Multiscale Computational Models of Complex Biological Systems

Joseph Walpole, Jason A. Papin, and Shayn M. Peirce

Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia 22908; email: smp6p@virginia.edu

Annu. Rev. Biomed. Eng. 2013. 15:137–54

First published online as a Review in Advance on April 29, 2013

The *Annual Review of Biomedical Engineering* is online at bioeng.annualreviews.org

This article's doi: 10.1146/annurev-bioeng-071811-150104

Keywords

data integration, model validation, systems biology, bioinformatics, biochemical networks, model design

Abstract

Integration of data across spatial, temporal, and functional scales is a primary focus of biomedical engineering efforts. The advent of powerful computing platforms, coupled with quantitative data from high-throughput experimental methodologies, has allowed multiscale modeling to expand as a means to more comprehensively investigate biological phenomena in experimentally relevant ways. This review aims to highlight recently published multiscale models of biological systems, using their successes to propose the best practices for future model development. We demonstrate that coupling continuous and discrete systems best captures biological information across spatial scales by selecting modeling techniques that are suited to the task. Further, we suggest how to leverage these multiscale models to gain insight into biological systems using quantitative biomedical engineering methods to analyze data in nonintuitive ways. These topics are discussed with a focus on the future of the field, current challenges encountered, and opportunities yet to be realized.

Contents

INTRODUCTION

Biological systems are inherently complex in nature; they are composed of multiple functional networks that operate across diverse temporal and spatial domains to sustain an organism's growth, development, and reproductive potential. These so-called multiscale systems extend from the most basic of amino acid substitutions that alter protein function to concerted multicellular signaling cascades that regulate hormone release throughout an entire lifetime. Computational models are uniquely positioned to capture the connectivity between these divergent scales of biological function, as they can bridge the gap in understanding between isolated in vitro experiments and whole-organism in vivo models.

Although its meaning may seem clear, multiscale should be defined carefully, as it can very quickly spiral into the realm of catch-all scientific jargon. Fundamentally, a multiscale model must explicitly account for more than one level of resolution across measurable domains of time, space, and/or function. To clarify, many models of physical systems implicitly account for multiple spatial scales by simplifying their boundary conditions into black boxes in which assumptions about other spatial or temporal domains are summarized by governing equations. Additionally, a multiscale model's explicitly modeled tiers of resolution must provide information beyond that which can be obtained by independently exploring single scales in isolation.

The classic engineering exercise of heat transfer through an insulated rod is an excellent case study in implicit multiscale modeling. Whether it is solved using continuous partial differential equations (PDEs) or a discrete finite element method (FEM), all solutions to this problem rely on carefully defining spatial boundary conditions, the fundamental laws of thermodynamics of a closed system, and material properties such as a thermal conductivity coefficient. Using these tools, engineering students unwittingly wrangle molecular motion at the femtometer scale to reliably predict the distribution of temperatures across an idealized one-dimensional landscape measured in meters. However, were the motion of each molecule of metal in the rod to be explicitly accounted for, would any additional information about the system be gained (assuming that this were not a computationally intractable challenge)? In this example, the governing equations of thermodynamics sufficiently capture the probabilistic distributions of molecules without requiring explicit representation in the model.

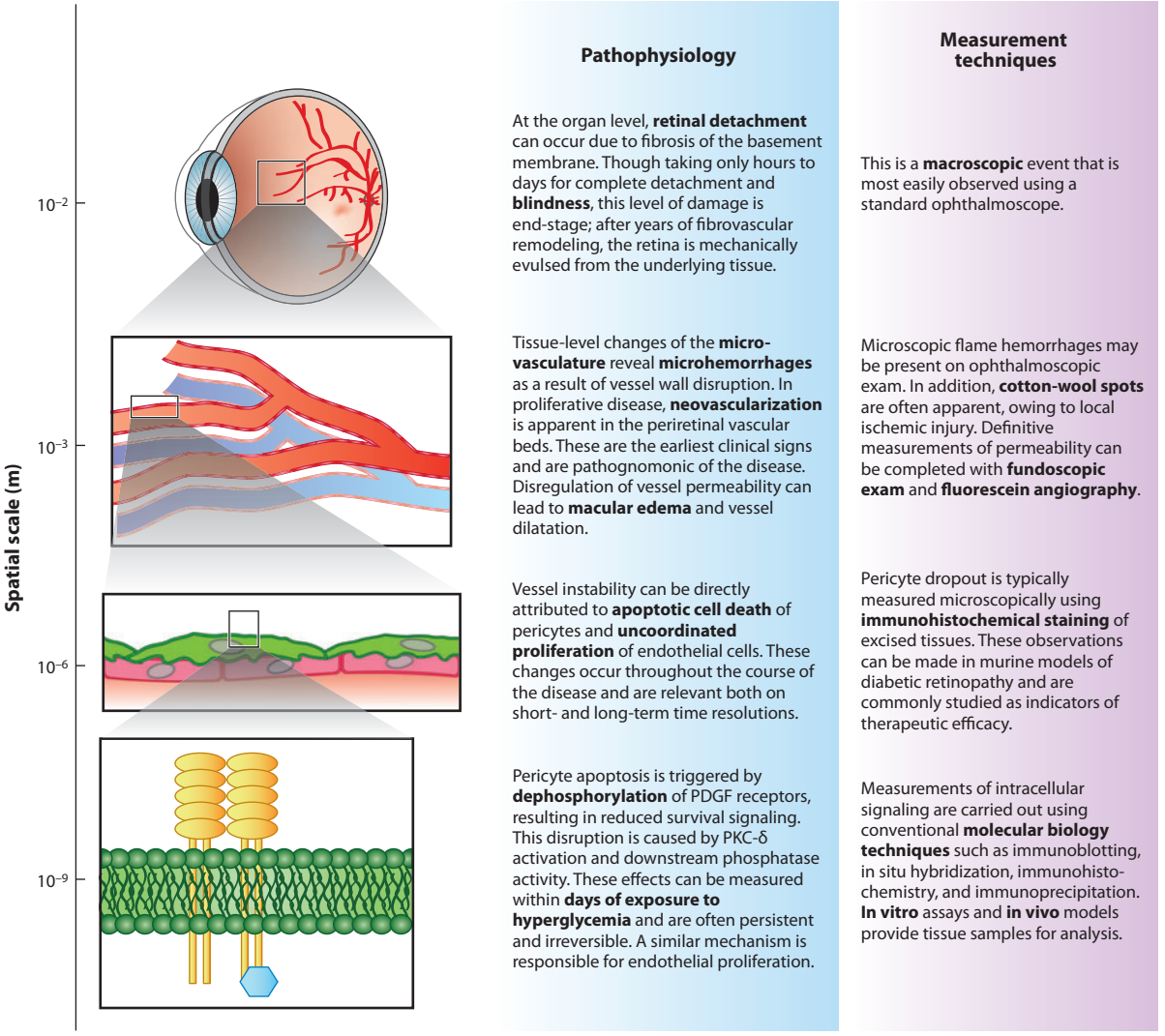

Figure 1

Diabetic retinopathy as a case study in multiscale pathophysiology. A detailed look at how the pathogenesis of diabetic retinopathy is a function of multiple spatial scales across biology. Abbreviation: PDGF, platelet-derived growth factor.

