Modeling of Biological Processes in an Object-Oriented Environment

John R. Hummel[1], Robert B. Silver[1,2], and John H. Christiansen[1]

[1]Advanced Computer Applications Center
Argonne National Laboratory
9700 S. Cass Avenue/DIS-900
Argonne, IL 60439-4832

[2]Marine Biological Laboratory
7 MBL Street
Woods Hole, MA 02543-1015
jhummel@anl.gov, rsilver@mbl.edu, jhc@anl.gov

The submitted manuscript has been created by the University of Chicago as Operator of Argonne National Laboratory ("Argonne") under Contract No. W-31-109-ENG-38 with the U.S. Department of Energy. The U.S. Government retains for itself, and others acting on its behalf, a paid-up, non-exclusive, irrevocable world-wide license in said article to reproduce, prepare derivative works, distribute copies to the public and display publicly, by or on behalf of the Government.

KEYWORDS: Cell biology, cell decisions, molecular biology, object-oriented, high performance computing

ABSTRACT

The detailed modeling of biological systems is one of the "grand challenges" in the computational world. Of particular interest is to learn the mechanism through which cells make decisions and regulate their life processes. This challenge entails representing known processes at multiple spatial and temporal domains and unraveling the nature of unknown processes. The spatial domains range from individual molecules to complex molecular assemblies to cells and tissues to complete organisms. The temporal domains range from femtosecond-duration reactions through the life span of an organism.

Calcium signals play an essential role in controlling cell decisions and processes, although the mechanism of calcium action is unresolved. Calcium signals are essential for the control of nuclear envelope breakdown, a step in cell division and reproduction. Quantitative analyses demonstrate that those signals are encoded as space-time-patterned calcium emissions called quantum emission domains (QEDs). These QEDs occur within microdomains at the micrometer and sub-micrometer scales.

Knowledge of the selective signaling mechanism and metabolic networks by which cells generate and control these signals will promote major new insights to understanding calcium-dependent processes such as cell division. Gaining that knowledge requires the integration of kinetic analyses, tomographic reconstructions, and modeling of the component facets of the calcium regulatory process.

A significant challenge is the development of a global model and data management systems to (a) make data retrievable and available; (b) more readily usable for quantitative analyses; and (c) amenable to complex queries that span multiple data and information formats and their complexities and file compositions. The successful global model will provide for bench-verification and bench-investigation of computationally discovered phenomena – with subsequent feedback to the theory and modeling.

In this presentation, we discuss how object-oriented modeling and simulation technologies will be used to support the modeling of biological systems. The goal is to establish a framework for studying dynamic signal and information processing capabilities of essential intracellular biological regulatory networks necessary for control of cell division.

The Challenge: Understand How Cells Make Decisions

Understanding how cells orchestrate and utilize their many components in a manner that yields mitotic cell division is an important goal for biologists. Mitotic cell division is an essential feature of the biology of cells, and one key aspect of mitosis is that of nuclear envelope breakdown (NEB). In the NEB process, cells disassemble the "shell" of the nucleus and thus expose the chromosomes to cytoplasmic components. That exposure allows chromosomes to be captured by the mitotic apparatus that apportions the correct number and complement of chromosomes to each new daughter cell. Without NEB, the chromosomes would not be segregated to the new cells, and those cells would malfunction and die. Thus, nuclear envelope breakdown is an important event in the mitotic process. Figure 1 is a schematic representation of the mitotic process.

Calcium signals play an essential role in controlling cell processes, although the mechanism through which calcium acts is unresolved. Knowledge of the selective mechanism and signaling network, by which cells generate and control these signals will promote major new insight to understanding calcium-dependent processes (Silver [1989 and 1996] and Llinás et. al. [1992]). Obtaining that knowledge requires integration of kinetic analyses, morphological reconstructions and modeling of the component facets of the calcium regulatory process.

