DOE/SC-0066

GENOMES to LIFE

BIOLOGICAL SOLUTIONS FOR ENERGY CHALLENGES

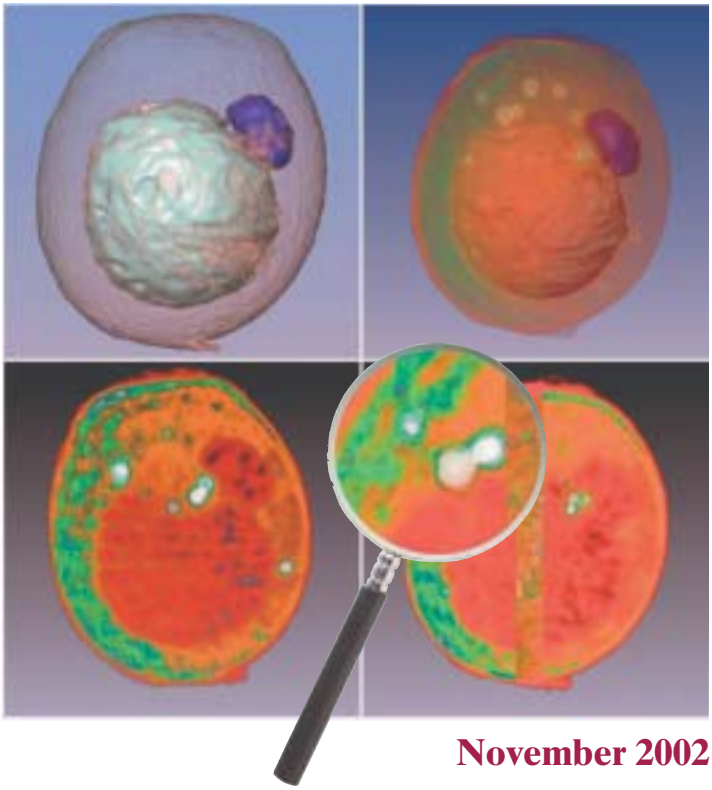
U.S. DEPARTMENT OF ENERGY

INNOVATIVE APPROACHES ALONG UNCONVENTIONAL PATHS

DOEGenomesToLife.org/technology

Report on the Imaging Workshop for the Genomes to Life Program

April 16–18, 2002



Sponsored by the
Office of Advanced Scientific Computing Research
and
Office of Biological and Environmental Research
of the
U.S. Department of Energy
Office of Science

November 2002

Genomes to Life Program

Contacts for program information

Marvin Frazier
U.S. Department of Energy (SC-72)
Office of Biological and Environmental Research
301/903-5468, Fax: 301/903-8521
marvin.frazier@science.doe.gov

Gary Johnson
U.S. Department of Energy (SC-30)
Office of Advanced Scientific Computing Research
301/903-5800, Fax: 301/903-7774
garyj@er.doe.gov

Publications

Documents, meeting reports, and image gallery via the Web:

- http://DOEGenomesToLife.org

Copies of this report and other Genomes to Life publications:

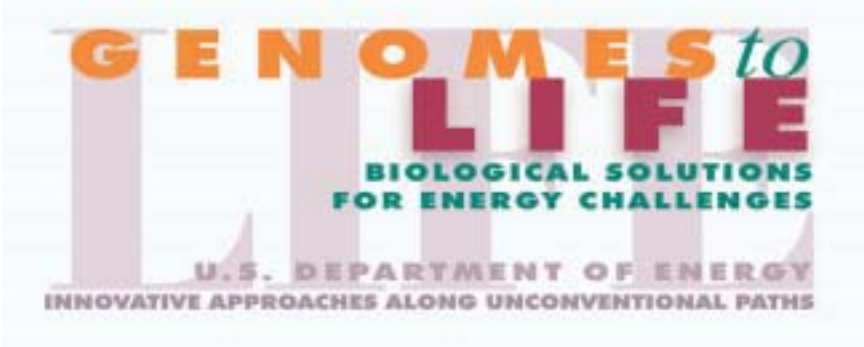
- Human Genome Management Information System
  Oak Ridge National Laboratory
  1060 Commerce Park, MS 6480
  Oak Ridge, TN 37830

  865/576-6669, Fax: 865/574-9888, mansfieldbk@ornl.gov

Electronic versions of this report and additional information:

- http://DOEGenomesToLife.org/technology

DOE/SC-0066



DOEGenomesToLife.org/technology

Report on the Imaging Workshop for the Genomes to Life Program

Charlotte, North Carolina
April 16–18, 2002

Co-Chairs
Steven Colson, Pacific Northwest National Laboratory
Damir Sudar, Lawrence Berkeley National Laboratory

Report prepared by the
Office of Advanced Scientific Computing Research
and
Office of Biological and Environmental Research
of the
U.S. Department of Energy
Office of Science

November 2002

Table of Contents

List of Figures v

Preface vii

Executive Summary ix

Section 1. Molecular Machines: Protein Complexes 1

Issues and Limitations 1
Required Instruments and Methods 2
Protein Identification and Characterization 3
Imaging Technology Overview 3
Cross-Cutting Needs 9
Protein or Sample Characterization 9

Section 2. Intracellular and Cellular Structure, Function, and Processes 13

Understanding the Genetic Basis for Microbial Function 13
Issues and Limitations 13
Recommendations 14
General Requirements 14
Specific Needs 14
Current Methodologies: Status and Development Needs 15

Section 3. Monoclonal and Heterogeneous Multicellular Systems, Cell-Cell Signaling, and Model Systems 21

Why Imaging is Needed 22
Imaging Issues 22
Improved Model Systems for Microbial Imaging 22
Interrogation of Model Systems 23
Required Instrumentation and Data Analysis 24

Section 4. Imaging Microbial Communities 25

Issues 25
Studying Microbial Communities 25
Challenges 26

Section 5.Computational Infrastructure for Imaging 31
Issues 31
Technology Needs 32
Section 6. Probe Development for Advanced Imaging Methods 39
Applications for Fluorescent Probes 39
Philosophy of Probe Development 41
Fluorescent Probe Design: Adaptation of Fluorophore To Create a Probe 45
Appendices 49
Appendix A: Workshop Attendees 51
Appendix B: Workshop Agenda 55
Appendix C: Imaging Methodologies for Genomes to Life Research 59
Appendix D: National Laboratory Capabilities and Imaging Technologies 69
Argonne National Laboratory (ANL) 70
Brookhaven National Laboratory (BNL) 70
Lawrence Berkeley National Laboratory (LBNL) 72
Lawrence Livermore National Laboratory (LLNL) 73
Los Alamos National Laboratory (LANL) 75
Oak Ridge National Laboratory (ORNL) 75
Pacific Northwest National Laboratory (PNNL) 77
Sandia National Laboratories (SNL) 79
Thomas Jefferson National Accelerator Facility 80

List of Images

Fig. ES.1. Confocal laser micrograph of a bacterial microcolony in a river biofilm community vi

Fig. ES.2. Synchrotron infrared images showing a small colony of natural bacteria vii

Fig. 1.1. One model system 1

Fig. 1.2. Imaging at multiple-length scales using different techniques to determine the function of molecular machines 2

Fig. 1.3. *Magnetospirillum magnetotacticum* 3

Fig. 1.4. Three-dimensional structural determination of single rubisco molecules with a simulated X-FEL, and direct-phase retrieval by oversampling technique 4

Fig. 1.5. A 3D reconstruction from electronmicrographs of flash-frozen specimens of the Ross River virus (left) and the Dengue 2 virus 4

Fig. 1.6. Helical structure of P-pili from *Escherichia coli*, showing evidence from X-ray fiber diffractions and scanning-transmission electron microscopy 5

Fig. 1.7. Cryo-X-ray microscopy of 3T3 fibroblast whole cells with no fixatives, stains, or contrast enhancement 5

Fig. 1.8. Section from a 3D electron tomographic reconstruction of *Caulobacter* 6

Fig. 1.9. Atomic force microscopy providing structural information on membrane proteins via imaging and singe-molecule force curves of protein unfolding 7

Fig. 1.10. Top: Two-channel fluorescence images (20 μm by 20 μm) of individual donor- or acceptor-labeled T4 lysozyme molecules tethered by cross-linker molecules to the hydrocarbon-modified glass surface of a cover slip under a pH 7.2 aqueous buffer solution 8

Fig. 1.10. Bottom: The crystal structure of wild-type T4 lysozyme with two covalently labeled dye molecules (fluorescence resonant energy transfer donor and acceptor: tetramethylrhodamine and Texas Red) 8

Fig. 2.1. Three-dimensional soft X-ray reconstruction of yeast 15

Fig. 2.2. Soft X-ray tomography 16

Fig. 2.3. X-ray fluorescence image of *Mycobacterium avium*, 24 h after infection of a fixed and air-dried macrophage 16

Fig. 2.4. (a) Soft X-ray micrograph and (b) X-ray fluorescence map of marine ciliate showing the cell's C content 17

Fig. 3.1. Relevant level of complexity inherent in microbial soil ecosystem 21

Fig. 3.2. Visualization of cell-cell communication in living cells 22

Fig. 4.1. Synchrotron infrared images showing a small colony of natural bacteria 27

Fig. 4.2. One-photon and two-photon images of the same location and depth, showing a laboratory-grown biofilm stained with a nucleic acid–specific fluorochrome 28

