BIOSILICO Vol. 1, No. 1 March 2003

BioSig: an informatics framework for representing the physiological responses of living cells

Bahram Parvin and Daniel E. Callahan

Subcellular experimental datasets and detailed cell models are required before modeling of whole organs. Cell modeling requires repeated interaction between simulation and experimental data. This review describes a coupled system of informatics and instrument control suitable for extracting information at the subcellular level. The BioSig informatics framework annotates time-series images with experimental variables and computed representation such as physiological responses, whereas visual servoing optical microscopy (VSOM) is used for adaptive perturbation and interaction with living cells.

Bahram Parvin*
Imaging and Informatics
Computational Research Division
Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 50B-2239
Berkeley, CA 94720, USA
*tel: +1 510 486 6203
e-mail: parvin@media.lbl.gov
Daniel E. Callahan
Cell and Molecular Biology Department
Life Sciences Division
Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 74-157
Berkeley, CA 94720, USA

▼ Progress in systems biology, systems physiology and physiomics has been recently reviewed [1–5]. In one example, detailed cellular models and cellular experiments were shown to play a crucial role in the successful construction of a whole-organ model of the heart [5]. The ability to trace genetic effects from abnormal protein function to cellular pathophysiology and then to organ abnormalities depended on a rich cellular dataset. Moreover, successful cell modeling in this example depended on the repeated interaction between simulation and the experimental dataset [5]. The BioSig (http://vision.lbl.gov/Projects/BioSig, http://vision.lbl.gov/Projects/vsom) imaging informatics system described here provides a framework for acquiring similar subcellular datasets for a variety of cell types.

BioSig has been used for digital imaging studies of microscope specimens that have been sectioned, fixed and stained [6]. However, BioSig also contains features designed for real-time interaction with living cells across a distributed system. Such features facilitate the cataloging and analysis of subcellular, living-cell responses that are observed using digital imaging fluorescence microscopy.

Advances in fluorescence microscopy and fluorescence-probe technologies have made it possible to collect responses of living cells across multiple channels of information [7,8]. BioSig has been designed to collect physiological responses together with experimental variables and the state of the instrument. Large amounts of multichannel spatio-temporal data are effectively represented in BioSig for further analysis, manipulation and remote access. The effective representation of these data enables dynamic microscope control for online studies while simultaneously making these representations accessible through the Internet.

Visual-servoing optical microscopy (VSOM) protocols and instrumentation accomplish real-time interactions with living cells by repeatedly performing the following steps:

(1) perturbation or stimulation of cells (immobilized on a surface in a perfusion chamber) via computer-controlled syringe pumps;
(2) analysis of the resultant physiological responses of individual cells and organelles; and
(3) adjustment of pump operation based on observed cell responses and previously archived cell responses [9].

The BioSig framework is designed to make the richly annotated cell responses of previous VSOM experiments available during this process. In this way, instrument control can be knowledge based because decisions can be made by referring to archived cell responses obtained under similar experimental conditions. In addition, instrument control can be adaptive, because perturbations are adjusted based on observed cell responses, or cell simulations running interactively or in parallel with on-line VSOM experiments. Such a coupling

1478-5282/03/$ – see front matter PII: S1478-5282(02)02230-4

between informatics, analysis, living cells and instrumentation will help automate and accelerate iterations of the research-cycle consisting of hypothesis, model testing and experimentation [2].

Problem-solving environments such as these have been applied to several scientific domains for the construction of virtual laboratories [10]. However, most of the present systems do not address the comparison of simulation and experimental data through automated analysis of time-series images.

The process of generating, modulating and capturing cell responses also provides a basis for the construction of extensive, cell-type specific, physiological response databases *in silico*. Such databases would consist of a large collection of physiological responses and subsequent biological endpoints such as apoptosis, proliferation or differentiation. Cell-type specific responses would be collected for a wide range of perturbations and stimuli that are administered under different micro-environmental conditions. These *in silico* VSOM datasets would represent comprehensive information on the dynamic behavior of individual cells as a system.

It has been noted that successful biological simulation begins with a foundation of rich, cellular-level biological data [5]. In this context, physiological responses can serve as a foundation for the construction of cellular physiomes. As *in silico* physiological response representations grow in size and complexity, they will become increasingly useful for offline data mining and online, automated discovery and optimization of the differences between cell types. This will aid the design of: (1) specific cell-based fluorescence assays suitable for multiwell plate screening systems; (2) pharmaceuticals and delivery systems that target specific cells, cell compartments, phenotypes or physiological states; and (3) *in vitro* cell culture conditions suitable for the selective propagation of specific cell types.