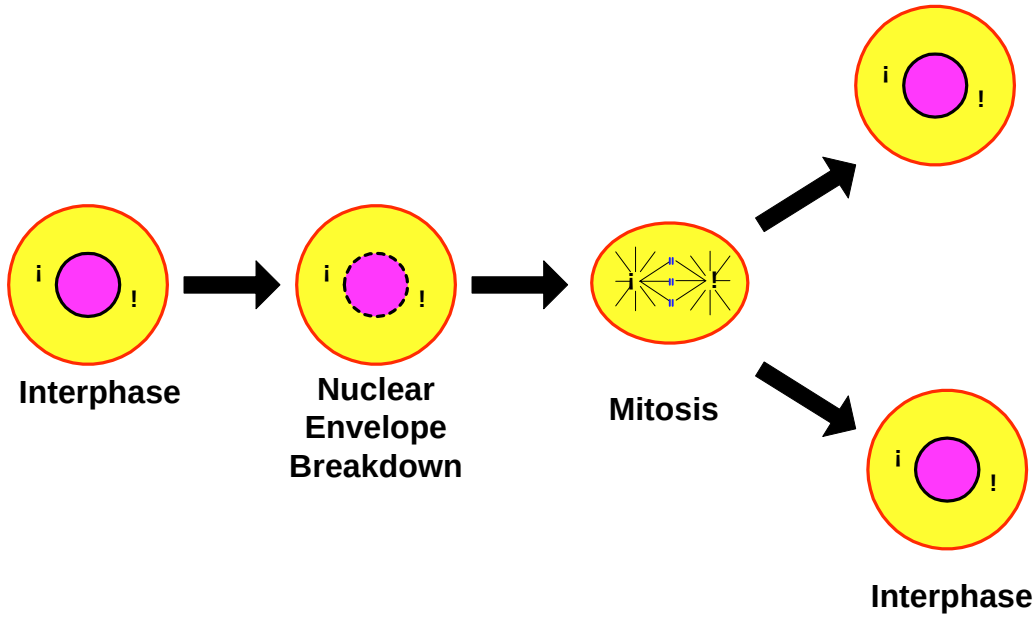


Figure 1. Schematic representation of the mitotic process.

The Role of Calcium in Cell Division

Calcium signals that control cell function present a profound conundrum. Calcium signals are considered ubiquitous in cell biology and are selective in exerting control on specific regulated processes amongst many functions within a common cytoplasm. How this complex messaging is achieved through a single ion - calcium - remains an unsolved puzzle. Among the fundamental cellular processes, a calcium signal precedes and is required for NEB.

Calcium regulation in mitosis is extremely complex. Regulation is achieved through a system of perinuclear calcium-regulatory endomembrane vesicles, i.e., 100-200 nm diameter "shells" bounded by phospholipids. Calcium signals derived from endomembrane stores coordinate and activate various biochemical processes required for events such as NEB, phosphorylations (e.g. Silver *et al.* [1980, 1996] and Silver [1986]. An increasing body of evidence indicates that temporally patterned Ca^{2+} signals, with durations ranging from milliseconds through tens of seconds, are essential for the regulation of cell function. These Ca^{2+} signals can be evoked through several combinations of agonists and Ca^{2+} pools, including 1,4,5-inositol trisphosphate (IP_3), cyclic ADP-ribose, NAADP, and Leukotriene B_4. Various calcium release mechanisms have been described and modeled, including: wholesale global release, waves, oscillations, and fine-structured space-time patterns of very limited spatial and temporal scope. Some calcium signals appear to spread in steps over large portions of cytoplasm, while others appear to remain localized to the site of release.

The Scientific Challenge Posed by the Physical Limits of Microscopy and Cell Biology

Accurate modeling of changes in local calcium concentration, and of the regulatory mechanisms that control the pre-NEB calcium signal, requires a detailed understanding of the cellular components that occupy the space in which control is implemented. A wide range of evidence implicates membrane-bounded vesicles in the sequestration of calcium**,** so our analysis of calcium regulation *in situ* must employ methods that allow us to identify the position and size of such vesicles and their endogenous regulatory enzymes and substrates with considerable precision. An accurate three-dimensional (3-D) reconstruction of the calcium regulatory endomembranes and their component enzymes, within the mitotic apparatus, is needed.

Two tools available for such reconstructions are light and electron microscopy. Light microscopy has the value of allowing work on living cells, facilitating the analysis of a significant volume from each specimen (as well as sampling many specimens), and allowing direct determination of calcium concentrations. Indeed, it is with multispectral video light microscopy that the calcium signals are observed. Light microscopy, including confocal microscopy, has a limitation, however; the spatial resolution of such data is limited to ~200 nm by the physics of image formation with visible light.