Fig. 5.1. Scales of imaging modalities 31

Fig. 5.3. Control cells (green) and engineered autocrine cells (orange) mixed overnight and followed for 4 h by two-color fluorescence time-lapse microscopy 35
Fig. 5.4. Computational biology 3D reconstruction of a mitochondrian crista 36
Fig. 5.5. Example of 3D volume visualization of a tomographic magnetic resonance imaging data set of a *Xenopus laevis* frog from oocyte after heat shock 36
Fig. 5.6. Automated uptake analysis for a population of cells: (a) image and corresponding data model, (b) active page with segmented result, and (c) time-series response of a selected cell from (b) 37
Fig. 5.7. An example of remote instrumentation for shared collaborative use—the DeepView architecture for remote microscopy funded by MICS 37
Fig. 5.8. Conceptual vision of the infrastructure for imaging 38
Fig. 6.1. FISH using multiple fluorescently labeled DNA probes 39
Fig. 6.2. Intracellular calcium in smooth muscle cells imaged using a cytosolic fluorescent calcium indicator Fura-2 (A) and the near-membrane calcium indicator FFP18 (B) 40
Fig. 6.3. (a) Fluorescence of Di-8-ANEPPS in lipid vesicles from excitation at two alternative wavelengths (440 nm and 530 nm). (b) The ratio of emission intensity is linearly proportional to membrane potential 40
Fig. 6.4. Localizing sites of receptor activation by a juxtacrine ligand 41
Fig. 6.5. A cytotoxic T cell (CTL) engaging an antigenic target immunostained for tubulin (green) and LFA-1 (red) 41
Fig. 6.6. Multicolor staining for different proteins (CD8, CD4, and HLA) allowing different cells to be distinguished in a mixed population 42
Fig. 6.7. Structures of a number of common fluorophores 42
Fig. 6.8. Series of cyanine dyes with an increasing number of conjugated double bonds 43
Fig. 6.9. Excitation and emission spectra shown for Cy3 through Cy7 43
Fig. 6.10. Subsurface cancer shown as more easily detected by the antibody-conjugated Cy7 dye than other Cy dyes because the emission properties better match the tissue's optical window ... 44
Fig. 6. 11. Photostability of a series of cyanine dyes 44
Fig. 6.12. Large spectral shifts (~260 nm) possible with lanthanide chelates because absorption is due to antennae attached to the macrocycle holding the metal, while metal ions provide fluorescence 44
Fig. 6.13. The fluorescein arsenide FLASH-EDT2 specifically labeling proteins containing the tetracysteine sequence CCXXCC 45
Fig. 6.14. A single optical section taken by two-photon excitation imaging through an intact pancreatic islet (~100-μm diameter) loaded with Fura-2/AM 45
Fig. 6.15. A single optical section of a fixed and permeabalized pancreatic islet 45
Fig. 6.16. Dyes that are excited at a single source but have emission at different, easily resolved wavelengths 46
Fig. 6.17. Plasmon-resonant particles useful for monitoring dynamics over long time scales 46
Fig. 6.18. STEM images of freeze-dried unstained phosphorylase kinase that has undergone exchange 47

Preface

This report is a result of the Imaging Workshop for the Genomes to Life (GTL) program held April 16–19, 2002, in Charlotte, North Carolina. The meeting was sponsored by the Office of Biological and Environmental Research and the Office of Advanced Scientific Computing Research of the U.S. Department of Energy's (DOE) Office of Science.

The purpose of the workshop was to project a broad vision for future needs and determine the value of imaging to GTL program research. The workshop included four technical sessions with plenary lectures on biology and technology perspectives and technical presentations on needs and approaches as they related to the following areas of the GTL program:

1. Molecular machines (protein complexes)
2. Intracellular and cellular structure, function, and processes
3. Multicellular: Monoclonal and heterogeneous multicellular systems, cell-cell signaling, and model systems
4. Cells in situ and in vivo: Bacteria in the natural environment, microenvironment, and in vivo systems

More than 60 attended the workshop, including scientists from most of the DOE national laboratories and partnering universities and GTL program managers. (A list of the steering committee and DOE advisors is included in Appendix A and the workshop agenda in Appendix B.) Participants were divided into seven writing teams that met after the technical sessions to draft this report, which was originally divided into the following sections:

- Executive Summary
- Molecular Engines
- Intracellular and Cellular
- Multicellular
- Cells In Vivo
- Image Data Computational Infrastructure
- Probe Development

The sections were compiled and edited at Pacific Northwest National Laboratory. Oak Ridge National Laboratory did the final editing and prepared the report for publication. Oak Ridge Institute for Science and Education made all arrangements for the workshop.

Executive Summary

Thought is impossible without an image.
— Aristotle, 325 B.C.E.

The overall goal of the U.S. Department of Energy's (DOE) Genomes to Life (GTL) program is to understand the composition and function of the biochemical networks and pathways that carry out the essential processes of living microbial organisms. Such understanding is essential for DOE to more effectively address its missions in energy security, carbon management, and environmental cleanup. Imaging of microbial organisms is an essential enabling component of GTL because it provides a method for linking genomic information to function. Imaging aids understanding of how cell function changes with time and environment. Innovations in imaging coupled with computational advances will accelerate scientific discovery and enable biological solutions to energy challenges.

GTL has four main goals:

Goal 1: Identify and characterize the molecular machines of life—the multiprotein complexes that execute cellular functions and govern cell form

Goal 2: Characterize gene regulatory networks

Goal 3: Characterize the functional repertoire of complex microbial communities in their natural environments at the molecular level

Goal 4: Develop the computational methods and capabilities to advance understanding of complex biological systems and predict their behavior

What Imaging Needs To Provide

Current imaging techniques display a wealth of information about eukaryotic (e.g., human) biological systems over a wide range of length and time scales. Imaging of these systems has led to significant advances in understanding cell function and complex cellular systems (see Fig. ES.1). Microbial systems with their smaller cells, however, present different challenges. New techniques are needed to connect genomic information with microbial functions spatially and temporally in model systems and in their natural environments. These new techniques will drive further advances in all fields of biology.

Imaging and the Molecular Machines of Life (GTL Goal 1)

Imaging will contribute directly to identifying and characterizing the molecular machines of life, giving a deeper understanding of their relationships. Imaging will help define interactions among proteins and other cellular components in the complex interacting networks of living cells. A real-time, molecular-scale description of protein interactions will reveal metabolic relationships that can be engineered to accomplish DOE missions. High-throughput methods

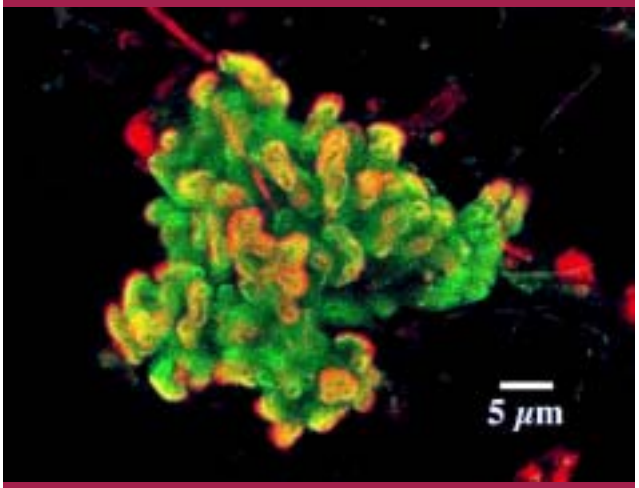


Fig. ES.1. Confocal laser micrograph of a bacterial microcolony in a river biofilm community. The colony is stained with the electrical-potential–sensitive fluorescent stain JCI; orange regions are areas of high potential. [Source: John R. Lawrence, National Water Research Institute, Canada]

(e.g., mass spectrometry) for characterizing protein complexes require validation of the existence and function of these complexes in living cells. Imaging can provide that validation.

New imaging methods will be required to define the state of a biological system in response to differing environmental conditions and enable the functional interpretation of traditional analyses of protein complexes. Imaging provides a direct link among the genomes of microorganisms and the atomic structures of the molecular machines that define their functions. Direct observations of protein complexes that comprise these machines will, in turn, provide an important link between genomic information and living cell function. Realizing these potentials will require innovative probes to visualize the structure and dynamics of molecular machines and to locate specific proteins. Substantial innovation will be needed to develop spectroscopies that enable measurements of dynamics (function); microscopies with sufficient spatial resolution and sensitivity to image individual proteins; methods to resolve their atomic structures; and computational methods to acquire, store, access, visualize, and interpret results (see GTL Goal 4).

Imaging to Characterize Gene Regulatory Networks (GTL Goal 2)

Imaging the location of regulatory proteins in vivo to identify their binding sites in DNA or other cellular structures is needed to understand the primary step of complex gene regulatory networks. The identities of most of these regulatory proteins are unknown. Methods must be developed to interrogate DNA-protein bound pairs on very rapid time scales during the cycles of cell growth and function.

To understand the functions of gene regulatory networks, the dynamic timing of gene expression needs to be known as a function of cell cycle and stimulating signals. This requires development of small fluorescent expression tags that can be imaged without delay in vivo. In particular, we must know the temporal sequence for expression and intracellular distribution of the regulatory proteins themselves to design computational models of the networks.