Ultimately, this model system is explicitly analyzed at the scale of the rod while implicitly accessing information about molecular thermal motion using established equations of thermodynamics. However, as a law of biological systems has yet to be codified into governing equations, the biomedical scientist lacks the means to similarly make assumptions across multiple tiers of measurable resolution in any accurate way. This challenge is further compounded by the complex nature of the system that is being investigated; that is to say, components of biological systems act differently in isolation than they do when integrated into the larger machinery of a living organism.

To further illustrate the need for explicit multiscale models in biology, consider the multiple levels of spatial, temporal, and functional scale that are known to operate in the pathophysiology of diabetic retinopathy (**Figure 1**). At its most advanced stage, proliferative diabetic retinopathy can

result in blindness due to retinal detachment at the macroscopic level. The detachment, however, is preceded by years of tissue damage caused by microvascular hemorrhage and fibrovascular remodeling of the retinal basement membrane. These defects in the vessel wall are the result of pericyte (abluminal vascular support cell) apoptosis, which leads to aberrant vessel growth and increased vessel permeability throughout the retina. Finally, pericyte apoptosis occurs owing to reduction of platelet-derived growth factor (PDGF) receptor survival signaling mediated by activation of PKC-δ and downstream phosphatases in the setting of chronic hyperglycemia (1, 2).

Which tier of resolution provides the most information for understanding the underlying mechanisms of this complex disease? Conversely, is there a tier of resolution that offers the least understanding of the disease? The debatable answers to these questions have driven model building for decades, as investigators attempt to develop the highest information yield from their intellectual investments in computational modeling approaches. More recently, however, investigators have been turning to multiscale modeling techniques to generate detailed information about complex biological systems. In these multiscale models, perturbations of fine-grained parameters (e.g., protein modifications) can generate observable and measurable changes to coarse-grained outputs (e.g., tissue patterning), and vice versa. This integration across functional, spatial, and temporal scales in biological systems introduces a powerful tool for capturing and analyzing biological information that is inaccessible through other modeling and experimental techniques.

In this review, we provide a meta-analysis of multiscale modeling, focusing foremost on recent publications from the biomedical engineering community. First, we describe the tiers of biological resolution that have been modeled and the computational techniques leveraged to obtain insightful conclusions. Our focus then shifts to a discussion of best practices in model verification and validation as we discuss challenges unique to multiscale modeling. Once we have covered the questions and the tools used to answer them, we expand on how multiscale models capture biologically relevant data that may be inaccessible using conventional wet-lab techniques. Finally, we look to the future of the field and propose a set of specific landmarks that, if accomplished, may provide even greater insight into the form and function of complex biological systems.

CURRENT MULTISCALE MODELING EFFORTS

Although computational models take many shapes and forms, we propose a simple taxonomy for defining characteristic styles that serves to define multiscale models for illustrative purposes (**Figure 2**). These model types are neither absolute nor comprehensive; rather, the taxonomy provides a simple reference that allows for criteria-based discussion of the examples used in this review. Type 1 models are iterative approaches in which data from a single scale of simulation are used as input for the next tier of resolution. Although outputs are available at each tier, they do not necessarily inform events at previously simulated tiers of resolution. Type 2 approaches rely on independently simulating each scale of resolution to generate outputs for other simulations. In this way, information is passed between tiers of resolution in discrete packets of input/output data. Lastly, Type 3 approaches use simulations run in parallel with constant communication between tiers of resolution. This approach can also be viewed as a way for a low-resolution simulation to control and receive information from many simulations at higher resolutions.

Scales of Biology: What Is Being Modeled?

Multiscale models are pervasive in the biological sciences, covering many tiers of resolution and many disciplines. Using a selection of literature from the past decade, we have highlighted and clustered broad biomedical disciplines based on the levels of spatial resolution they are investigating

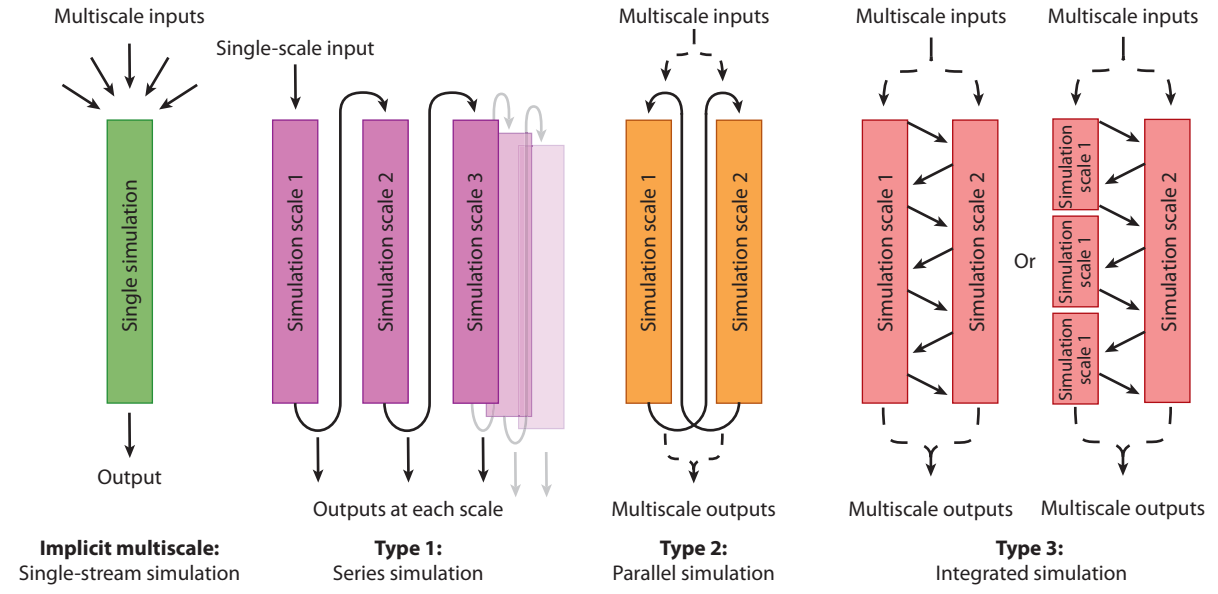

Figure 2

Multiscale model taxonomy. A proposed system of nomenclature for categorizing the techniques used to generate multiscale models for the purposes of this review.

with multiscale models (**Figure 3**). A clear trend exists, in which metabolomics and genomics research are clustered separately because they are uniquely focused on subcellular resolutions (3–7). Further, most work at the organ and multisystem scales consists of studies of tissue mechanics and disciplines interested in cellular trafficking (e.g., cancer and immunology) (8–18).

As might be expected, much effort is focused on the interrogation of biology at many resolutions, from signaling networks (i.e., subcellular simulations in which proteins are not explicitly modeled) to cell networks (i.e., tissue-level simulations comprising more than a single cell). In particular, the fields of cell biology, developmental biology, vascular biology, and cardiovascular research all share very similar methods of modeling at these tiers (19–42). A common theme among these fields is a desire to understand how subcellular networks may influence tissue-level patterning through the actions of individual cells.

Of course, it should be emphasized that these are trends from a subset of papers that have been broadly classified based on the field of biological research and the explicitly modeled tiers of resolution. This meta-analysis is also purely an evaluation of the quantity of publications in a given field and not the quality of the models being developed. Clearly, other disciplines are also using multiscale modeling to their advantage, and even in the disciplines shown, there are researchers whose work does not neatly conform to the selected scales. This meta-analysis does, however, demonstrate a clear trend in the literature, which may provide some insight into current gaps in computational coverage within the disciplines of interest.