Instrumentation

VSOM instrumentation consists of an automated, inverted fluorescence microscope and computer-controlled perfusion pumps that alter the microenvironment of living cells in a micro-perfusion chamber. Solutions are perfused at a known flow rate into the chamber. Software has been developed for automatic collection of time-lapsed multichannel images, which are annotated with experimental variables and the state of the instrument. Furthermore, each image is automatically annotated with a summary of observed physiological responses corresponding to specific subcellular compartments of individual cells.

Adaptive extraction of physiological responses can be achieved in two operational modes. The instrumentation can be programmed with predefined thresholds (representing the magnitude of physiological responses) and run independently in a stand-alone mode, or the instrument control can be coupled with the BioSig informatics framework across a computer network for knowledge-based instrument control via real-time access to archived annotations, perturbations and cellular responses. In the first operational mode, images, log files and computed annotations are deposited into a file system for subsequent archiving into a database. In the second mode, VSOM accesses the database for matching on-line computed physiological responses against archived ones. Our previous work in distributed instrumentation and visual-servoing for micropositioning also lays the foundation for future distributed VSOM instrumentation (multiple systems running in parallel) and repetitive repositioning capabilities (for large cell-population, multiple field-of-view VSOM studies) [11–13].

Informatics

BioSig maintains an integrated view of the lab notebook, the state of the instrument, and time-lapsed images that are analyzed for quantification of physiological responses. As a result, different experimental results can be compared for validation, instrument control, exploratory analysis, simulation and modeling. The informatic system comprises three components: (1) the data model, (2) the presentation manager and (3) the query manager. The data model couples experimental variables with the computed representation of images for subsequent analysis. The model is object-orientated and enables bi-directional tracking of annotation and measured feature data. The presentation manager provides mapping between the data model and the user interface, and the display functionality of a particular query, in either text or graphics. The query manager maps high-level queries to the Java objects that implement the data model.

The data model (Fig. 1) provides navigational links between the different modules: (1) experimental variables and the state of the instrument, (2) time-series images, (3) detailed quantitative annotations and (4) the system biology markup language (SBML, http://www.sbw-sbml.org/index.html). The SBML provides an important component for the construction of kinetic models from a collection of physiological responses under known perturbation conditions. The significance of this data model is that it supports both fixed-specimen and living-cell VSOM studies. Experimental variables include cell type, conditioning reagents, pre-treatments (e.g. pharmaceutical exposure) and detailed VSOM protocols. Captured images are compressed, annotated and archived in the database. Physiological responses and other quantitative annotations