Knowledge of gene regulatory networks for both microorganisms and eukaryotic systems and imaging of the expression schedule and subsequent distribution of each gene product can provide a basis for understanding the molecular machines of life and their coordinated function in complex microbial communities in natural environments. The gene regulatory networks act as a digital (molecule-by-molecule) computer to specify the identity and level of expression of target genes. Computer models must be developed to enable broad interpretation of experimental results, leading to predictions of biological function (see GTL Goal 4).

Imaging to Characterize Complex Microbial Communities in Model and Natural Environments at the Molecular Level (GTL Goal 3)

The past decade's advances in techniques for imaging living biological material have revealed that microbial communities are dynamic structured assemblages with compartmentalized (e.g., metabolic) activities. Imaging methods with multiple modalities enable interrogation of these spatially and temporally organized features in a multiplexed manner. Then, to understand life in naturally occurring communities, we must know the identities of the constituent species, their functions in the community, how they perform these functions, and how the communities change over space and time. Attaining these objectives will require development of an advanced suite of probes, imaging devices, and computational methods (see Goal 4). Understanding the complexity of natural systems will require direct study coupled with research on better-defined model systems (see Fig. ES.2).

Imaging and Developing Computational Methods and Capabilities (GTL Goal 4)

Once images are acquired, new data-processing methods will be required to access and manage image data, enable visualization, and make possible the quantitative analysis of biological systems and their components. Such processing will enable development of predictive computer models that will be essential for addressing DOE missions.

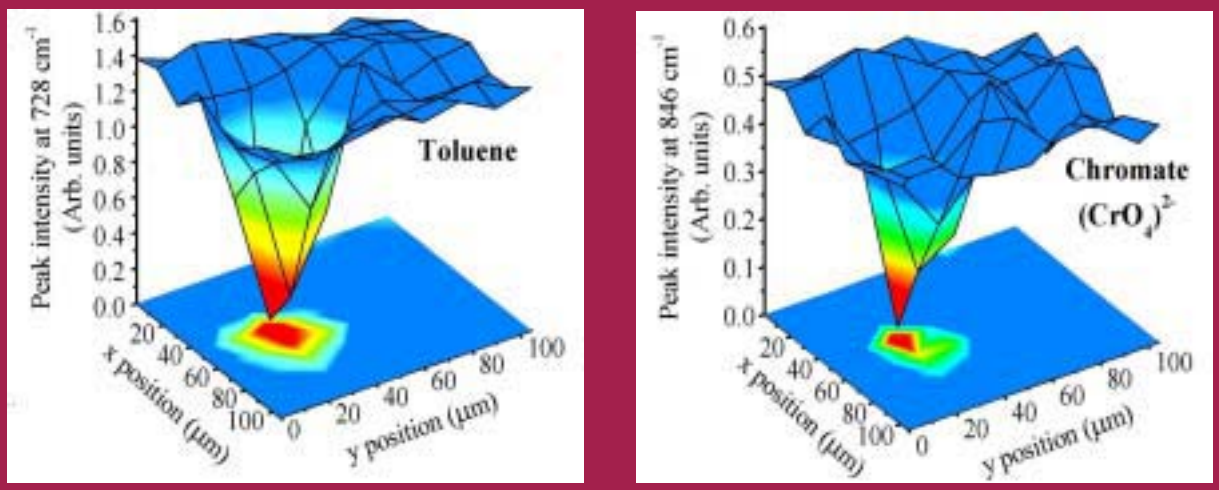


Fig. ES.2. Synchrotron infrared images showing a small colony of natural bacteria. A common organic contaminant, toluene, was used to accelerate the reduction of a carcinogenic form of chromium (chromate CrO_4^{-2}) to its environmentally safe form. [Source: H.-Y. Holman et al., "Real-Time Characterization of Biogeochemical Reduction of Cr (VI) on Basalt Surface by SR-FTIR Imaging," *Geomicrobiology Journal* **16**(4), 307–24 (©1999). Reproduced by permission of Taylor & Francis, Inc. (www.routledge-ny.com)]

Recommended Investment Strategy and Timeline for Accomplishments

We recommend a GTL imaging program that integrates technical approaches and biological needs. This program should draw on existing capabilities from other disciplines and advance existing methods to address the unique requirements of the GTL program. It should immediately initiate research toward the most significant imaging challenges, including innovative research with high risks and high potential payoffs. The GTL imaging program should include single- and multi-investigator projects and multi-institutional research programs, funds to develop and maintain essential capabilities at DOE national laboratories, and investments in education. Bringing together scientists from different backgrounds who will create interdisciplinary approaches to important technical challenges will be critical.

Expected Time Frame for Accomplishments

- Within 2 to 3 years: Capture, integrate, extend, and apply existing imaging technologies to microbes and microbial communities.
- Within 5 years: Identify fundamental bottlenecks and limitations on reaching the potential of existing imaging technologies; develop capabilities for initial conversion of these potentials to practice; develop a pathway for addressing the most challenging potentials.
- Within 20 years: Develop technologies to quantify functions of machines, cells, and communities in real time and in situ with minimal perturbation to the systems; apply these technologies to relevant organisms and establish predictive models.

Impact of Imaging on GTL Goals and DOE Missions

Coupled with computational modeling and the use of genomic, proteomic, and related analytical information, imaging will quantify functions of machines, cells, and communities, which in turn will enable the use of microbial systems to solve problems in DOE mission areas. Visualization and quantitative image analysis of biological systems and their components provide a level of understanding of complex systems that cannot be obtained in any other manner.

1. Molecular Machines: Protein Complexes

SECTION 1

The human genome is a sophisticated and complex coding system for the formation of thousands of different proteins. Complex assemblies of these proteins, the molecular machines of the cell, perform the majority of cellular functions, from enzymatic activity to cell motility. Goal 1 of the Genomes to Life (GTL) program is to "identify and characterize the molecular machines present in the microbial cell." This goal is intricately linked to the ultimate GTL program goal: to model cells and cell systems using a predictive, computational biology approach for understanding cellular behavior. A key to developing these modeling capabilities is the ability to produce comprehensive functional characterization of gene products acting in concert in large, multiprotein molecular machines. Understanding the constituents and functions of the machines is critical to the overall goals of the program (Figs.1.1a and 1.1b).

A comprehensive approach to molecular machines requires the identification of all molecular machines in the cell, their protein constituents, the interactions that bind the constituents together, and how these machines are assembled, regulated by their environment, and perform their function. This effort will be aided by comparative genomics among species to furnish information critical to the identification of novel molecular machines. The size of the task should not be underestimated. The active and existing complexes in any given cell depend on (1) the genome, (2) the cellular environment, and (3) time. Key to progress will be identification of the complexes most important to understanding cell function. Imaging will play a key role in this process.

Issues and Limitations

Proteins and other molecules that comprise molecular machines may be identified through the mapping of protein-protein and protein-ligand

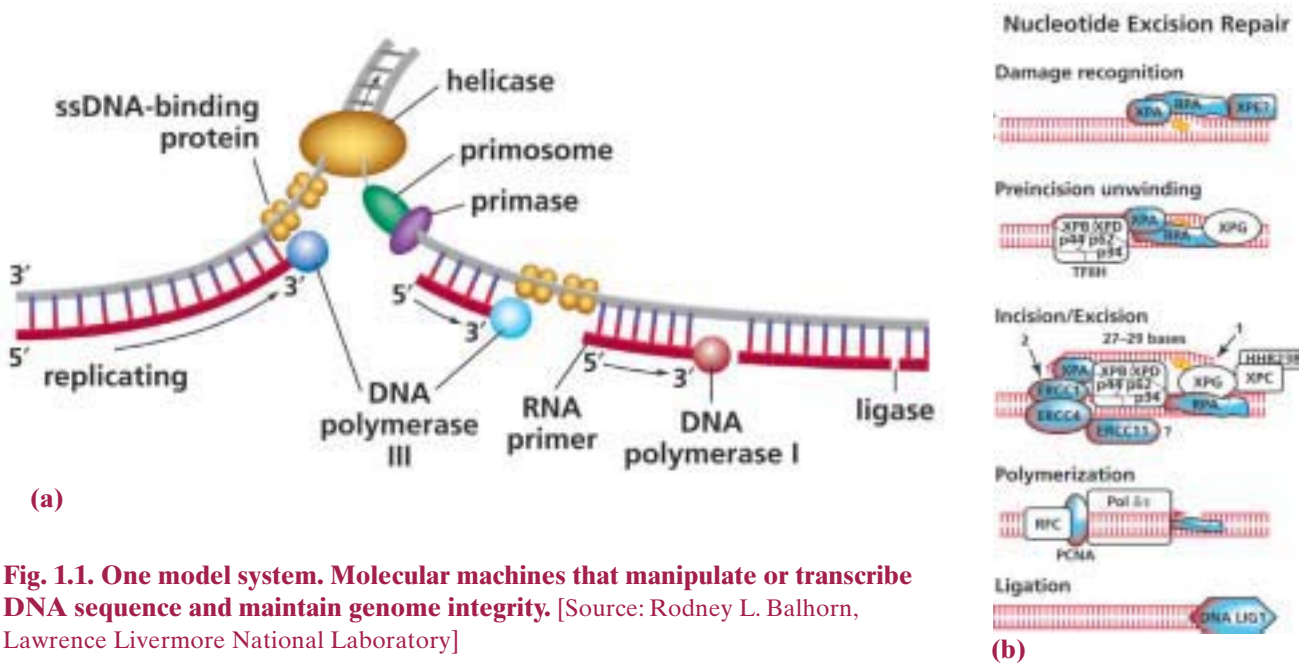


Fig. 1.1. One model system. Molecular machines that manipulate or transcribe DNA sequence and maintain genome integrity. [Source: Rodney L. Balhorn, Lawrence Livermore National Laboratory]

interactions using phage display, two-hybrid techniques, and direct isolation of protein complexes. These methods allow independent enumerations of intermolecular interactions, and correlations derived by comparing results from different functions will provide significant information for elucidating signaling pathways and the many other functions performed by molecular machines; each method, however, has intrinsic biological limitations. These methods need optimizing (by other parts of the GTL program) so they can deliver a more rapid identification of the constituents of molecular machines. This is a significant need that will directly support future imaging efforts.