Most importantly, this analysis demonstrates that a major goal of the field is yet to be realized: No single comprehensive gene-to-organism multiscale model has been developed. Based on our observations, within each of the listed disciplines, there are many open avenues of research where multiscale efforts are either sparsely represented or completely nonexistent. This deficit is not a shortcoming, but rather an opportunity to push the boundaries of knowledge in these biomedical

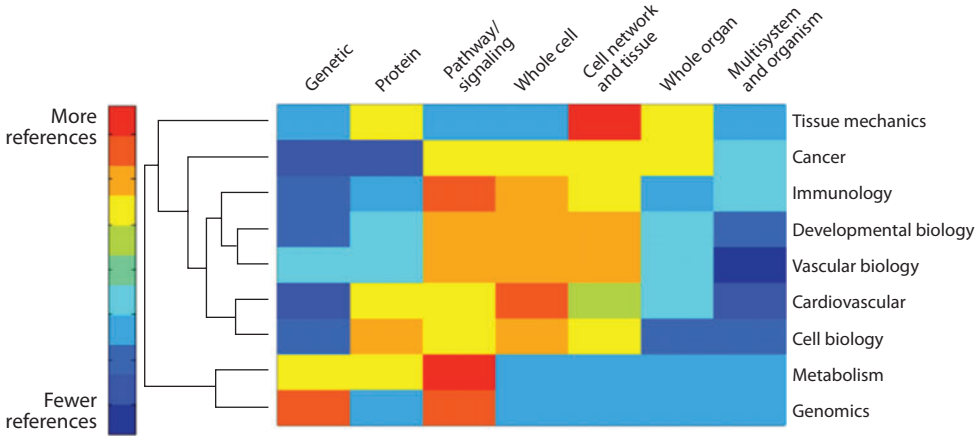

Figure 3

Clustergram of multiscale models as a function of biological discipline and spatial resolution. Each publication was scored as containing (1) or not containing (0) a biological scale within the described multiscale model. For each discipline, the Boolean values were summed and then normalized to the total number of publications within that category, such that the weighted heat map is scaled from 0 to 1. For example, of the 7 publications in vascular biology, 5 involved whole-cell components, resulting in a weighted score of $5/7 = 0.71$. A total of 39 publications were included in this analysis.

investigations, using multiscale modeling as a platform for high-throughput, high-yield hypothesis generation and testing.

Models Within Models: Components of a Multiscale System

All modeling methodologies have strengths and weaknesses with regard to their ease and fidelity in capturing biological system dynamics. Typically, these techniques are broadly classified into continuous and discrete modeling strategies based on how the solution space is acquired. Classification into deterministic and stochastic models is an alternative method that divides systems based on whether they contain a degree of randomness that allows for multiple solutions to the same initial conditions. Importantly, though not an exhaustive list, the modeling techniques presented here are all taken from published multiscale models; these examples have already been validated against experimental data and, therefore, serve as a foundation for future computational efforts.

Continuous modeling strategies often use systems of ordinary differential equations (ODEs) and PDEs to reach steady-state solutions. Solutions to these continuous systems are deterministic, as they obey the Picard-Lindelöf theorem on existence and uniqueness (43). Because numerical methods for solving PDEs, such as FEMs and finite volume methods (FVMs), rely on reduction to a system of ODEs, the theorem's assumption of uniqueness holds despite the possibility that the PDEs contain stochastic elements.

Generally, systems of ODEs using the law of mass action kinetics are leveraged to represent chemical reactions within the cytosol and nucleus of the cell (13, 15, 18, 19, 44–46). As the kinetics of molecular binding, conformational switching, and diffusion often occur over very small timescales, the assumption of steady state in the overall model architecture (which may be discretized into hours, days, weeks, etc.) is typically valid. Sun et al. (47) employed a Type 3 approach using a system of ODEs executed with the COmplex PAthways SImulator (COPASI)

to explicitly model the function of transforming growth factor (TGF)-β1 in a multiscale model of epidermal wound healing. Using this technique, the authors expanded on a previous single-scale model and were able to decouple the promigratory and antiproliferative effects of TGF-β1 on various cell types in an in silico skin wound closure model over time. Analogous reasoning and techniques have also been used for analysis of metabolic and signaling networks in which a steady-state flux is desired for informing higher tiers of function (3, 5, 6, 45, 48).

Multiscale models of reaction diffusion kinetics are also typically executed in continuous time and are often used to represent intra- and extracellular molecular binding and diffusion (29, 38, 39, 41, 49). These models differ from the diffusion/pathway models discussed above, as they typically rely on systems of PDEs that are solved using numerical approaches. Broadly speaking, FEMs (and related FVMs) are uniquely suited for monitoring geometrically constrained properties, such as cell surface interfaces, and mechanical properties of tissues across all scales (17, 38, 50–54). Aguado-Sierra et al. (35) generated a patient-specific three-dimensional model of heart failure in which a finite element mesh was fitted to echocardiographs and mechanical parameters were directly estimated from a combination of MRI and cardiac ultrasound (modified Type 2 approach). Such work highlights the clinical value of computational models by using patient data to generate electrical conduction and mechanical contractility maps with the potential to inform interventional decisions as processing cost and time decrease. It is important to note that these approaches are actually a hybrid of continuous and discrete strategies, as FEMs rely on discretization of continuous equations to generate numerical solutions for otherwise irreducible PDEs.

Discrete stochastic modeling techniques are a heterogeneous group of computational foundations that rely on nondeterministic solutions to generate constrained distributions of outputs. These techniques include methods such as Markov chains, whose probabilistic transition matrices are well suited to biological systems with functions that can be discretized into independent states. Along with the related class of discrete state–based Boolean networks, these techniques have modeled receptor activation states (e.g., cardiomyocyte ion channels), compartmentalized signaling networks, and functional protein conformations (19, 20, 34, 37, 40, 55). Barua et al. (5) recently developed an algorithm, *Gene*Force, to explore the Boolean rules in metabolic signaling networks and correct for inconsistencies between experimental results and model predictions. The model forces an optimized output by allowing for a set degree of rule violation; these perturbations to the original set of rules revealed incorrectly silenced gene transcription, which when corrected allowed for agreement with experimental results. This Type 3 approach generated as much as an 8% improvement in model predictive accuracy and was applied to well-curated metabolome libraries for organisms such as *Escherichia coli*. From a multiscale perspective, this model identifies gene-level phenomena that impact metabolomic signaling outputs.

Recently, agent-based modeling (ABM) has become a very popular and powerful tool for representing discrete stochastic biological processes as either compartmentalized or spatially defined models. These models include geometries in one-, two-, and three-dimensional configurations and may be scaled such that each fundamental agent is as large (groups of organisms) or as small (subcellular membrane components) as is desired. Zahedmanesh & Lally (52) incorporated a lattice-free ABM with a FEM approach to explore the effects of porosity, compliance, cyclic strain, and flow-induced shear stress on tissue-engineered blood vessels (Type 2 approach). This investigation was able to explore how these complex and nonintuitive parameters combined to affect development of intimal hyperplasia over time; this technique has the potential to analyze timescales that cannot be investigated using in vitro techniques, allowing for predictions of long-term performance. Owing to their diversity of scale, ABMs have been used to describe multicellular processes including tissue electrical conduction, cell trafficking, tissue mechanics, immunomodulation, arterial remodeling, inflammation, and many others (10, 14, 18, 21, 37, 38, 44, 56, 57).