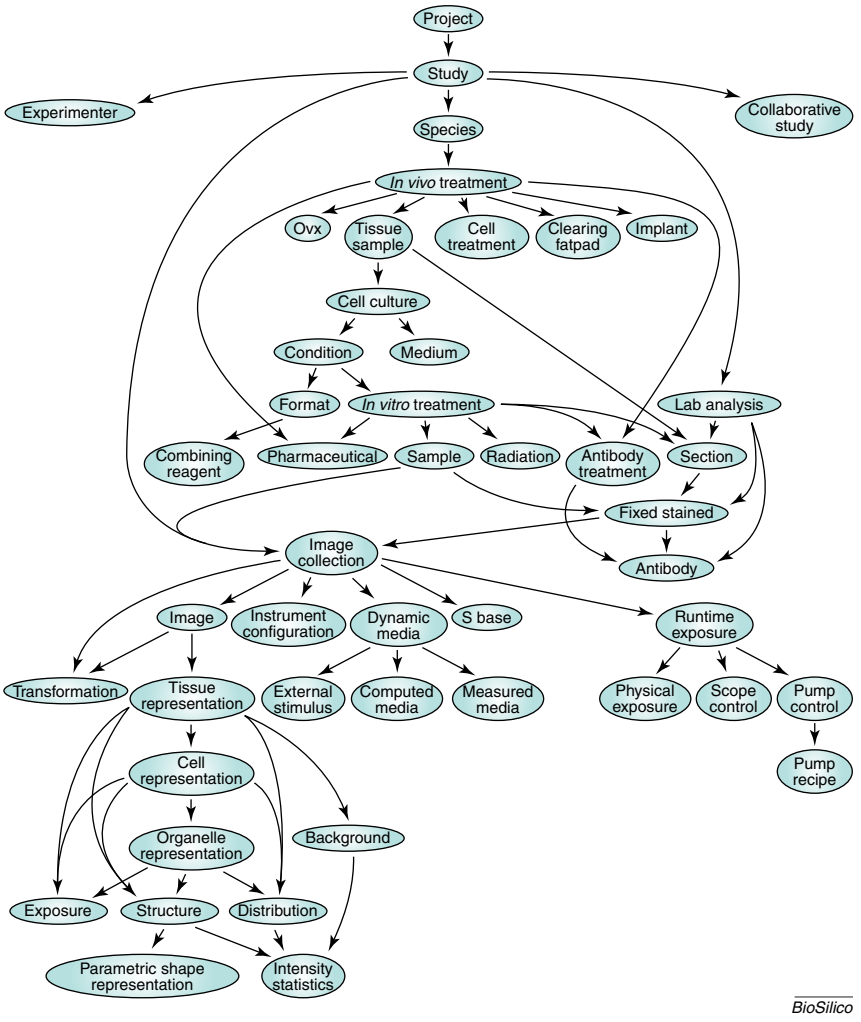

Figure 1. Data model for capturing experimental variables, images and quantitative results. Each block represents a class, and the edge shows the navigation path between classes. The model is designed to capture quantitative information at the image scale (either tissue or an ensemble of cells), cell scale (cell representation) and organelle scale (organelle representation). This model links computed physiological responses of each organelle of each cell to their external stimuli (runtime exposure). The link to system biology markup language (SBML) is made through the collection of time-series images for each experiment. The model is designed to accommodate studies of both cultured cells and multicellular systems.

correspond to feature-based representation of time-lapsed microscopy images. Such an annotation is hierarchical for representing information at image scale (all cells under observation), cell scale and organelle scale.

The presentation manager supports features for browsing the database and visualizing the result of a query function. Browsing of the database is performed against a pre-defined schema that captures annotation data, images and corresponding features. The data model in Fig. 1 is represented in XSD (XML schema), and the presentation manager constructs a view into the database using this representation and the corresponding style sheets (XSL) for browsing and updating. In this context, hardwiring of a graphical user interface (GUI) is bypassed in favor of a more flexible and dynamically generated user interface.

The query manager provides a set of predefined operators and templates to assist in information visualization and hypotheses testing. These operators help to draw contrast between computed features and their corresponding annotation data, and compute a variety of statistical measures such as analysis of variance and principal component analysis. These templates translate a query into a Java program that manipulates the database to retrieve required information. Through its deep-fetch mechanism, the object-orientated database simplifies sensitivity analyses, such as the analysis of variance, because each computed feature has to be mapped to its source (e.g. a tissue section or cell line). An example of such a high-level operator includes correlation of a particular computed feature with respect to an independent variable. For example, what is the correlation between physiological responses of two similar cell specimens that have been incubated with the same fluorescent compound at different concentrations? In this case, the physiological response could be described as the washout curve of the fluorescent compound from a specific subcellular compartment. The retention of fluorescent compounds has been used as an indicator of multidrug resistance in cancer cells [14].

Automated analysis

Quantitative analysis and change-detection at the subcellular level is an important process that can lead to a detailed understanding of physiological responses as a function of micro-environmental perturbations or genetic differences. The crucial issue is to delineate each subcellular region through a process known as segmentation. There are several ambiguities associated with correct delineation of objects of interest, which emanate from noise, and touching compartments (e.g. touching nuclei or mitochondria). In our system, noise is detected, then removed and interpolated from the image, and geometric constraints are