Required Instruments and Methods

A molecular-level understanding of a molecular machine's function requires imaging of the machine at various levels of detail using techniques that measure many variables (e.g., chemistry and structure) and enable collection of dynamic information on numerous time scales. Imaging of individual proteins and their complexes will be required to

- Provide a molecular-level understanding of their function through direct observations,
- Test hypotheses about their function and regulation through designed mutations, and
- Develop probes that will enable the visualization and tracking of individual molecular machines inside living cells.

In general, this will require the coordinated use of multiple imaging techniques. The goal will be to monitor structure, function, stimuli, and response in real time; for example, the onset of a change in calcium ion concentration while simultaneously observing calcium-induced changes in molecular structure.

Ultimately, the molecular machine must be imaged in the context of its cellular environment to understand where and when it functions and to identify other cellular components or systems that may interact with it during its functioning. As a starting point, the intact machine must be imaged when isolated from other cellular components to define the machine's extent. Machine component positions may be determined by using affinity tags (such as antibodies or antibody domains) specific to individual components. Machine components (e.g., constituent proteins) may need to be imaged individually at very high resolution. Images of dynamic activities, which can provide considerable insight into the machine's function and molecular mechanisms, should be obtained wherever possible.

Imaging at the whole-cell level is discussed in detail in other parts of this report but may be carried out, for example, using optical microscopy (e.g., fluorescence microscopy, Fig. 1.2), 2D or 3D confocal microscopy, X-ray microscopy (2D or 3D X-ray tomography), or electron microscopy (2D or 3D electron tomography). Optical microscopy provides images of living cells and enables the collection of time-domain information. Three-dimensional scanning multiphoton microscopy can provide 3D images of the cell's dynamic behavior. However, significant advances in its time resolution would greatly advance our ability to study the activity of living cells in great detail.

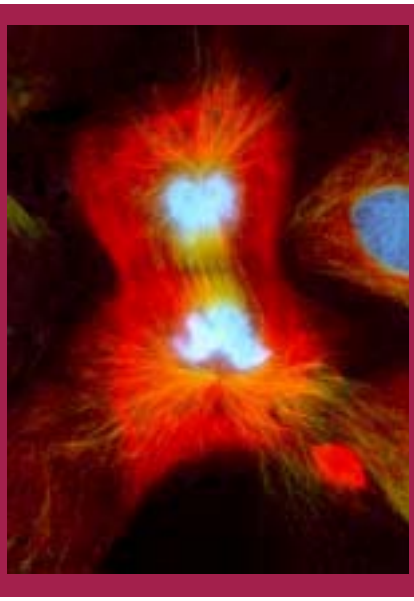


Fig. 1.2. Imaging at multiple-length scales using different techniques to determine the function of molecular machines. This triple-exposure fluorescent micrograph is of a newt lung cell in early telophase of mitosis. The red stain is keratin, and the yellow is tubulin; blue depicts chromosomes. [Image courtesy of Conly L. Rieder; Division of Molecular Medicine; Wadsworth Center; New York State Department of Health; Albany, New York]

Protein Identification and Characterization

Protein identification and characterization are currently slow, labor-intensive processes. New, more rapid methods are needed that will enable identification of the protein constituents comprising molecular machines, types and location of post-translational modifications, and ligands that participate in their function. More efficient protein-sequencing methods will enable identification of the protein, and, when coupled with mass spectrometry, the sites of post-translational modifications.

Current Methods

Affinity tags (e.g., antibodies) against a molecular machine will be needed to (1) enable isolation of the entire molecular machine, as well as its component proteins, and (2), when combined with imaging, enable localization of both the molecular machine within the cell and a particular constituent within the machine. Affinity tags can be produced using a variety of techniques, among them phage and yeast display; however, advances in high-throughput approaches are needed for wholesale production of affinity tags for the cell's many molecular machines.

New Approaches

In addition to developing new probes for tagging proteins in vivo, new methods are needed that will enable the tagging of proteins at a single site within their sequence in vivo with a chromophore, electron-dense material, or other suitable probes. This may include the development of new vectors or gene constructs that will result in the expression of proteins covalently attached to unique, modifiable functional groups (e.g., thioester) or relatively short peptide sequences that preferentially (and when possible, covalently) bind fluorescent ligands (e.g., metals, lanthanides, and particles) or contribute their own signal.

Imaging Technology Overview

A complete understanding of molecular machines will require information covering a wide range of dimensions and time scales. Structural data such as molecular dimensions, subunit structure, and overall molecule shape are typically obtained with high-resolution techniques at the angstrom to nanometer scale. Information on dynamics and function of molecular machines, however, is usually obtained on larger spatial and shorter time scales using methods that interrogate a molecular machine within its immediate environment. A variety of techniques can be used to access the structural and dynamic information of protein complexes. However, there is much room for improvement in several areas, including the possibility of exploiting new phenomena.

Short-Wavelength Techniques

Electrons and X rays both have wavelength ranges that can give diffraction-limited resolution in a range useful for directly imaging the spatial arrangement of subunits in isolated protein complexes. X-ray crystallography has provided atomic-resolution structures of proteins, their subunits, and small machines such as ribosomes. Electron microscopy (EM) is well established for viewing complexes at 0.5- to 2.0-nm resolution. This resolution is useful for direct structure determination or for positioning subunit structures from X ray into an overall atomic-resolution complex. Samples are rapidly frozen and viewed in thin amorphous ice. Particles are typically viewed by tomography either in 2D arrays or as isolated objects. If the sample is freeze-dried, quantitative measurement of electron scattering can give direct mass distribution useful for determining subunit stoichiometry or confirming the faithfulness of reassembly (Fig. 1.3).

Fig. 1.3 *Magnetospirillum magnetotacticum*. Electron microscopy of intact, frozen-hydrated bacterial cells with no stain or chemical fixatives. [Source: Kenneth H. Downing, Lawrence Berkeley National Laboratory]

Intact molecular machines composed of multiple proteins often are too large or too flexible to make a crystallographic approach feasible. Furthermore, to obtain any information about substructure, resolutions significantly greater than that of light microscopies are necessary. X-ray microscopy is promising for the future when improvements in optics and instrumentation will take full advantage of the 2.4-nm wavelength used for image formation to yield fine details of protein complexes. At present, this size is best suited to EM.

The imaging of flash-frozen specimens of molecular machines has huge potential for providing large amounts of structural and functional information about the operation of molecular machines. Flash-frozen specimens maintain their native conformation and can survive multiple exposures so that a partial tilt series of micrographs can be produced on a single specimen. Intermediate voltage electron micrographs (300 to 400 keV) are particularly well suited to this type of study because these relatively high energy electrons have a greater penetration depth and cause less radiation damage than the more typical 100-keV electrons used in most transmission electron microscopes (TEM). Three-dimensional reconstructions can be calculated from images of a molecular machine taken at many different angles (much like the production of 3D radiological images from a CAT scan). Images of affinity-tagged specimens give significant information about the relative positioning of different gene products within the machine (Figs. 1.4 and 1.5).

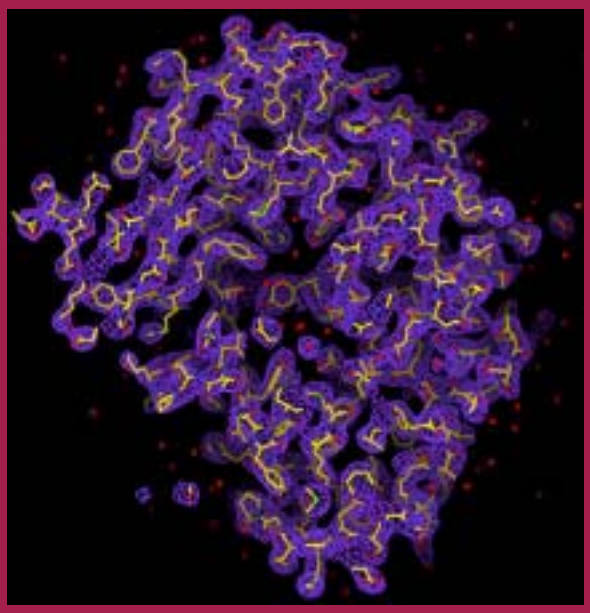


Fig. 1.4. Three-dimensional structural determination of single rubisco molecules with a simulated X-FEL and direct-phase retrieval by oversampling technique. The map shows the 3D electron-density map of the rubisco molecule (contoured at 2 sigma) reconstructed from the 3D diffraction pattern of 3×10^5 identical copies of the rubisco molecule. Poisson noise was added (Rl – 16.6%), and the central $3 \times 3 \times 3$ pixel intensity was removed. The edge of the simulated X-FEL diffraction pattern corresponds to a resolution of 2.5 Å. [Source: Jianwei Miao, Keith Hodgson, and David Sayre, "An Approach to Three-Dimensional Structures of Biomolecules by Using Single-Molecule Diffraction Images," *PNAS* **98**(12), 6641–45. ©2001 National Academy of Science, U.S.A.]