Selecting a Computational Method Based on Function and Spatial Resolution

The computational techniques presented above are examples that are currently being employed in multiscale models. We classified the techniques as continuous-deterministic or discrete-stochastic (with some exceptions and hybrids), while highlighting specific spatial and temporal domains that the models are suited to represent. This classification forms the basis for a discussion of how multiscale models can be designed by selecting the best computational techniques for the task, rather than forcing a modeling technique to approximate a system for which it is poorly suited. To this end, we propose guidelines for how these individual techniques can be combined across scales (**Figure 4**).

As a class of modeling techniques, network analyses include discrete state–based techniques (e.g., Markov chains, Boolean networks) as well as continuous systems biology approaches (e.g., flux balance analysis). These methods are well suited to modeling the smallest tiers of resolution: genomic, proteomic, and metabolomic. It is important to note that these are different scales of network connectivity and represent their own tiers of resolution. Genomic data are spatially compartmentalized to the nucleus and are temporally independent of proteomic concentrations and posttranslational modification. Similarly, events occuring at the metabolomic scale, though dependent on proteomic data, integrate multiple proteins and their relative concentrations and localizations. As such, these subcellular regimes are independent modules for simulation, despite being biochemically codependent.

Milne and colleagues (6) recently constructed a composite gene-protein-reaction (GPR) model to simulate regulation of butanol production as a function of growth conditions (e.g., growth medium, atmosphere, etc.). Their findings supported the hypothesis that *Clostridium beijerinckii* is an ideal candidate for biofuel applications. The *i*CM925 model contained 925 genes coding for 938 reactions involving 881 metabolites—approximately 18% of the protein-coding genome of *C. beijerinckii*. This level of detail and network annotation for a relatively understudied organism was captured and analyzed with linear algebraic equations that were defined by a homogeneous ODE constrained with mass balance principles. Simply put, a vast amount of multiscale data (genetic expression and metabolomic network connectivity) was integrated using a single computational technique through a Type 1 approach.

In the subcellular regime, continuous-deterministic systems of ODEs and PDEs are also ideal for monitoring concentrations of signaling molecules in both the intra- and extracellular domains. Owing to a paucity of relevant kinetic parameters, these systems are often less comprehensive than the previously described network analyses; however, they excel at explicitly accounting for binding kinetics and monitoring rates of reactions as a function of time. Sample & Shvartsman (29) demonstrated the use of these continuum approaches to solve for gradients of morphogens within the developing *Drosophila* embryo. In their model, solving for compartment-dependent degradation rates was integral to understanding nuclear-cytoplasmic shuttling of morphogens, which are responsible for long-range patterning (hybrid of Type 1 and 2 approaches). Although focusing on only a single protein, this technique expands the resolution from purely intracellular reactions to subcellular components within intercellular interactions.

At this point, we have explicitly accounted for many of the internal cellular components (i.e., genome, proteome, and signaling networks); the next tier of resolution, the whole cell, requires additional consideration, as the functions of interest are again interwoven with the scale of investigation. The cell may be viewed as a mechanical entity with discretized membrane segments and interconnected cytoskeletal components, or it may alternatively be viewed as the smallest component of the system. This biological scale is a natural transition point at which both continuum

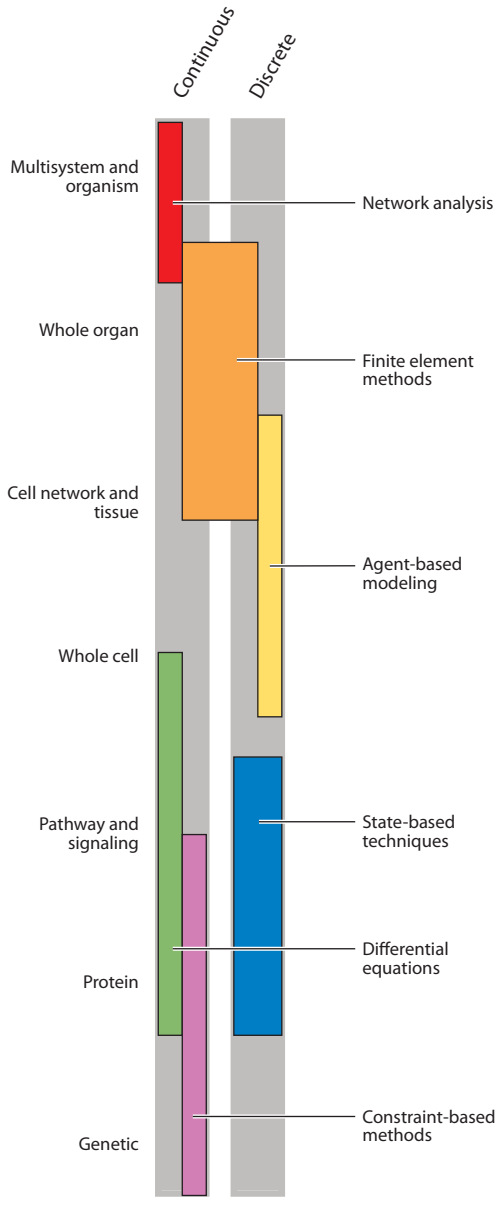

Figure 4

Map of modeling techniques by scale. Conceptual map of modeling techniques divided into continuous and discrete categories across the spatial scales for which they are most suited.

and discrete modeling approaches have been successful, and it falls to the investigator to make the final decision, guided by the hypothesis to be tested. Practically, if the cell is the largest entity in the system (i.e., only a single cell is being modeled), a more fine-grained approach is necessary. The converse is also true: If the cell is part of a larger tissue network, it must be more coarsely resolved to make observations feasible given limited computing resources.

For the sake of simplicity, here we consider the cell to be a transition state between the subcellular and supercellular domains (this notably excludes mechanical analyses of single cells, which are often performed at the whole-cell level). Such a view favors a discrete-stochastic approach to cell behavior, as this captures a degree of biological noise and allows for easy representation in physical space. ABMs are well suited to this task because they can be specifically adapted to represent cells as either single- or multiagent entities within the system. Bentley et al. (57) chose the latter approach and represented a capillary as a linear array of 10 endothelial cells, each comprising 1,288 membrane agents. This representation was necessary because their analysis required discrete membrane localization of receptors, as well as detailed filopodial sprouting within a three-dimensional extracellular space (Type 1 approach). Bailey et al. (24) opted for the former approach, representing each endothelial cell in the network as a single agent from which to generate a larger microvascular system (Type 1 approach). Again, this selection was reasonable based on the analysis at hand: leukocyte extravasation as a function of adhesion molecule expression in a tissue bed.

Tiers of resolution beyond the cell network and tissue levels, as demonstrated in **Figure 3**, remain largely unexplored as components of multiscale models. This may be due to technical limitations in that computational power is not yet available to track discretized agents throughout an entire organism. Larger, whole-organ models do exist and typically adopt a finite element approach in which each cell is represented as part of the discretized mesh. Moreno and colleagues (20) had great success with this technique, using an FVM approach to analyze the effects of antiarrhythmic pharmaceuticals on cardiac conduction through a fully rendered three-dimensional human heart (several Type 1 iterations). This model stands out, particularly because the smallest explicitly resolved element was the well-studied cardiomyocyte sodium channel. This voltage-gated channel was modeled using Markov states that were altered in the presence of various inhibitors. Most notably, cardiotoxic concentrations of antiarrhythmics could not be predicted at the single-cell scale; however, when cells operating with the same parameters were linked into a network (and ultimately a complete tissue), the model very closely matched clinically observed data.