The calcium regulatory vesicles, upon which the metabolic regulation is conducted, appear as roughly spherical bodies of 100–200 nm in diameter. This is well below the resolving power of light microscopy. While

these are readily observed with a transmission electron microscope, the process of specimen preparation kills the cells.

The unit actions of enzymes and ion channel occur at temporal scales between 10^{-10} and 10^{-3} seconds. Such rates far exceed current video detection methods, even for high-speed video.

Thus, the "kernals" of the cell biology we study occur at spatial and temporal scales that are beyond our immediate grasp; we are faced with real physical limits to these studies. We are also faced with analyses of data in many formats from many different sources (*i.e.*, video records; spectroscopic and spectrophotometric; high-resolution electron microscope images; parameters, constraints, and output of various models). As an example, the multispectral video microscopy in this study generates 0.5 GBytes of image data per minute, *i.e.*, 0.34 TBytes per six-hour experiment for four channels of video. To further complicate matters, these data are acquired at different sites, and needed metadata may be incomplete.

The Computational Challenge to Analyze and Model the Cell Division Process

The computational load for this project is enormous, both from the perspective of analyzing the data and modeling the physics. An example of the challenge from the analysis perspective is shown in Table 1, which summarizes the computational issues in the processing of video microscopy data. The results in the table show the computational times associated with the application of simple spatial and temporal filters and the reconstruction and rendering of the multispectral data will need 5.86×10^{13} floating point operations (flops). Analysis of the data adds to that load.

To realize the vast potential of computational signal analysis, modeling, and simulation, the computation should occur within the experimental loop, *i.e.*, real-time teraop-class supercomputing. Recently, we estimated that a single iteration of a numerical solution for reaction-diffusion of the metabolites and calcium would require 1.73×10^{17} flops, roughly 30 hours on all of the processors of Nirvana Blue running at 1.6 teraflops. Our goal is to analyze Ca^{2+} QED patterns from pre-NEB signals mitotic cells to compare the experimental and model results. Such an analysis will enable us to ascertain the primary kinetic mechanisms responsible for the complex calcium signal observed.

Taking this approach, we begin with modeling on desktop workstations and small clusters of workstations. Gathering adequate statistics on the particle models requires supercomputing. For example, one 50-second realization of the particle model simulating only the LtB_4 receptor/channel takes several hours on a desktop computer. This estimate is derived from the following considerations: we assume that 0.1-millisecond time steps resolves the fastest time scales, that the peak number of cytosolic Ca^{2+} ions is about 10^4, and that the equivalent of about 100 flops is required to update the position of each of these ions. This translates to 50 seconds $\times$ 10^4 time steps/second $\times$ 10^4 particles $\times$ 100 flops per time step per particle = 5×10^{11} flops per realization. A 500 Mflop desktop machine working at ten percent of peak performance will take about 10^4 seconds, or about 3 hours, to generate a single realization. This increases about tenfold if the remainder of the main pathway from membrane lipid through arachidonic acid, 5-HPETE, LtA_4, and finally LtB_4 is included in the simulation. To compare to experiment, several thousand realizations should be generated and the resulting signals compared to experiment. This process is iterated to obtain estimates of the rate constants. The computational effort for a single realization of the model is modest. However, to generate statistically meaningful results, the computational effort becomes substantially larger. These simulations can be run in parallel, and we are planning to exploit available supercomputer resources to perform these simulations.

Table 1. Overview of the computational load associated with the processing of video microscopy data of cell division processes.

	Floating Point Operations Required to Analyze Different Lengths of Video Data			
Processing Task	**33 msec**	**1 sec**	**1 minute**	**1 hour**
Application of Spatial Filter	2.62×10^7	7.86×10^8	4.72×10^{10}	2.83×10^{11}
Application of Temporal Filter	7.86×10^7	2.36×10^9	1.42×10^{11}	8.52×10^{12}
Reconstruction of Data	1.31×10^7	3.93×10^8	2.36×10^{10}	1.42×10^{11}
Image Rendering	1.31×10^7	3.93×10^8	2.36×10^{10}	1.42×10^{11}
Total Required Operations	$\mathbf{1.31 \times 10^8}$	$\mathbf{3.93 \times 10^9}$	$\mathbf{2.36 \times 10^{11}}$	$\mathbf{2.44 \times 10^{12}}$