Theoretical studies of the number of images needed to produce very high resolution images of a molecular machine have been put forward by a number of investigators. Theoretically, a 3D reconstruction of a molecular machine can be produced at a level of resolution approaching that of X-ray crystallography.

This process is not instrument limited but rather computation limited. Estimates predict that an image with 4-Å resolution of the ribosome's large subunit could be computed from 500,000 particle images and an image with 3-Å resolution from 2 million particle images. Even though hundreds of particle images can be collected on one electron micrograph, with CCD detectors rather than film, the ability to process and recombine this

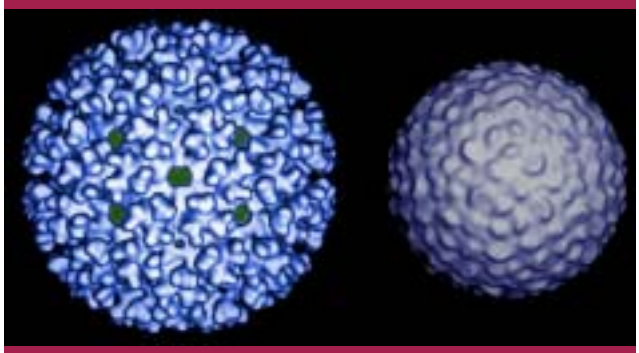


Fig. 1.5. A 3D reconstruction from electronmicrographs of flash-frozen specimens of the Ross River virus (left) and the Dengue 2 virus. [Source: Tim Baker and Richard Kuhn, Purdue University]

number of images is still far beyond the capabilities of existing algorithms and computational resources. The huge potential of this approach, however, could justify a significant investment in research directed toward achieving that capability.

Scanning transmission electron microscopy (STEM) provides information complementary to TEM techniques by showing a direct link to the biochemistry. Measurement of mass-per-unit length of helical complexes and total mass of other complexes supplies important information for determining the stoichiometry of the complexes and comparing native and reassembled structures. Use of metal clusters such as undecagold as labels for specific constituents makes STEM a valuable method for localizing gene products within complexes (Fig. 1.6).

Imaging molecular machines within cells can be achieved using soft X-ray (XM) or EM. High-resolution techniques such as XM can be carried out at atmospheric pressure on whole hydrated cells up to 10 μm thick and can obtain information at 25- to 30-nm resolution using existing optics.

Better resolution can be obtained using TEM of thin specimens placed in a vacuum chamber and held stationary. A promising method for preserving specimens for both XM and TEM is flash-freezing, in which the specimen is frozen so quickly that water molecules do not have time to crystallize but instead solidify into a vitreous ice. This well-developed technology enables cryo-X-ray tomography of whole cells yielding structural information at better than 50-nm resolution and electron tomography of isolated complexes and 2D arrays approaching 1-nm resolution (Fig. 1.7).

The resolution limit of electron tomography is not that of the microscope but is related to the obtainable number of distinct views of the cell. Because every cell is different, images from different cells cannot be averaged (as is typically done with images of molecular complexes). Instead, an entire tilt series must be carried out on a single, small cell (e.g., bacteria <0.5 μm in diameter) or on sections of cells. Micrographs resulting from the bacterium's tilt series or of all sections of a eukaryotic cell can then be recombined into a 3D reconstruction approaching 10 nm in resolution (Fig. 1.8). Methods for lowering radiation damage are needed to allow for larger numbers of images to be obtained on a single specimen. This may be possible using higher-voltage electrons (>500 keV) or by imaging specimens maintained at liquid helium temperatures.

Plastically Stressed Adhesion Pili

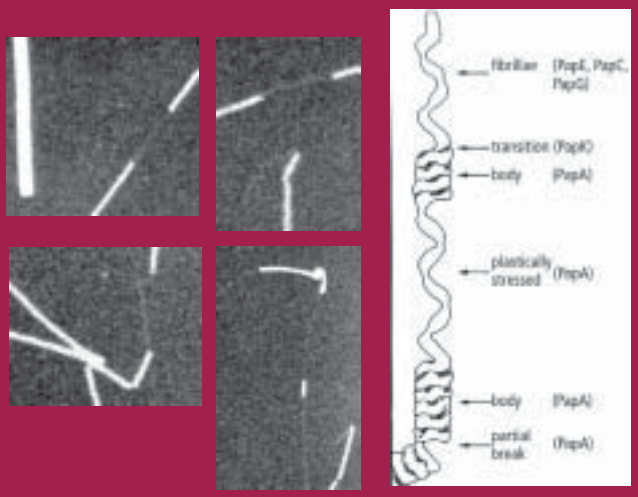


Fig. 1.6. Helical structure of P-pili from *Escherichia coli*, showing evidence from X-ray fiber diffraction and scanning-transmission electron microscopy. [Source: Minfang Gong and Lee Makowski, "Helical Structure of P- pili from *Escherichia coli*: Evidence from X-Ray Fiber Diffraction and Scanning-Transmissions Electron Microscopy," *Journal of Molecular Biology* **228**(3), 735–42 (1992). Reprinted by permission of the publisher Academic Press, an imprint of Elsevier Science]

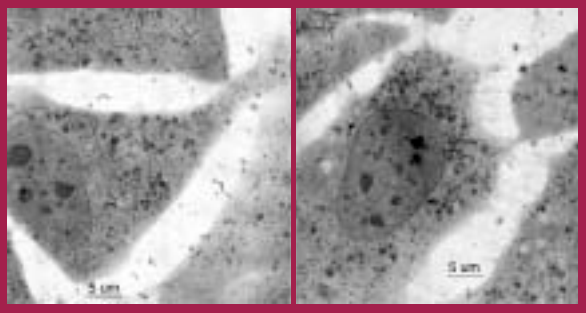


Fig. 1.7. Cryo-X-ray microscopy of 3T3 fibroblast whole cells with no fixatives, stains, or contrast enhancement. [Source: W. Meyer-Ilse et al., "High-Resolution Protein Localization Using Soft X-Ray Microscopy," *Journal of Microscopy* **201**, 395–403 (2001)]

Dynamic information on an individual specimen that cannot be generated with EM can be obtained by flash-freezing a specimen at various times after an applied stimulus. To study structural changes, for instance, Ken Taylor at Florida State University is flash-freezing a specimen of an insect flight muscle in the act of contracting.

Recently, electron microscopes using energy filters have come into use. One potential use of energy filters is the development of energy-specific tags to absorb electrons at one energy but not at others. These tags could be used in a manner analogous to that of green fluorescent protein in light microscopy—to localize tagged gene products within cells at EM resolution. Used with electron tomography, these tags could localize specific molecular machines to within 10 nm in the cellular environment.

Optical Microscopy

Single-molecule microscopy and spectroscopy can give direct insight into the mechanism, dynamics, and kinetics of molecular machines. These molecular complexes are dynamic in their behavior, lending themselves to study at the single-molecule level. This approach can provide a broader view of population distributions than techniques that probe the ensemble average.

Laser-based optical probe techniques show great promise for real-time monitoring of the motions of molecular machines. Techniques are evolving that side-step diffraction limits of focused light and that could monitor individual molecules in vivo and in vitro. Multiphoton and multicolor techniques hold promise for high selectivity and the simultaneous detection of multiple species. Advances in detector technology are still needed to permit the real-time imaging of multiple species, making possible the simultaneous monitoring of chemical composition and molecular motion. Multiplexed measurements performed in real time could uncover details of the mechanisms for building and operating molecular machines.

A highly desirable goal is the capability to directly view and study the workings of molecular machines (i.e., protein complexes). This accomplishment is constrained by the sizes of complexes that range from angstroms to tens of nanometers. The size range is much smaller than the resolution of conventional optical microscopic techniques, which are limited by the diffraction limit of light. Other approaches, such as scanning electron microscopy (SEM), TEM, STEM, and atomic force microscopy (AFM, Fig. 1.9), have the resolution to directly image at these scales; however, these techniques require experimental conditions that do not lend themselves to real-time study of dynamic systems. Strategies for observing the operation, motion, and assembly of a molecular machine as a function of time are possible and should be further developed.

One limitation of current methods is that external tags must be used to observe the structure or motion of the system of interest. These tags must be highly selective, nonperturbative, easily detected, resistant to degradation, and have unique spectroscopic signatures. The most common probes for this purpose are fluorescent markers, even though they do not have optimal properties. Several new types of probes being developed for monitoring molecular motion include phosphorescent structures, quantum dots, and fluorescent microspheres, as well as surface-enhanced Raman particles and plasmon-resonant particles.