To summarize, function and spatial resolution dictate modeling technique. Based on our current understanding and computational limitations, it is necessary to view some biological processes as continuous equations and others as discrete states. Ascending from sub- to supercellular resolutions, continuous models that were once exceptionally accurate begin to lose resolving power. Conversely, discrete models are often computationally expensive and become most useful at lower resolution for cell networks and tissues where cells are easily viewed as individual modules. Yet, larger systems may require a return to network approaches to account for spatial distances and boundaries between organ systems that are too large to be explicitly modeled at the cellular level.

VERIFICATION AND VALIDATION OF MULTISCALE MODELS

Validating Across Multiple Scales

As with all computational models, prior to their use as experimental constructs, multiscale approaches must be rigorously tested against independent data sets for proper validation. Recently, Qu and colleagues (58) reviewed how information is translated between scales of models and highlighted several of the challenges associated with validation across tiers of resolution. A key observation from this review is that inherently noisy stochastic systems and noiseless deterministic systems can generate dramatically different outputs when used to model the same biological phenomena (Keizer's paradox). Furthermore, adding noise to a previously noiseless system by combining deterministic and stochastic models may increase the likelihood of phase transitions,

thus increasing the number of stable solutions. These additional solutions may be biologically relevant; however, they may also become problematic, as the addition of stochasticity to a system may generate results that disagree with previously validated solutions in deterministic systems.

Ultimately, this incongruence between solutions reduces to the simple fact that insufficient computational resources exist to explicitly model every protein in a living organism simultaneously. Multiscale models must rely on techniques such as those mentioned above (selecting appropriately resolvable approaches based on function and spatial scale, using integrative systems biology, etc.) to capture accurate and robust information from each tier of resolution. It stands to reason that by linking potentially divergent modeling techniques, inconsistencies may be introduced into multiscale systems. To reach model agreement (both intermodel agreement and agreement with biological experiments), a validation strategy that is both theoretically sound and computationally practical must be selected.

Individual Verification Versus Complete Multiscale Verification

Multiscale models often originate through linking of individual models from two different scales to generate a composite system. In cases in which each tier of a model has been independently published, the models must, by definition, be validated at the single-scale level before validating at the multiscale level. Our lab has, in collaboration with others, followed this strategy to generate Type 2 multiscale models from successfully implemented single-scale models (21, 59, 60). In one particular example, the multiscale model captured continuum elements (extracellular matrix composition, fluid dynamics, etc.) as well as discrete elements (mechanical properties as determined by cell number and orientation) to generate a blood vessel wall for measuring adaptation to chronic hypertension.

As explored by Hayenga and colleagues (61), prior to a comprehensive model being generated, the continuum and discrete systems shared common outputs that were independently validated. Importantly, despite sharing independently validated outputs, the models were not in complete agreement, as they drew on data from different scales. The discrete ABM was generated from cell-level data acquired primarily from reduced in vitro systems that no longer maintained systems-level responses. Conversely, the continuous constrained mixture model was based on tissue-level data from studies of tissue parameters in which different systems-level responses were potentially still intact. Disagreement between the models presented a significant challenge, as neither was, strictly speaking, incorrect.

Ultimately, to reconcile these differences between scales and allow for comprehensive model validation, agreement on shared variables was required. As such, each model was deemed equally unreliable for the purposes of weighting parameters for a genetic algorithm approach to parameter estimation. Agreement between the continuous and the discrete models was achieved for shared parameters, allowing for validation of independent terms using a shared data set.

Fedosov et al. (62) described a multiscale model of erythrocyte membrane mechanics in the context of malaria infection and how changes in material properties and cell geometry impact bulk blood viscosity (Type 1 approach). In this example, validation was performed at the cellular level using optical tweezers and optical magnetic twisting cytometry to measure deformability of erythrocytes during different stages of malaria parasite development. Bulk blood viscosity was validated against a separate data set to demonstrate that each tier of model resolution independently achieved agreement with biologically relevant data sets. In order to perform these validations, previously dimensionless particle models had to be scaled using erythrocyte diameter as a reference length. This example highlights how careful selection of units and appropriate parameter selection are necessary to achieve multiscale validation.

Multiscale models are subject to scrutiny at both individual and integrated tiers of resolution. To appropriately parameterize a model and achieve validation, each module or computational technique must be in agreement with biological data before investigators can advance to a complete multiscale simulation. Further validation of the multiscale model is required to test the reliability of data transfer between computational scales such that cross talk between continuous and discrete systems does not introduce artifacts or discrepancies. As with all modeling efforts, thorough and thoughtful validation is key to achieving acceptance in the biological community; the predictive power of a model is dependent on the rigor of this validation.

BIOLOGICAL INSIGHT FROM MODELS

Measuring the Unmeasurable

Most modeling endeavors begin with a hypothesis that cannot be easily tested using even the most cutting-edge experimental assays. Tracking individual macrophages in real time in vivo, measuring chemokine concentration gradients throughout an entire tissue region, determining frequency responses to mechanical stimuli in the human ear, observing capillary and lymphatic filling as a function of muscle contraction, quantifying the effects of drug therapy on granuloma formation over the course of 300 days with receptor-level resolution—these are just a few examples of recent investigations that would not be possible without multiscale modeling approaches (24, 39, 42, 44, 51). Multiscale models are capable of quantifying any explicitly implemented variable as an output across all tiers of resolution.

In addition to quantifying individual variables with relative ease, multiscale models also allow for simultaneous observation of multiple parameters across resolution domains. Tracking multiple variables across a range of parameter values allows for construction of valuable phase planes to describe systems-level behaviors (13, 20, 29, 57). Bifurcations in these phase-plane analyses offer insights into system stability and potential interventional targets that may yield higher likelihoods of maintaining transitions from one equilibrium state of a biological system to another. For example, Kim & Maly (63) explored reorientation of individual $CD8^+$ natural killer T (NKT) lymphocytes in the two-dimensional parameter space defined by microtubule length and initial centrosome orientation relative to a target cell (Type 2 approach). This analysis revealed complex relationships between the parameters, suggesting certain combinations that would render the NKT cell unable to properly orient itself for productive cytolytic activity. Such incompatible orientations could not be predicted by either parameter alone, emphasizing the need for more rigorous analysis.

Similarly, as shown in Holland et al. (22), acquiring many data across multiple scales of resolution allows for more informed selection of reducible components within complex systems (Type 2 approach). Using a graphical approach in a normalized phase plane to study the kinetics of β-adrenergic signaling, this investigation demonstrated a method to identify reactions that can be reduced reasonably to steady state when evaluating system dynamics. Each reaction trajectory in the system was compared with steady-state values: Trajectories in the phase plane with greater deviations from steady state had larger hysteresis loops and could be identified as necessary for capturing dynamics of systems behaviors.

