a new compound. It can do that using a predictive mapping from the space of its parameter values to the event-specific space of physiochemical properties for the *m* already studied xenobiotics. The expectation is that the process of discovering and validating such mappings for a given event will be much less complicated than current ADME prediction methods. The reason is that each micromechanistic event is very simple. It is their collective influence, on clearance for example, at the liver level that is complex. We hypothesize that only a few physiochemical properties will be most influential in each event illustrated in Fig. 4 and so will account for most of the variance of each micromechanistic event parameter. This is a critical hypothesis to test. The recent, yet still early stage reports of work aimed at developing this capability have been encouraging [Sheikh-Bahaei and Hunt, 2006; Sheikh-Bahaei and Hunt, 2007; Yan et al., 2008a].

Summary

The business of discovering and developing new therapeutic health and disease interventions requires acquiring and exploiting deep insight into the complex, generative, mechanistic networks that are responsible for the phenomena of interest. Doing so in the face of incomplete data and multiple uncertainties requires building and challenging better mechanistic models. The VS and methods outlined above promise to facilitate the accrual of validated and falsified mechanisms, help form concrete hypotheses for translation, raise awareness of and focus on concrete detail for particular interventions, formalize analogue and component sharing, facilitate open method and data standards in biological modeling and simulation, and increase the ease with which accumulated mechanistic knowledge is measured and visualized. These features could drastically improve our understanding of the biological systems under study and make drug development much more efficient and effective.

Acknowledgements

We thank Jesse Engelberg, Teddy Lam, Sean Kim, Andy Gewitz, and Laura Gaetano for helpful commentary on the manuscript.

Tables

Table 1. Analogues within the visionary scenario, including ISLs, will exhibit these capabilities. As engineered devices, they must also be very concrete yet robust to experimental context.

1. *Transparency*: One can visualize and measure simulation details as they unfold, and comparable them to referent wet-lab and animal models.
2. *Experimental indistinguishability*: When compared during experimentation at a similar level of abstraction, attribute measurements of an analogue and the corresponding wet-lab model are, to domain experts, indistinguishable.
3. *Mappings*: Design in silico observables and components to be consistent with in vitro and in vivo counterparts. Doing so enables clear, iteratively concretizable, in silico-to-in vitro mechanism mappings.
4. *Local mechanisms*: Simulated physiological phenomena generated during analogue execution are consequences of local component interactions (components can be heterogeneous).
5. *Reusability*: By using relational grounding (avoiding absolute grounding) [Hunt et al., 2009], analogue components can be quasi-autonomous, and reconfigured easily to represent different mechanistic hypotheses, aspects of the referent wet-lab system, and experimental conditions.
6. *Flexibility*: 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. Doing so enables cycles of scientific modeling and simulation: testing, falsification, and refinement, as the space of targeted attributes expands.
7. *Adaptability*: An analogue and/or its components must be constructed so that it can be adapted easily to function as a component in other analogues. E.g., a liver comprised of LOBULES can fit and function within human analogue; components from LOBULES and HEPATOCYTES can be used in a lung analogue, etc.
8. *Components articulate*: To study alternate mechanistic and micromechanistic hypotheses, and 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.

References

Engelberg JA, Ropella GE, Hunt CA. 2008. Essential operating principles for tumor spheroid growth. BMC Syst Biol. Dec 23;2:110.

Hung DY, Chang P, Cheung K, McWhinney B, Masci PP, Weiss M, Roberts MS. 2002. Cationic drug pharmacokinetics in diseased livers determined by fibrosis index, hepatic protein content, microsomal activity, and nature of drug. J Pharmacol Exp Ther. Jun;301(3):1079-87.

Hunt CA, Ropella GE, Lam TN, Tang J, Kim SH, Engelberg JA, Sheikh-Bahaei S. 2009. At the biological modeling and simulation frontier. Pharm Res. Nov;26(11):2369-400.

Hunt CA, Ropella GE, Yan L, Hung DY, Roberts MS. 2006. Physiologically based synthetic models of hepatic disposition. J Pharmacokinet Pharmacodyn. Dec;33(6):737-72.

Hunt CA, Baranzini S, Matthay MA, Park S. A Framework and Mechanistically Focused, In Silico Method for Enabling Rational Translational Research. 2008 AMIA Summit on Translational Bioinformatics, San Francisco, CA, March 10-12, 2008.