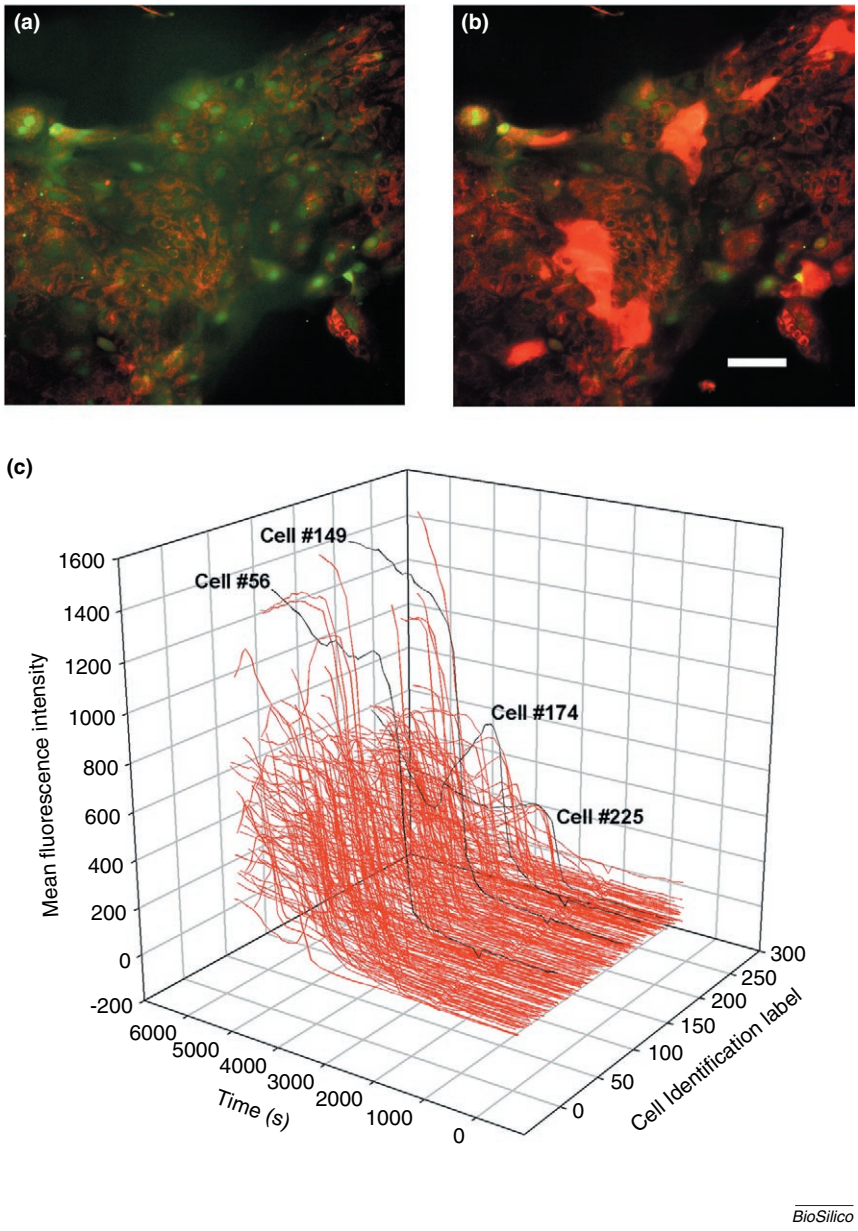

Figure 2. Time-dependent response of 225 living human mammary epithelial cells (HMEC 184B5 cells) after perfusion (stimulation) with 2 μM thapsigargin (473 s; first stimulation) and 5 μM ionomycin (3941 s; second stimulation). Both agents cause an increase in intracellular calcium that is sufficient to induce mitochondrial membrane depolarization and subsequent apoptosis [8]. **(a)** Cells before stimulation stained with FURA-PE3 (a calcium-sensitive dye for ratiometric imaging; green channel) and LDS-751 (an RNA and DNA binding dye that initially localizes in cellular mitochondria; red channel). **(b)** Cells at a later time-point (2645 s; between the first and second stimulation) when LDS-751 fluorescence intensity has increased owing to mitochondrial membrane depolarization. Scale bar, 100 mm. **(c)** 3D graph of all cell responses (indicated by increases in LDS-751 fluorescence intensity) in the red channel shows both weaker, early responses to the first stimulation (cells 225 and 174; black response lines), and strong, later responses to the second stimulation (cells 56 and 149; black response lines). Cells were also stained in the blue channel (not shown) with the cell-permeable, DNA-specific dye Hoechst 33342. The nuclei of individual cells were detected and segmented in this channel using automated image analysis [15]. Increases in red channel mean fluorescence intensities in the nuclear cell compartment are plotted in the 3D graph.

applied to delineate regions of interest based on self-similar properties. Details of the nuclear segmentation algorithm are available [15], and an example of application of this technique for capturing subcellular physiological responses is shown in Figure 2.

Concluding remarks

Initial applications of the BioSig and VSOM technology have been implemented in the fields of fluorescence cellular assay development [14] and antisense mRNA imaging agent development [16–23]. We are currently using multidrug resistance protein inhibitors to generate unique cellular responses (representing calcein accumulation and retention). In addition, we are using VSOM to develop dual-labeled (fluorescent and radiolabeled) antisense compounds and delivery vehicles for imaging mRNA in living cells, animals and humans.