Developing a Model of the Calcium Signal Process

We are developing a modeling framework to study the role of calcium in the cell division process based on simulation and information technologies developed at Argonne National Laboratory. The advantage of this approach is that it integrates data from analysis of calcium signals observed with multispectral video microscopy, tomographic reconstructions of light and electron microscopy data, and biochemical and enzymological studies, including Michaelis-Menten kinetics. This level of integration of "wet" and modeling experiments and analysis will be unique in biology, and we believe it will present a new standard for quantitative analytical studies of complex biological systems.

The level of complexity of the biological system is vast and far beyond lists of differential equations of S-system compartment models, which assume that the system being modeled is essentially uniform. Cell and regions of a cell morphologically and biochemically exhibit discontinuous structure and organization. Clearly then, the analyses of cellular metabolism and regulation require a similar form of analysis. Considering cellular components, microdomains, signals and processes as discrete objects is consistent with the structure and organization of cells and their metabolic and regulatory machinery and mechanisms. Figure 2 shows an overview of the modeling and analysis processes we are trying to represent, while Figure 3 describes the activation pathways that need to be modeled.

One of the tools being exploited in this effort is the Dynamic Information Architecture System (DIAS), an object-oriented simulation framework. DIAS has been used in a number of extensive modeling and simulation applications, including the simulation of a health care organization and a number of environmental problems.

DIAS has a large library of existing objects, many of which can be applied to this problem. For example, a set of generalized transport objects has been designed that can be adapted for chemical transport at the molecular level. Furhter, a Person object which can exhibit physiological behaviors has been developed which can be extended to encompass processes down to the organ, tissue, or, in the extreme, molecular level.

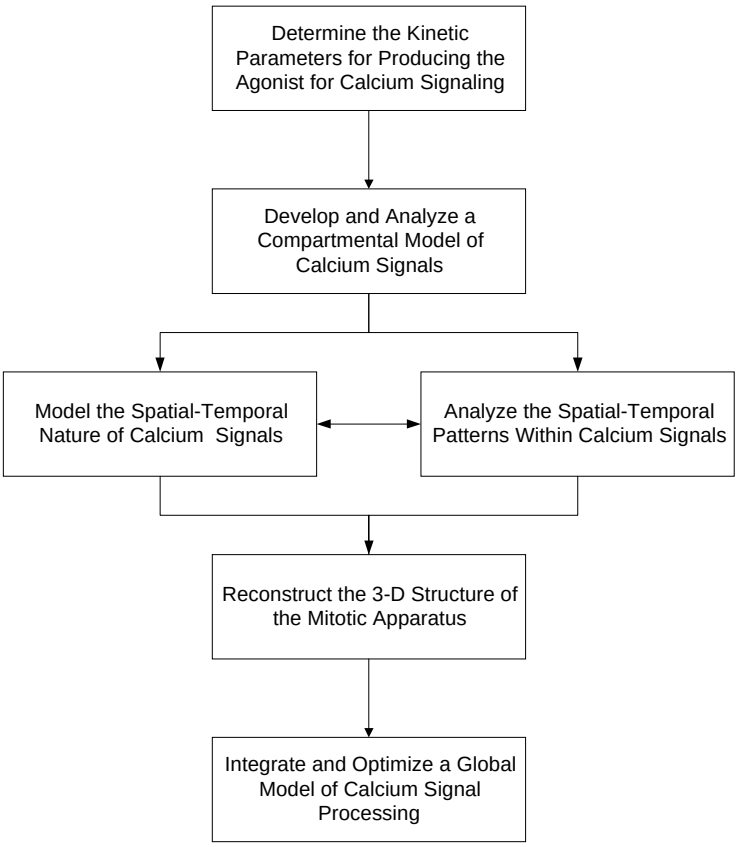


Figure 2. Representation of the modeling and analysis processes being developed to study to the role of calcium in cell division.