These examples highlight how traditional engineering approaches to capturing systems behaviors can be applied to biological systems. However, as these approaches generally require large amounts of quantitative data to be useful, traditional wet-lab experiments do not easily translate to these analysis techniques. Computational models, in particular multiscale models, offer an alternative source of data that can be acquired across many parameter values in a

high-throughput manner. As experimental methodologies develop, these predictions can be independently validated, or they may provide insight into unexplored hypotheses that can be tested immediately.

The Virtual Bench: In Silico Perturbations

Beyond simply capturing otherwise inaccessible measurements with high resolution across large changes in scale, multiscale modeling allows for precise manipulation of network variables to distinguish true effects from experimental artifacts. Even the most precise RNA interference strategies (including small interfering, short hairpin, and microRNA) are capable of producing off-target effects either directly (e.g., silencing alternative binding sites) or indirectly (e.g., diminution of native RNA translation), resulting in confounding or erroneous observations (64). This statement does not imply that models are error-free; models simply are capable of isolating perturbations as defined by the user's constraints without concern about unknown interactions.

Thus, at first glance, multiscale modeling offers the unique advantage of being able to introduce perturbations across any tier of resolution. This ability is potentially very powerful, as it allows for not only knockdown and overexpression experiments but also very precise changes to the degree of expression of a single gene or set of genes. This control is, quite simply, not possible with current molecular biology techniques. Though not a replacement for experimental investigation, these approaches can serve to contextualize data and reconcile discrepancies that may be caused by off-target effects. Further, computational methods may also refine experimental approaches by surveying all possible perturbations to narrow the scope of experimental interrogation.

For example, Fallahi-Sichani and colleagues (13, 44) have described the effect of modifying NF-κB signaling mechanisms at the level of transcript stability with implications for temporal variables (e.g., degradation rate, activation rate) affecting the outcomes of *Mycobacterium tuberculosis* infection (Type 2 approach). In this work, pharmacological therapies were applied using a system of ODEs to capture intracellular signaling pathways, and cellular behaviors were executed as a discrete probabilistic ABM at the tissue-level scale. This union of subcellular pathway manipulation and multicellular function allowed for direct investigation of pharmacological intervention on a relevant pathophysiological outcome that would otherwise be unobtainable by modeling at an individual tier of resolution.

Alternatively, for systems being modeled with a top-down approach, more general questions can be answered by completely removing subsystems from multiscale models. In these cases, functional impairments are evaluated, as opposed to specific physiological interventions. Using this approach, Shirinifard and colleagues (9) demonstrated unique growth patterns in avascular tumors by removing the capability for angiogenic growth from their multiscale model of solid tumors. Insights from such a broad phenotype perturbation (i.e., complete abrogation of angiogenesis rather than impairment of a single component in the pathway with downstream effects) allow for investigations into the minimal functions necessary for individual system behaviors.

Further, the concept of in silico perturbations can be extended by direct analogy to benchwork—with the exception that it can be executed at high throughput with low resource allocation. More so than single-resolution models, multiscale models can be mapped directly to biological assays for both experimental validation and hypothesis testing. Multiscale models have proven predictive for optimizing biopolymer scaffolds based on altering material properties to investigate extracellular matrix mechanotransduction and cell seeding (16, 52, 56). Small parameter changes that were easily completed with computational iteration in these studies would have required extensive material cost and time commitment to generate comparable data sets on the benchtop.

LOOKING FORWARD

Throughout this review, we have highlighted the currently available computational tools for multiscale modeling and the best practices for their implementation. As shown in **Figure 3**, many disciplines of biological research have yet to fully leverage the power of multiscale modeling across more than a few tiers of resolution. That being said, examples do exist that span the spectrum from the most fundamental genetic modifications to organ-level perturbations. Combining these tools simultaneously across all of these scales may seem at this point an intractably difficult problem; however, some preliminary efforts are already emerging.

The Physiome Project is a collection of biological databases, mathematical models, and utilities being gathered with a single purpose: integration (65, 66). Models from every spatial, temporal, and functional scale are being curated as individual modules so that they can be preserved for integration into larger, multiscale simulations. The efforts of this project are ongoing, as it recognizes that, owing primarily to computational limitations, a single, whole-organism model that explicitly incorporates all tiers of biological resolution has yet to be realized. As we noted above, the majority of information from multiscale models is concentrated near the cellular level with decreasing availability of models and data at the genetic and whole-organism levels. This ongoing effort shows much promise as a means to begin generating larger multiscale models from validated, optimized modules that have been assembled with integration in mind.

Beyond implementing better and more comprehensive multiscale models, the future of the field also holds potential to advance other recently accelerating fields of biomedical engineering. In particular, efforts in synthetic biology are using multiscale data and analysis to inform design optimization and control systems theory of novel biological systems. In a recent publication, Nawroth and colleagues (67) outlined the design, development, and implementation of a synthetic jellyfish capable of self-propulsion, which has been dubbed the Medusoid. They described the reverse engineering process as occurring over several orders of space and time in order to capture the necessary information to generate synthetic muscle fibers capable of productive, concerted contraction. Callura et al. (68) have similarly begun to use multiscale approaches as they scale up from a single gene to a composite genetic switchboard. Capable of regulating four metabolic genes in *E. coli*, their synthetic regulatory system reliably shunted flux through different carbon-utilizing pathways as measured by mRNA levels and direct quantification of metabolites. This effort demonstrates in a strictly in vitro sense how multiscale theory can be applied to better understand and engineer biological systems.

Multiscale modeling strives, above all, to better understand the fundamental processes that sustain biology. Unquestionably, effects at the genetic level are responsible for both subtle and dramatic phenotypic expression of an entire organism. We are only just starting to construct computational models that can explicitly demonstrate this same degree of emergent pattern phenomena through appropriate interscale connectivity. It is our hope that the techniques and practices presented here will guide future efforts in this field toward high-quality multiscale model implementation.

SUMMARY POINTS

1. Multiscale models are explicitly executed simulations of complex biological systems that have been integrated across temporal, spatial, and functional domains. Through simultaneous evaluation of multiple tiers of resolution, multiscale models provide access to systems behaviors that are not observable using single-scale techniques.

2. A combination of multiple computational techniques, including both continuous and discrete systems, is optimal for efficiently capturing information across biological scales. Each spatial scale can be summarized by the biological functions occupying that tier of resolution, allowing for modeling techniques to be implemented based on how well they represent these functions.
3. Multiscale models more closely recapitulate traditional benchtop experimentation while allowing for high-throughput hypothesis generation and testing, quantitation of values that cannot otherwise be measured, and translation to in vivo systems. Perturbations of high-resolution parameters (e.g., protein-binding constants) can generate low-resolution outputs that are biologically relevant (e.g., tissue developmental patterning), allowing for simultaneous access to quantifiable values across all scales of biology.

FUTURE ISSUES

1. Fundamental to the model-building process, sensitivity analyses are performed to explore the parameter space for potentially interesting and useful tuning variables on which system outputs are strongly dependent. Multiscale models must be thoroughly investigated to determine whether sensitivities are truly a function of system behavior or an artifact of coarse-graining lower-resolution outputs. This area will require further investigation through the continued development of multiscale and complex systems models.
2. Appropriate parameter selection remains a concern in the computational modeling community, as many of the parameter values required to develop multiscale models are either difficult or impossible to measure. Values obtained from in vitro data may not be suitable for multiscale models operating at a tissue-network or larger spatial scale. As such, exploration of parameter estimation techniques may be required to better parameterize multiscale models. Alternatively, emerging in vivo molecular imaging techniques may grant access to previously unobtainable parameter values.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors acknowledge funding from the National Institutes of Health (R01 GM088244 to Jason A. Papin and NIH-HL082838 to Shayn M. Peirce) and the Cystic Fibrosis Research Foundation (grant #1060 to Jason A. Papin).