Acknowledgements

Our research is funded through: (1) The Office of Biological and Environmental Research, Molecular Nuclear Medicine Program 'Imaging Gene Expression in Health and Disease,' in the Office of Science, US Department of Energy, under contract DE-ACO3-76SF00098; and (2) the Department of Defense, Breast Cancer Research Program, under the US Army Medical Research and Materiel Command, contract number DAMD17-98-1-8177. We would also like to thank Gerald Fontenay and Qing Yang of our development team. The publication number is LBNL-51445.

References

1 Chong, L. and Ray, L.B. (2002) Whole-istic biology. *Science* 295, 1661
2 Kitano, H. (2002) Systems biology: a brief overview. *Science* 295, 1662–1664
3 Csete, M.E. and Doyle, J.C. (2002) Reverse engineering of biological complexity. *Science* 295, 1664–1669
4 Davidson, E.H. (2002) A genomic regulatory network for development. *Science* 295, 1669–1678

5 Noble, D. (2002) Modeling the heart – from genes to cells to the whole organ. *Science* 295, 1678–1682
6 Parvin, B. *et al.* (2002) Biosig: an imaging bioinformatic system for studying phenomics. *IEEE Comput.* 35, 65–71
7 Taylor, D.L. *et al.* (1994) Potential of machine-vision light microscopy in toxicologic pathology. *Toxicol. Pathol.* 22, 145–159
8 Plymale, D. R. *et al.* (1999) Monitoring simultaneous subcellular events *in vitro* by means of coherent multiprobe fluorescence. *Nat. Med.* 5, 351-355
9 Callahan, D.E. and Parvin, B. (2001) Visual-Servoing Optical Microscopy. WIPO-PCT Patent Application WO 01 94528 A2
10 Hoffman, C. M. *et al.* (1994) SOFT LAB – a virtual laboratory for computational science. *Math. Comput. Simulat.* 36, 479–491
11 Parvin, B. *et al.* (2001) Declarative flow control for distributed instrumentation. *Proc. First IEEE/ACM Int. Symposium on Cluster Computing and the Grid* (May 2001; Brisbane, Australia) 48–55
12 Parvin, B. *et al.* (1997) Visual servoing for on-line facilities. *IEEE Comput.* 30, 56–62
13 Parvin, B. *et al.* (1997) Method and apparatus for accurately manipulating an object during microelectrophoresis. US Patent 5671086
14 Legrand, O. *et al.* (1998) Pgp and MRP activities using calcein-AM are prognostic factors in adult acute myeloid leukemia patients. *Blood* 91, 4480–4488
15 Yang, Q. and Parvin, B. (2002) Harmonic cut and regularized centroid transform for localization of subcellular structures. *Proc. Int. Conf. Pattern Recog,* (August 2002; Quebec, Canada) IEEE Press
16 Matsuo, T. (1998) *In situ* visualization of messenger RNA for basic fibroblast growth factor in living cells. *Biochim. Biophys. Acta* 1379, 178–184
17 Sokol, D.L. *et al.* (1998) Real time detection of DNA–RNA hybridization in living cells. *Proc. Natl. Acad. Sci. U. S. A.* 95, 11538–11543
18 Tavitian, B. *et al.* (1998) *In vivo* imaging of oligonucleotides with positron emission tomography. *Nat. Med.* 4, 467–471
19 Tavitian, B. *et al.* (2002) Characterization of a synthetic anionic vector for oligonucleotide delivery using *in vivo* whole body dynamic imaging. *Pharm. Res.* 19, 367–376
20 Thaden, J. and P.S. Miller (1993) Photoaffinity behavior of a conjugate of oligonucleoside methylphosphonate, rhodamine, and psoralen in the presence of complementary oligonucleotides. *Bioconjugate Chem.* 4, 386–394
21 Kuhnast, B. *et al.* (2002). Fluorine-18 labeling of peptide nucleic acids. *J. Label. Comp. Radiopharm.* 45, 1–11
22 Kuhnast, B. *et al.* (2000) General method to label antisense oligonucleotides with radioactive halogens for pharmacological and imaging studies. *Bioconjugate Chem.* 11, 627–636
23 Kuhnast, B. *et al.* (2000) Fluorine-18 labeling of oligonucleotides bearing chemically-modified ribose-phosphate backbones. *J. Label. Comp. Radiopharm.* 43, 837–848