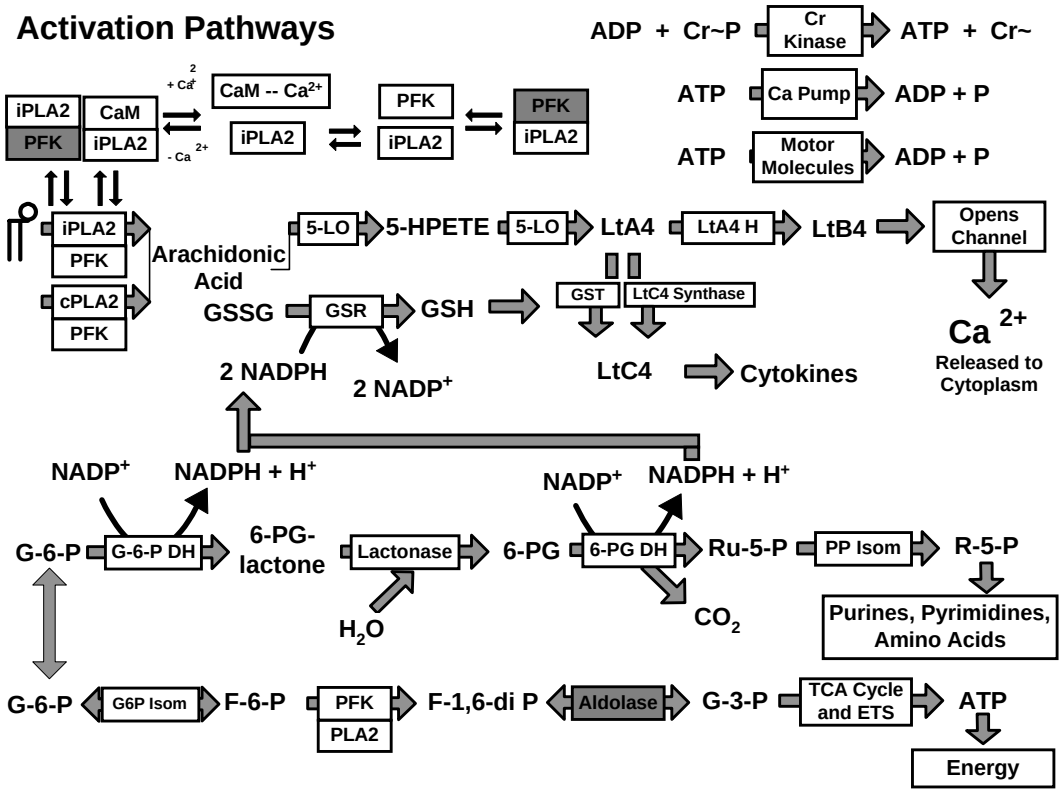


Figure 3. Activation pathways being studied to understand the role of calcium in cell division.

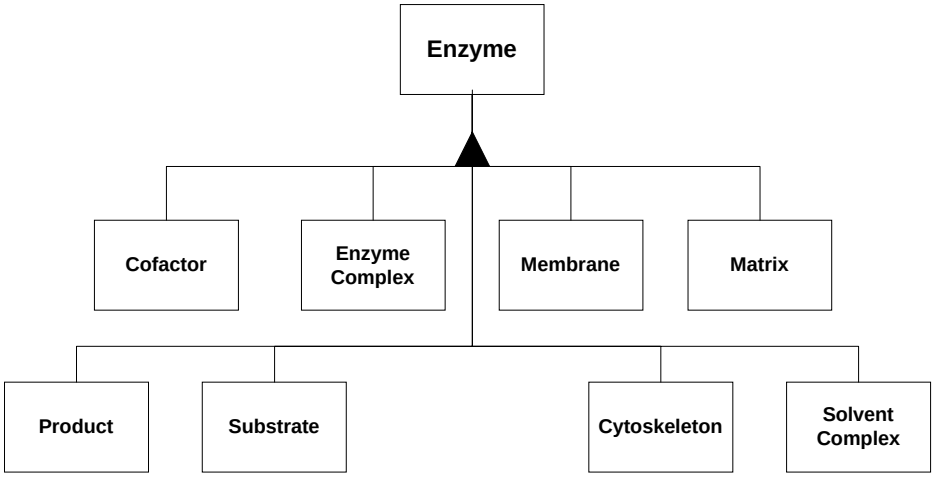
Adopting the DIAS framework represents a significant, and important, step forward in modeling of biological systems. With the DIAS approach, for the first time, the full suite of attributes of each enzyme or other cellular component or process – from data and metadata – acting within a spatial and temporal structure can be assessed in a dynamic fashion that more faithfully represents cellular processes and behavior.