LITERATURE CITED

1. Frank RN. 2004. Diabetic retinopathy. *N. Engl. J. Med.* 350:48–58
2. Geraldes P, Hiraoka-Yamamoto J, Matsumoto M, Clermont A, Leitges M, et al. 2009. Activation of PKC-δ and SHP-1 by hyperglycemia causes vascular cell apoptosis and diabetic retinopathy. *Nat. Med.* 15:1298–306

3. Modi SR, Camacho DM, Kohanski MA, Walker GC, Collins JJ. 2011. Functional characterization of bacterial sRNAs using a network biology approach. *Proc. Natl. Acad. Sci. USA* 108:15522–27
4. Seok J, Xiao W, Moldawer LL, Davis RW, Covert MW. 2009. A dynamic network of transcription in LPS-treated human subjects. *BMC Syst. Biol.* 3:78
5. Barua D, Kim J, Reed JL. 2010. An automated phenotype-driven approach (*Gene*Force) for refining metabolic and regulatory models. *PLoS Comput. Biol.* 6:e1000970
6. Milne CB, Eddy JA, Raju R, Ardekani S, Kim PJ, et al. 2011. Metabolic network reconstruction and genome-scale model of butanol-producing strain *Clostridium beijerinckii* NCIMB 8052. *BMC Syst. Biol.* 5:130
7. Cosgrove BD, Alexopoulos LG, Saez-Rodriguez J, Griffith LG, Lauffenburger DA. 2009. A multipathway phosphoproteomic signaling network model of idiosyncratic drug- and inflammatory cytokine-induced toxicity in human hepatocytes. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2009:5452–55
8. Eissing T, Kuepfer L, Becker C, Block M, Coboeken K, et al. 2011. A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks. *Front. Physiol.* 2:4
9. Shirinifard A, Gens JS, Zaitlen BL, Poplawski NJ, Swat M, Glazier JA. 2009. 3D multi-cell simulation of tumor growth and angiogenesis. *PLoS ONE* 4:e7190
10. Folcik VA, Broderick G, Mohan S, Block B, Ekbote C, et al. 2011. Using an agent-based model to analyze the dynamic communication network of the immune response. *Theor. Biol. Med. Model.* 8:1
11. Caputo KE, Hammer DA. 2009. Adhesive dynamics simulation of G-protein-mediated chemokine-activated neutrophil adhesion. *Biophys. J.* 96:2989–3004
12. Pospieszalska MK, Zarbock A, Pickard JE, Ley K. 2009. Event-tracking model of adhesion identifies load-bearing bonds in rolling leukocytes. *Microcirculation* 16:115–30
13. Fallahi-Sichani M, Kirschner DE, Linderman JJ. 2012. NF-κB signaling dynamics play a key role in infection control in tuberculosis. *Front. Physiol.* 3:170
14. Brown BN, Price IM, Toapanta FR, DeAlmeida DR, Wiley CA, et al. 2011. An agent-based model of inflammation and fibrosis following particulate exposure in the lung. *Math. Biosci.* 231:186–96
15. Scheff JD, Mavroudis PD, Calvano SE, Lowry SF, Androulakis IP. 2011. Modeling autonomic regulation of cardiac function and heart rate variability in human endotoxemia. *Physiol. Genomics* 43:951–64
16. Sander EA, Stylianopoulos T, Tranquillo RT, Barocas VH. 2009. Image-based multiscale modeling predicts tissue-level and network-level fiber reorganization in stretched cell-compacted collagen gels. *Proc. Natl. Acad. Sci. USA* 106:17675–80
17. Sharafi B, Ames EG, Holmes JW, Blemker SS. 2011. Strains at the myotendinous junction predicted by a micromechanical model. *J. Biomech.* 44:2795–801
18. Adra S, Sun T, MacNeil S, Holcombe M, Smallwood R. 2010. Development of a three dimensional multiscale computational model of the human epidermis. *PLoS ONE* 5:e8511
19. Greenstein JL, Winslow RL. 2011. Integrative systems models of cardiac excitation-contraction coupling. *Circ. Res.* 108:70–84
20. Moreno JD, Zhu ZI, Yang PC, Bankston JR, Jeng MT, et al. 2011. A computational model to predict the effects of class I anti-arrhythmic drugs on ventricular rhythms. *Sci. Transl. Med.* 3:98ra83
21. Thorne BC, Hayenga HN, Humphrey JD, Peirce SM. 2011. Toward a multi-scale computational model of arterial adaptation in hypertension: verification of a multi-cell agent based model. *Front. Physiol.* 2:20
22. Holland DO, Krainak NC, Saucerman JJ. 2011. Graphical approach to model reduction for nonlinear biochemical networks. *PLoS ONE* 6:e23795
23. Schmidt BJ, Papin JA, Lawrence MB. 2009. Nano-motion dynamics are determined by surface-tethered selectin mechanokinetics and bond formation. *PLoS Comput. Biol.* 5:e1000612
24. Bailey AM, Thorne BC, Peirce SM. 2007. Multi-cell agent-based simulation of the microvasculature to study the dynamics of circulating inflammatory cell trafficking. *Ann. Biomed. Eng.* 35:916–36
25. Friedl P, Wolf K. 2010. Plasticity of cell migration: a multiscale tuning model. *J. Cell Biol.* 188:11–19
26. Grima R. 2008. Multiscale modeling of biological pattern formation. *Curr. Top. Dev. Biol.* 81:435–60
27. N'Dri NA, Shyy W, Tran-Son-Tay R. 2003. Computational modeling of cell adhesion and movement using a continuum-kinetics approach. *Biophys. J.* 85:2273–86