Fig. 1.8. Section from a 3D electron tomographic reconstruction of *Caulobacter*. This cell is budding off from the parent cell (right), while developing a flagellum at the left end. The sample was prepared in vitreous ice with no chemical fixatives or stains. Internal features such as membranes are clearly seen, as are protein complexes about the size of ribosomes and patches of the periodic S-layer. [Source: Kenneth H. Downing and Grant J. Jensen, Lawrence Berkeley National Laboratory]

Each probe has trade-offs such as size limitations and optical activity lifetime. Examples of trade-offs include a comparison of single-fluorescent molecule and fluorescent-microsphere tagging approaches. The single-fluorescent molecule tag provides a probe that is small relative to the machine under investigation, which ensures that the probe will not significantly alter the dynamics and kinetics—a concern when the microsphere used is much larger than the system under investigation. However, the single molecule has a much shorter optically active lifetime of milliseconds to minutes compared with the microsphere, which can be active for hours. Additionally, the instrument response of single-molecule emission is complicated by dynamic photophysical behavior. Despite these limitations, tags are currently used to directly observe biomolecular movement at a much higher resolution than the diffraction limit associated with laser focus. This is accomplished by fitting the optical image with the known or estimated instrument-response function. If the fit parameters accurately reflect the response function, the tag location can be determined down to the nanometer level.

Another approach to investigate molecular movement consists of observing an indirect signal that can be correlated with the movement. One such approach is single-molecule fluorescence resonance energy transfer (SM-FRET). This method measures angstrom-scale distances between two fluorescent tags. A fluorescent molecule that has absorbed a photon and becomes electronically excited has several pathways for removing this energy. It can fluoresce, internally relax (internal conversion or intersystem crossing), be quenched, or transfer its excitation to another chromophore. Any of these processes can be sensitive to the molecule's environment, and measurements of their rates can be used to probe that environment. Energy transfer from donor to acceptor chromophore is particularly interesting because of its sensitivity to the distance between chromophores. This energy-transfer behavior is a dipole-dipole interaction and can measure molecular-scale distances (Fig. 1.10).

Currently, the typical FRET experiment determines the degree of energy transfer by measuring the total amount of light detected through two different spectral filters. The average distance between FRET pairs can be calculated from the amount of light originating from each member of the pair. Using FRET on single pairs makes possible the determination not only of the average distance between FRET pairs but also of distance distribution within a sample. Individual molecule-correlation functions can be calculated and averaged to obtain overall correlation functions and therefore the power spectrum. From this information and the quantum yields of donor and acceptor, we can determine energy-transfer efficiency and distances between chromophores. If one of the FRET pair chromophores is attached to one part of a molecular motor and the other to a different part, their relative motions can be monitored.

Other approaches can include the observation of such optical properties as polarization state, fluctuations of excited-state lifetime, and fluorescence depolarization. All require labeling of the molecular complex under investigation, which in some cases is needed in multiple

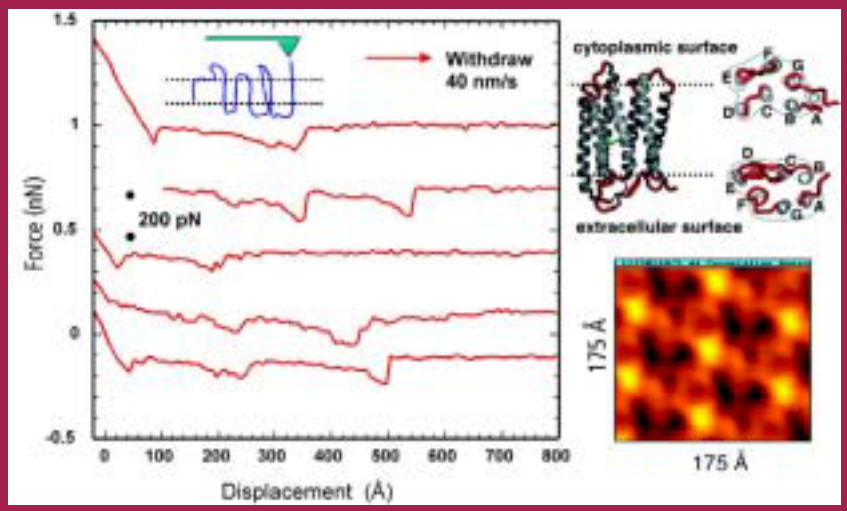


Fig. 1.9. Atomic force microscopy providing structural information on membrane proteins via imaging and single-molecule force curves of protein unfolding. [Source: Alan R. Burns, Sandia National Laboratories]

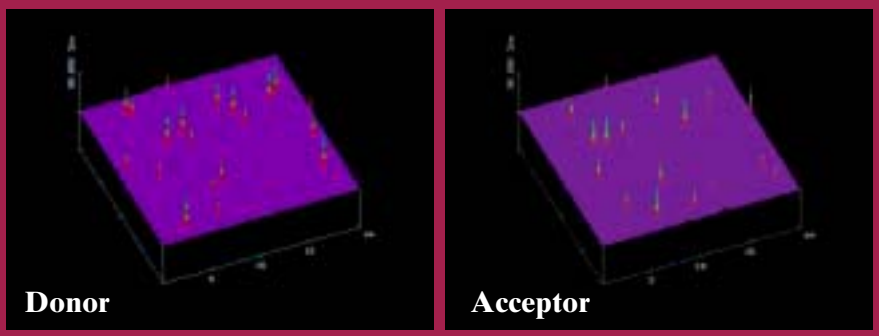
sites. Another optical approach involves monitoring movement or structural changes based on the intrinsic spectroscopic properties of the system under study. This makes the use of a tag unnecessary. Surface-enhanced Raman scattering is a possible strategy that uses a common spectroscopic property to produce a fingerprint spectrum of the structure and show correlated spectral fluctuations with molecular structure. The use of surface-enhanced approaches and predetermined spectral features correlated with the movement of interest should be capable of millisecond time scales over extended periods.

Scanning Probe Microscopy (SPM)

SPM has proven useful for determining biological structure and function. It can be performed in situ (on live organisms, if necessary) without staining. This technique potentially provides 1-nm topographic resolution. SPM can be combined with and enhance simultaneous fluorescence imaging. It also can be used to image individual proteins, protein complexes, and molecular arrays (membrane proteins). Phase-sensitive tapping mode provides information about chemical changes and "softness" (compliance). Functionalized tips can be used to generate force vs displacement curves, providing the ability to obtain single-molecule bond strength information that lends itself to the characterization of protein-protein or protein-ligand interactions and tension-induced protein unfolding.

Diffraction

Information about the size and shape of purified structures randomly oriented in solution can be obtained by X-ray or neutron scattering. Neutron diffraction is particularly useful for defining the volume of the structure that excludes water through the use of scattering in varying ratios of H_2O and D_2O. Variation of H_2O and D_2O ratios can be used to contrast deuterated proteins, making possible the study of ligand binding, including conformational changes induced by the interaction. Wide-angle solution scattering also has the potential to provide scattering fingerprints that can be used to determine the structural class of a protein in solution. For instance, this technique can readily distinguish between an immunoglobulin fold and a beta-barrel.



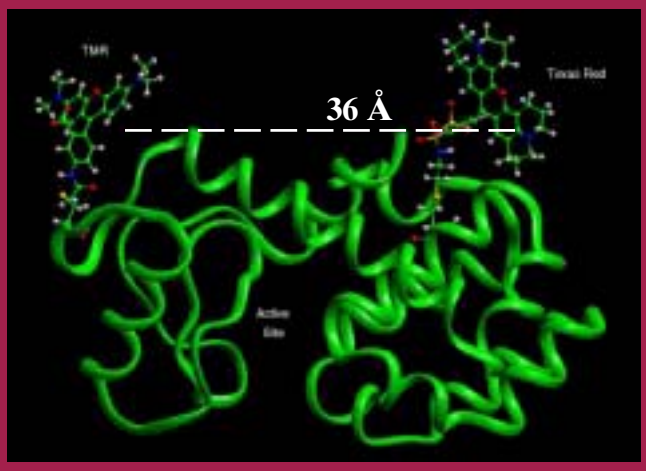


Fig. 1.10. Top: Two-channel fluorescence images (20 μm by 20 μm) of individual donor- or acceptor-labeled T4 lysozyme molecules tethered by cross-linker molecules to the hydrocarbon-modified glass surface of a cover slip under a pH 7.2 aqueous buffer solution. The donor and acceptor emissions were detected separately by a pair of avalanche photodiode detectors after passing through appropriate isolation filters. The images were taken with an inverted fluorescence microscope by raster-scanning the sample with a focused 10- to 100-nW, 532-nm laser beam. Each individual peak is attributed to a single T4 lysozyme molecule. Bottom: The crystal structure of wild-type T4 lysozyme with two covalently labeled dye molecules (fluorescence resonant energy transfer donor and acceptor: tetramethylrhodamine and Texas Red). [Source: H. Peter Lu, Pacific Northwest National Laboratory]

Cross-Cutting Needs

In future we anticipate a need for laboratory-scale sources for X-ray diffraction and electron scattering that will give scientists dedicated instruments for determining the structure of individual proteins and protein complexes comprising a cell's molecular machines. New techniques capable of sequencing single proteins online also will be needed for instant identification of a molecular machine's constituents and its location in the genome. Some of these advances will be enabled by microfluidics and microchemistry lab efforts. We anticipate needing new techniques for the specific labeling of proteins and protein domains as well as for manipulating the genome to better characterize and customize molecular machines.