Figure 4 represents an initial design concept for an Enzyme object. In the figure, we show an initial description of the subclasses required to describe an Enzyme, and we list some of the behaviors we need to be able to model, along with some of the external factors that can impact these behaviors.

The process of using object-oriented design, in turn, will significantly facilitate analyses and quicken our enhanced understanding of the impact of interactive metabolic dynamics, i.e., how different enzymes and pathways and their positive and negative feedbacks influence cellular processes and thus cellular decisions in a way that cells actually use.

For biologists, who typically eschew mathematical approaches (often seeming to be "equation-phobic"), such an approach may seem to represent an unwelcome extreme. However, this approach, with its inherent tolerance for incomplete data sets, may well hasten analyses of integrated, multiscale systems. Once developed for the cell decisions of cell division, and the attendant metabolic processes, other kernals can be analyzed in a similar fashion. While each of such systems can be viewed as an independent study, as is typically the contemporary fashion, DIAS will be able to consider each of these kernals as separate objects unto themselves. Within this paradigm, DIAS will then be able to help researchers identify interactions, influences, commonalties, and unique attributes of each of the cellular systems and each different type of cell; truly a "bottom-up" integrated approach to modeling a complex system as vast as a cell.

Cellular decisions are computed on a scale far beyond that which we can presently achieve with silicon-based microprocessors. Trying to understand the cell decision process is a significant challenge that crosses several disciplinary lines. The computational challenges we face are enormously complex and in keeping with the importance of the questions. In the end, our models will arrive at numerical solutions that more faithfully simulate cellular biology and help deepen our understanding and appreciation of normal and diseased cells, ultimately having important implications for fundamental knowledge and public health.



Behaviors:

- **Reactions (Forward, Backward, & Secondary)**
- **Bifunctional Activity**
- **Substrate Channeling**
- **Metabolic Pathway**
- **Interacting Pathways**
- **Cell Cycle Regulation**
- **Developmental Controls**
- **Gene Expression**
- **Apoptosis**
- **Transformation of Cell**

Constraining External Factors:

- **Substrate, Cofactor, and Product Pools**
- **Restricted Diffusion**
- **Complexed Enzymes, Proteins, Lipids, Carbohydrates, DNA, & RNA**
- **Local and Bound Ions**
- **Local and Bound Water**
- **Total Solvent**
- **Local Viscosity**
- **Temperature**
- **Pressure**
- **Conductance of Milieu**
- **Photonic and Ionic Spectrum**

Figure 4. Preliminary design of an enzyme object, including a list of behaviors and constraining external factors.

ACKNOWLEDGMENT

This work is sponsored by the U. S. Department of Energy under contract number W-31-109-ENG-38 and under grants to R.B.S. from the M.C.B. Program of the National Science Foundation (*e.g.*, NSF MCB-99082680), and that support is gratefully acknowledged.

REFERENCES

Llinás, R., M. Sugimori, and R. B. Silver, 1992. Microdomains of high calcium concentration in a presynaptic terminal, *Science*. 256:677-679.

Silver, R.B., R. D. Cole, and W. Z. Cande, 1980. Isolation of mitotic apparatus containing vesicles with calcium sequestration activity, *Cell* 19:505-516.

Silver, R.B., 1986. Mitosis in sand dollar embryos is inhibited by antibodies directed against the calcium transport enzyme of muscle, *Proc. Nat. Acad. Sci. U.S.A.* 83:4302-4306.

Silver, R.B., 1989. Nuclear envelope breakdown and mitosis in sand dollar embryos is inhibited by microinjection of calcium buffers in a calcium-reversible fashion, and by antagonists of intracellular Ca^{2+}-channels, *Develop. Biol.* 131:11-26.

Silver, R.B., 1996. Calcium, BOBs, QEDs, Microdomains and a Cellular Decision: Control of Mitotic Cell Division in Sand Dollar Blastomeres, *Cell Calcium.* 20:161-179.