28. Brodland GW, Chen X, Lee P, Marsden M. 2010. From genes to neural tube defects (NTDs): insights from multiscale computational modeling. *HFSP J.* 4:142–52
29. Sample C, Shvartsman SY. 2010. Multiscale modeling of diffusion in the early *Drosophila* embryo. *Proc. Natl. Acad. Sci. USA* 107:10092–96
30. Tomlin CJ, Axelrod JD. 2007. Biology by numbers: mathematical modelling in developmental biology. *Nat. Rev. Genet.* 8:331–40
31. von Dassow G, Meir E, Munro EM, Odell GM. 2000. The segment polarity network is a robust developmental module. *Nature* 406:188–92
32. Meir E, Munro EM, Odell GM, Von Dassow G. 2002. Ingeneue: a versatile tool for reconstituting genetic networks, with examples from the segment polarity network. *J. Exp. Zool.* 294:216–51
33. Longo D, Peirce SM, Skalak TC, Davidson L, Marsden M, et al. 2004. Multicellular computer simulation of morphogenesis: blastocoel roof thinning and matrix assembly in *Xenopus laevis*. *Dev. Biol.* 271:210–22
34. Silva JR, Rudy Y. 2010. Multi-scale electrophysiology modeling: from atom to organ. *J. Gen. Physiol.* 135:575–81
35. Aguado-Sierra J, Krishnamurthy A, Villongco C, Chuang J, Howard E, et al. 2011. Patient-specific modeling of dyssynchronous heart failure: a case study. *Prog. Biophys. Mol. Biol.* 107:147–55
36. Newman SA, Christley S, Glimm T, Hentschel HG, Kazmierczak B, et al. 2008. Multiscale models for vertebrate limb development. *Curr. Top. Dev. Biol.* 81:311–40
37. Das A, Lauffenburger D, Asada H, Kamm RD. 2010. A hybrid continuum-discrete modelling approach to predict and control angiogenesis: analysis of combinatorial growth factor and matrix effects on vessel-sprouting morphology. *Philos. Trans. R. Soc. A* 368:2937–60
38. Liu G, Qutub AA, Vempati P, Mac Gabhann F, Popel AS. 2011. Module-based multiscale simulation of angiogenesis in skeletal muscle. *Theor. Biol. Med. Model.* 8:6
39. Vempati P, Popel AS, Mac Gabhann F. 2011. Formation of VEGF isoform-specific spatial distributions governing angiogenesis: computational analysis. *BMC Syst. Biol.* 5:59
40. Bauer AL, Jackson TL, Jiang Y, Rohlf T. 2010. Receptor cross-talk in angiogenesis: mapping environmental cues to cell phenotype using a stochastic, Boolean signaling network model. *J. Theor. Biol.* 264:838–46
41. Hashambhoy YL, Chappell JC, Peirce SM, Bautch VL, Mac Gabhann F. 2011. Computational modeling of interacting VEGF and soluble VEGF receptor concentration gradients. *Front. Physiol.* 2:62
42. Causey L, Cowin SC, Weinbaum S. 2012. Quantitative model for predicting lymph formation and muscle compressibility in skeletal muscle during contraction and stretch. *Proc. Natl. Acad. Sci. USA* 109:9185–90
43. Coddington EA, Levinson N. 1955. *Theory of Ordinary Differential Equations*. New York: McGraw-Hill. 429 pp.
44. Fallahi-Sichani M, Flynn JL, Linderman JJ, Kirschner DE. 2012. Differential risk of tuberculosis reactivation among anti-TNF therapies is due to drug binding kinetics and permeability. *J. Immunol.* 188:3169–78
45. Quo CF, Moffitt RA, Merrill AH, Wang MD. 2011. Adaptive control model reveals systematic feedback and key molecules in metabolic pathway regulation. *J. Comput. Biol.* 18:169–82
46. Laise P, Di Patti F, Fanelli D, Masselli M, Arcangeli A. 2011. Deterministic and stochastic aspects of VEGF-A production and the cooperative behavior of tumoral cell colony. *J. Theor. Biol.* 272:55–63
47. Sun T, Adra S, Smallwood R, Holcombe M, MacNeil S. 2009. Exploring hypotheses of the actions of TGF-β1 in epidermal wound healing using a 3D computational multiscale model of the human epidermis. *PLoS ONE* 4:e8515
48. Haggart CR, Bartell JA, Saucerman JJ, Papin JA. 2011. Whole-genome metabolic network reconstruction and constraint-based modeling. *Methods Enzymol.* 500:411–33
49. Tveito A, Lines GT, Edwards AG, Maleckar MM, Michailova A, et al. 2012. Slow calcium-depolarization-calcium waves may initiate fast local depolarization waves in ventricular tissue. *Prog. Biophys. Mol. Biol.* 110:295–304
50. Thompson AB, Gayzik FS, Moreno DP, Rhyne AC, Vavalle NA, Stitzel JD. 2012. A paradigm for human body finite element model integration from a set of regional models. *Biomed. Sci. Instrum.* 48:423–30
51. Zhang X, Gan RZ. 2011. A comprehensive model of human ear for analysis of implantable hearing devices. *IEEE Trans. Biomed. Eng.* 58:3024–27

52. Zahedmanesh H, Lally C. 2011. A multiscale mechanobiological modelling framework using agent-based models and finite element analysis: application to vascular tissue engineering. *Biomech. Model. Mechanobiol.* 11:363–77
53. Goktepe S, Abilez OJ, Parker KK, Kuhl E. 2010. A multiscale model for eccentric and concentric cardiac growth through sarcomerogenesis. *J. Theor. Biol.* 265:433–42
54. Du P, O'Grady G, Davidson JB, Cheng LK, Pullan AJ. 2010. Multiscale modeling of gastrointestinal electrophysiology and experimental validation. *Crit. Rev. Biomed. Eng.* 38:225–54
55. Trayanova NA, Rice JJ. 2011. Cardiac electromechanical models: from cell to organ. *Front. Physiol.* 2:43
56. Artel A, Mehdizadeh H, Chiu YC, Brey EM, Cinar A. 2011. An agent-based model for the investigation of neovascularization within porous scaffolds. *Tissue Eng. Part A* 17:2133–41
57. Bentley K, Gerhardt H, Bates PA. 2008. Agent-based simulation of notch-mediated tip cell selection in angiogenic sprout initialisation. *J. Theor. Biol.* 250:25–36
58. Qu Z, Garfinkel A, Weiss JN, Nivala M. 2011. Multi-scale modeling in biology: how to bridge the gaps between scales? *Prog. Biophys. Mol. Biol.* 107:21–31
59. Valentin A, Humphrey JD. 2009. Evaluation of fundamental hypotheses underlying constrained mixture models of arterial growth and remodelling. *Philos. Trans. R. Soc. A* 367:3585–606
60. Valentin A, Humphrey JD. 2009. Parameter sensitivity study of a constrained mixture model of arterial growth and remodeling. *J. Biomech. Eng.* 131:101006
61. Hayenga HN, Thorne BC, Peirce SM, Humphrey JD. 2011. Ensuring congruency in multiscale modeling: towards linking agent based and continuum biomechanical models of arterial adaptation. *Ann. Biomed. Eng.* 39:2669–82
62. Fedosov DA, Lei H, Caswell B, Suresh S, Karniadakis GE. 2011. Multiscale modeling of red blood cell mechanics and blood flow in malaria. *PLoS Comput. Biol.* 7:e1002270
63. Kim MJ, Maly IV. 2009. Deterministic mechanical model of T-killer cell polarization reproduces the wandering of aim between simultaneously engaged targets. *PLoS Comput. Biol.* 5:e1000260
64. Singh S, Narang AS, Mahato RI. 2011. Subcellular fate and off-target effects of siRNA, shRNA, and miRNA. *Pharm. Res.* 28:2996–3015
65. Bassingthwaighte JB. 2010. Predictive modeling and integrative physiology: the Physiome Projects. *Open Pacing Electrophysiol. Ther. J.* 3:66–74
66. Bassingthwaighte JB, Chizeck HJ. 2008. The Physiome Projects and multiscale modeling. *IEEE Signal Process. Mag.* 25:121–44
67. Nawroth JC, Lee H, Feinberg AW, Ripplinger CM, McCain ML, et al. 2012. A tissue-engineered jellyfish with biomimetic propulsion. *Nat. Biotechnol.* 30:792–97
68. Callura JM, Cantor CR, Collins JJ. 2012. Genetic switchboard for synthetic biology applications. *Proc. Natl. Acad. Sci. USA* 109:5850–55

Annual Review of Biomedical Engineering
Volume 15, 2013

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Biomedical Engineering* articles may be found at http://bioeng.annualreviews.org/