Other DOE programs could meet several imaging needs. These include new probes, advances in protein sequencing and characterization, sample preparation and validation, computational simulations, and data management.

Rapid Methods for Structure Determination of Single-Protein Molecules or Complexes

Current methods for determining protein structure take years to provide a finished structure. Nuclear magnetic resonance (NMR) structure determination is currently limited to proteins of 40 to 60 kD, and X-ray diffraction methods require that the proteins be crystallized. Crystallization is the slowest step in the process, often taking many years of concerted effort. In addition, many proteins never crystallize, among them membrane proteins. Frequently, molecular signaling transverses the membrane, and more than 30% of all proteins are known to be tightly associated with membranes. New biochemical and biophysical approaches will be required to identify the molecular structures of these proteins. Creating tabletop or lab-scale X-ray sources would be a high-impact development that would allow the scientific community to take full advantage of the use of X rays in imaging protein machines.

New developments leading to the production of a "next-generation" beamline, such as the Linac Coherent Light Source at Stanford, offer the promise of obtaining high-resolution (2 Å or less) diffraction data from single-protein molecules, protein complexes of any size, viruses, and even cells without crystallization using a coherent X-ray source ten times brighter than we have today. To aid the development of this technology, which may become available as early as 2007, single-molecule diffraction techniques using this beamline will need to be established. Other needs include sample preparation and introduction, rapid-readout detection arrays, and computer algorithms for combining and interpreting diffraction data from tens of thousands of molecules.

The development of new molecular probes applies to almost all aspects of GTL imaging-based research. Section 6 summarizes research needs in this area. Identification and characterization of molecular machines will require a wide range of probes. For example, optical methods require probes with high efficiency and stability for single-molecule research as well as probes designed for use with multiphoton imaging methods. Short-wavelength methods require electron-dense probes for use as high-resolution labels for individual protein subunits in complexes.

Protein or Sample Characterization

Development of High-Throughput Methods for Identifying Protein Function

Identifying protein function can be a daunting task. The GTL program would be advanced significantly if methods could be developed to enable the rapid identification of an unknown protein's function. Continued increases in numbers of known (experimentally derived) protein structures with different folds will aid the analysis of protein sequences as a first step to predict function using such computational methods as comparative modeling and threading. Function may be inferred or predicted by identifying unique folds, particular combinations of structural motifs, or structural features that match known substrate, cofactor, or other binding sites.

New array or chip-based protein-function assays should be developed for experimentally screening large numbers of proteins for function or functional attributes. This could be accomplished, for example, using arrays that test proteins for their ability to bind to a wide assortment of small molecules (e.g., metals, cofactors, and carbohydrates), macromolecules (e.g., ssDNA, dsDNA, RNA, and specific DNA sequences), antibodies, or other proteins. Direct assays that actually monitor enzymatic function also could be developed in similar array format.

Specimen Integrity and Functionality

Challenges for the imaging of molecular machines remain significant. First, new methods must be developed for producing enough of these machines for biochemical, biophysical, and imaging studies. In many cases, protein constituents of molecular machines are insoluble on their own in the absence of other machine constituents. Only by coexpression of multiple constituents—ideally all constituents—is it possible to produce specimens containing soluble samples of these machines. Although some coexpression vectors are available, additional ones need to be developed to enable the production of the many types of molecular machines to be studied in the GTL program.

The integrity of the purified molecular machine also must be assured. Integrity depends on the isolation protocol, extra components that may be required for stability, other preparative procedures (e.g., freezing), and deposition on the substrate used for imaging. Ideally, a functional assay should be available to ensure the viability of the complexes as prepared for imaging. Alternately, the consistency of results obtained by different imaging modalities and specimen preparations must be assessed. Fluorescence-based single-molecule (complex) spectroscopy combined with AFM structure mapping has the potential to assay for sample purity and reveal the relationship between secondary structure and dynamics or function.

Computational Simulations

Simulations of the structure, activity, and assembly of a molecular complex will need to be carried out on multiple-length scales and various levels of detail. For instance, the most detailed simulations will involve molecular models including atomic coordinates of all constituents. Subunit structures from X-ray crystallography may be packed together to provide atomic-resolution models. Less-detailed models will be useful for simulating the dynamics of complex formation, mechanical-force generation, and docking. Other models will be needed to integrate data on distances between fluorescent tags, assess the consistency of data collected by different experimental techniques, and link data available from techniques operating at very different length scales. Ultimately, all image information should be integrated into a dynamic computer simulation with molecular detail that will allow the investigator to observe, at the highest level of detail, the structure, assembly, and operation of the entire molecular machine and its interactions with other cellular components relevant to its functioning. This simulation should allow for changing the parameters of the system and predicting the resulting changes in function. Modeling of molecular dynamics on the time scale (nsec to msec) of many biological processes (e.g., protein refolding during complex formation) is an important, unsolved challenge.

An equally important mission of simulations will be the prediction of instrumentation performance. Details of the image-formation process will identify potential image artifacts and provide a basis for quantitative analysis of image imperfections including, for instance, the potential to correct for aberrations or AFM tip geometry. These simulations also will support design of advanced instrumentation with optimized imaging modalities.

Mechanical or radiative damage resulting from imaging may also be assessed through analysis of the process of damage followed by simulation. This analysis will allow for optimized data-collection and data-analysis strategies to compensate for the effects of damage.

Data Management

Any imaging techniques applicable to the study of molecular machines will result in the generation of huge amounts of data. Storage, manipulation, visualization, registration, and comparison of image information will represent a very

significant challenge. Separate images of identical complexes using either the same or different techniques will need to be compared and contrasted. Extracting information that will lead to functional insight from the vast amount of image data will require significant advances in visualization and image classification.

The creation of comprehensive image databases also represents a very significant technical challenge. Data intrinsic to images in one, two, or three dimensions, not to mention dynamic (fourth dimension) or correlation information (additional dimensions), are not readily organized in a universally applicable format. Instrumental and specimen parameters also will be essential parts of any useful database. A detailed analysis of image database needs should be addressed at the earliest possible time. Standardization requirements for these data protocols need to be defined early in the program. This will aid in the design of experimental protocols that will define the parameters necessary to maintain the long-term usefulness of the data collected. Public access to image-analysis software and databases of images will be an important product of the GTL program.

Because of the importance of computational tools and methods to the use of imaging in biological research, they are discussed in detail in Section 5 of this report.

Production of Lab-Scale X-Ray Sources

One high-impact development that would allow the scientific community to take full advantage of the use of X rays in imaging protein machines would be the creation of tabletop or lab-scale X-ray sources.

2. Intracellular and Cellular Structure, Function, and Processes

SECTION 2

This section addresses needs for molecular imaging of protein machines and other key species within living cells. In response to GTL Goal 1, "Identify and Characterize the Molecular Machines of Life—the Multiprotein Complexes that Execute Cellular Functions and Govern Cell Form," the focus is on understanding the role of these machines in living systems.

Obtaining information on sub- or intracellular processes from living cells while they exert their normal function is highly desirable. These processes involve transduction pathways within a cell, cell reaction to external stimuli or signaling, and internal or external processes that lead to cell aging and apoptosis. To achieve such understanding, experimental challenges have to be met:

- Imaging techniques that provide an overview of cellular processes have to be reasonably fast, ideally on the order of a few images per second or faster. Imaging techniques with high resolution need to be noninvasive or minimally invasive so as not to perturb the cellular function.

- Probes or local reporters need to be developed that can easily be taken up by and have minimal effect on the cell. These probes also have to be extremely stable (neutral) to photon interactions and should adapt to cellular environments. An ideal probe would change, for instance, its property (and therefore report) changes in pH, oxygen concentration, or glucose concentration, while maintaining its response to external probing. Such probes should also be highly localizable within the cell and convertible to X-ray– and electron-dense markers.

Promising approaches to these problems are the newly developed vibrational microscopies and spectroscopies, X-ray microscopy, and electron microscope (EM) tomography.

Understanding the Genetic Basis for Microbial Function

Developing novel methodologies would be valuable in helping to establish the relationship between the genome and microbial function in living cells. First, several questions need answering: Can cell-imaging methods be devised to determine the direct correlation of gene mutations and gene expression to function in microbial cells? Can methods be developed to relate gene activation to the simultaneous detection of metabolism and secretion in microbial cells? Can high-throughput, live-cell imaging methodologies be formulated for approximate and selective genotyping? Studies of community dynamics under environmental perturbations would be greatly enhanced through real-time, genetic classification of individual microbial cells within a mixed culture, even if this classification is only approximate and must be further refined by more rigorous genotyping methods.

Issues and Limitations

The following will require technological advances:

- Imaging of single molecules (and sets of molecules) within living cells.

- Improved spatial resolution and optical probes for co-localization of data from different techniques below the resolution limits of standard optical images.

- Development of activity-neutral dyes and other optical probes.

- High-throughput approaches that enable the rapid analyses of large numbers of cells.

Capabilities developed for this element of the GTL program are expected to have wide applications to the study of cellular biology.

Recommendations

The following actions are recommended:

- Coordinate program to develop all methodologies, in particular co-localization of data from different techniques.
- Establish databases (e.g., protein spectra).
- Develop open source, common data formats.
- Stimulate partnerships between biologists and instrumentation scientists.
- Focus on specific needs of each methodology (e.g., some need more facilities, and others need stimulated development).