Kim SH, Matthay MA, Mostov K, Hunt CA. 2010. Simulation of lung alveolar epithelial wound healing in vitro. J R Soc Interface. Aug 6;7(49):1157-70.

Kim SH, Debnath J, Mostov K, Park S, Hunt CA. 2009. A computational approach to resolve cell level contributions to early glandular epithelial cancer progression. BMC Syst Biol. Dec 31;3:122.

Kim SH, Park S, Mostov K, Debnath J, Hunt CA. 2009. Computational investigation of epithelial cell dynamic phenotype in vitro. Theor Biol Med Model. May 28;6:8.

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

Lam TN, Hunt CA. 2010. Mechanistic insight from in silico pharmacokinetic experiments: roles of P-glycoprotein, Cyp3A4 enzymes, and microenvironments. J Pharmacol Exp Ther. Feb;332(2):398-412.

Luke S, Cioffi-Revilla C, Panait L, Sullivan K, Balan G. 2005. MASON: a multiagent simulation environment. Simulation. 81:517-27. http://cs.gmu.edu/~eclab/projects/mason/ accessed 28 June 2010.

Minar N, Burkhart R, Langton C, and Askenazi M. 1996. The Swarm simulation system: a toolkit for building multi-agent simulations. Working Paper 96-06-042, Santa Fe Institute, Santa Fe, NM. http://swarm.org/index.php/Main_Page/ accessed 28 June

North M, Collier NT, and Vos JR. 2006. Experiences creating three implementations of the Repast agent modeling toolkit. ACM TOMACS 16:1-25. http://repast.sourceforge.net/ accessed 28 June 2010.

Park S, Ropella GE, Kim SH, Roberts MS, Hunt CA. 2009. Computational strategies unravel and trace how liver disease changes hepatic drug disposition. J Pharmacol Exp Ther. Jan;328(1):294-305.

Park S, Kim SH, Ropella GE, Roberts MS, Hunt CA. 2010. Tracing multiscale mechanisms of drug disposition in normal and diseased livers. J Pharmacol Exp Ther. Jul;334(1):124-36.

Sheikh-Bahaei S, Maher JJ, Hunt CA. 2010. Computational experiments reveal plausible mechanisms for changing patterns of hepatic zonation of xenobiotic clearance and hepatotoxicity. J Theor Biol. Aug 21;265(4):718-33.

Sheikh-Bahaei S, Hunt CA. 2006. Prediction of In Vitro Hepatic Biliary Excretion Using Stochastic Agent-Based Modeling and Fuzzy Clustering. In: L. F. Perrone, et al., eds., Proceedings of the 37th conference on Winter Simulation, Monterey, CA, Dec 03 - 06, 1617-24.

Sheikh-Bahaei S, Ropella GE, Hunt CA. 2006. In Silico Hepatocyte: Agent-based Modeling of the Biliary Excretion of Drugs In Vitro. In: L. Yilmaz, et al., eds., Proc 2006 Agent-Direc Simul Symp (Huntsville, AL, Apr 3-9), SCS Press, San Diego, CA, 157-163.

Sheikh-Bahaei S, Hunt CA. 2007. Parameter estimation via analysis of fuzzy clusters (PEAF): an algorithm to estimate parameters of agent-based models. In G. Hu, ed., Proc ISCA 20th Intern Conf Comp Appl Ind Engr, ISCA, Cary, NC (ISBN: 978-1-880843-65-9), San Francisco, CA, November 7-9, 23-27.

Sheikh-Bahaei S, Ali KA, Bera SR, Hunt CA. 2010. Predicting human hepatic clearance using hypernet neural networks. 19th International Conference on Software Engineering and Data Engineering (SEDE-2010), San Francisco, CA, June 16-18, ISBN: 978-1-880843-77-2, 355-359.

Tang J, Hunt CA. 2010. Identifying the rules of engagement enabling leukocyte rolling, activation, and adhesion. PLoS Comput Biol. Feb 19;6(2):e1000681.

Yan L, Sheihk-Bahaei S, Park S, Ropella GE, Hunt CA. 2008. Predictions of hepatic disposition properties using a mechanistically realistic, physiologically based model. Drug Metab Dispos. Apr;36(4):759-68.

Yan L, Ropella GE, Park S, Roberts MS, Hunt CA. 2008. Modeling and simulation of hepatic drug disposition using a physiologically based, multi-agent in silico liver. Pharm Res. May;25(5):1023-36.