General Requirements

Recently, many promising techniques have been developed for studying intracellular and cellular structure and function, as well as the time evolution of cellular mechanisms via kinetics and dynamics. Ultimate requirements are capabilities for imaging functional groups and studying kinetics and dynamics in samples of a few microns in different environments, in vivo, and with a spatial resolution of tenths of a nanometer. Clearly, no one technique will offer the optimum in all required categories. Further, integration of data from techniques at different size scales is necessary. Real progress is likely to rely on multidisciplinary teams (e.g., physicists, chemists, and biologists) working together.

Current techniques can provide spatial resolution of a few tens of nanometers, with both chemical and elemental specificity and dynamical information on the picosecond time scale. Work on techniques that perform at this level is preliminary at many laboratories, and a database is needed to enable comparison of techniques and to provide benchmarks of technical progress in advancing particular imaging tools. Sample-preparation standards will allow comparable samples (and sometimes the same samples) to be measured by a variety of techniques for correlation of results.

The current evolution of techniques is toward higher spatial and spectral resolution, establishing data that, if properly organized and shared, can be used to identify similarities among the thousands of prokaryotic and eukaryotic organisms and their constituents.

Specific Needs

Most high-resolution (<50-nm) imaging techniques work well with thin (10-μm) specimens. Recently, however, cells were shown to function differently in 2D and 3D cultures. Consequently, an increasing number of scientists are growing cells in 3D matrices. Although these cultures more closely resemble the cell's natural environment, they present unique imaging challenges. Obtaining information about molecular interactions in these cultures without excessive and potentially damaging processing is extremely important. We need to develop tools that can collect data about the location of DNA, RNA, and proteins, as well as molecular interactions, at better than 10-nm resolution without having to process the cells (e.g., dehydrate, embed, and section the cultures).

We have the ability to obtain high-resolution images of whole, hydrated cells. We also can obtain exquisite information about the genetic and protein composition of cell populations through DNA and protein microarray analyses. Combining the power of these two technologies could enable the imaging of live cells and furnish information about the cell's genetic and protein composition.

Understanding the internal functioning of cells in their natural environments is important and particularly challenging. Analysis often requires the ability to visualize the cells within thick, dense matrices. For example, specific molecular (DNA, RNA, and protein), structural, and functional information must be obtained from specific bacteria within biofilms in their natural environment.

Powerful tools exist to study molecular machines in vitro as well as to obtain structural information about these machines from crystallography and STEM techniques. We need to be able to examine these machines within living cells to obtain functional information (in addition to that obtained using optical methods such as FRET) at better than 50-nm resolution.

Tools also exist to examine protein-protein interactions in living cells both in vitro and in situ using fluorescent imaging techniques. We need to be able to obtain structural and functional information about molecules (DNA, RNA, proteins, and protein-protein interactions) within living cells in living organisms at better than 10-nm resolution.

Methods are needed that will ultimately allow us to track the movement of macromolecules, vesicles, molecular machines, and other such structures within a cell with nanometer-scale precision. Current developments in this field are based on the detection and accurate position determination of bright labels such as semiconductor quantum dots, dye-loaded spheres, and light-scattering particles. Other attractive probe developments being made in the area of delayed fluorescence can detect single probes against the large background of cell autofluorescence by time-gating the photon collection.

Current Methodologies: Status and Development Needs

X-Ray Microscopy

X-ray microscopy today uses electromagnetic radiation with photon energies in the range from 250 eV to 30 keV, with corresponding wavelengths from 45 nm Å to 0.4 Å. Because of the short wavelength, suboptical spatial resolution can be obtained. Several different contrast mechanisms can be used to form an image, thereby making X-ray microscopes useful for probing different properties of matter. X-ray fluorescence mapping can be applied to mapping trace metals such as Fe, Cu, Zn, As, Se, Cd, Hg, Pb, or U in organisms including microbes or eukaryotic cells, while simultaneously collecting distribution from P, S, Cl, K, and Ca.

The critical element of any X-ray microscope is the zone plate used for imaging or focusing. The outermost zone width and aspect ratio define the achievable spatial resolution and diffraction efficiency. Therefore, a major investment in zone-plate technology is to improve the resolution. Achieving an outermost zone of 10 nm or less, while maintaining a thickness in the range of 200 nm for soft X rays and >>1 μm for hard X rays, would enable a major breakthrough in all applications. Another challenge is X-ray tomography. While transmission contrast imaging lends itself to 3D imaging in both the soft and hard X-ray region (Figs. 2.1 and 2.2), 3D imaging using X-ray fluorescence—even though potentially very powerful (Figs. 2.3 and 2.4)—is currently not feasible because of the long exposure times involved. A major improvement could be achieved by using novel detectors to maximize the solid angle of detection and by improving the brilliance of hard X-ray storage rings.

X-ray imaging, an emerging field, will be invaluable for studying hydrated organic matter at high resolution, particularly organic and inorganic composites such as bacterial colonies (e.g., in soil, microbial communities, and biofilms). X-ray tomography can be carried out on whole cells or colonies of bacteria up to 10 μm thick, without the time-consuming sectioning required for TEM, and generate a 3D reconstruction of the cell or colony within 1 h at ~50-nm resolution.

In addition, the entire arsenal of heavy-metal labeling methods for TEM can be used for X-ray microscopy. As a result, correlated live-cell light

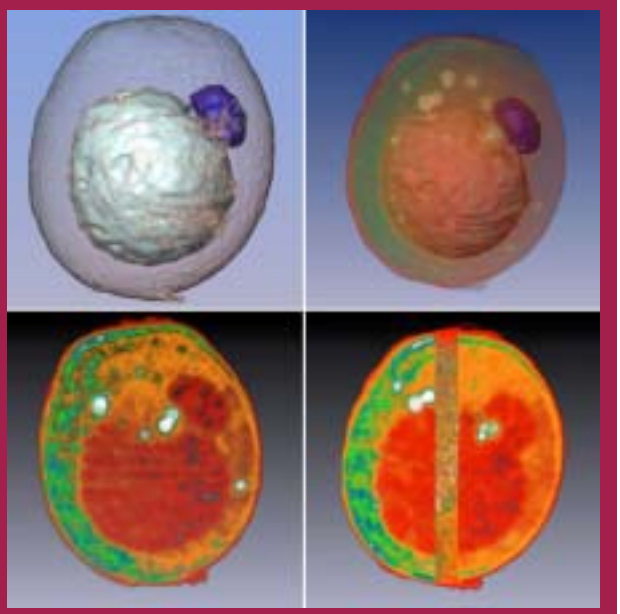


Fig. 2.1. Three-dimensional soft X-ray reconstruction of yeast. Example of 3D volume visualization of a tomographic data set showing various pixel- and surface-classification techniques. [Source: Carolyn A. Larabell and Mark A. LeGros, Lawrence Berkeley National Laboratory]

and high-resolution X-ray imaging analyses can be used to examine both the dynamics of specific proteins and the ultrastructural distribution of those proteins in the same cell.

Future implementations of mixed 3D imaging involving X-ray, zone-plate direct imaging, and lensless diffraction imaging promise to extend the spatial resolution of 3D X-ray methods to the nanometer level in thick, hydrated samples. X-ray imaging will, therefore, play a unique role in quantitative 3D measurements of protein and elemental distribution at high spatial resolution.

Optical Microscopy

Optical methods vary in complexity, from inexpensive and routine tools to major interdisciplinary investments at the national laboratories. The accessible distance scale ranges from the "optical limit" of about half the wavelength to macroscale or, in other words, from the bacterial and cellular organelle scale to sample sizes limited only by optical transparency. Image production providing rich information content is a result of a wide variety of physical and chemical properties. In favorable circumstances, the detection limit may be a single fluorescent molecule. We can classify optical imaging methods according to difficulty and application to define a role for GTL in meeting the needs of the biological community. This classification also is intended to show where multidisciplinary teams of physicists, chemists, and biologists should work together to produce the capabilities required for the next generation of methods and instruments. Each of these methods is discussed in some detail in Appendix C.

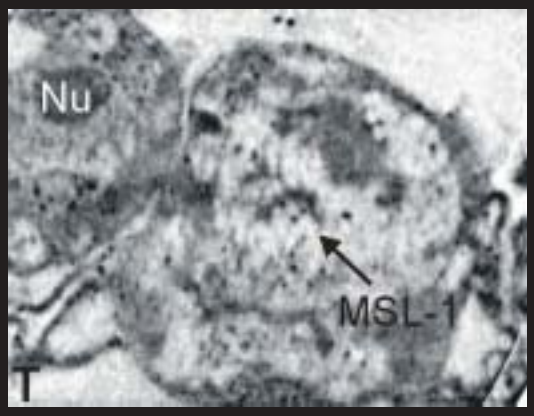


Fig. 2.2. Soft X-ray tomography. Section through a tomographic reconstruction of a *Drosophila melanogaster* cell. Cells were labeled against the MSL-1 protein using 1-nm colloidal gold conjugated antibodies and silver enhancement. Cell nucleolus (Nu) and MSL-1 distribution positions are marked. [Source: G. Schneider et al., "Nanotomography of Labelled Cryogenic Cells," *SPIE Proceedings* **4503**, 156 (2001)]

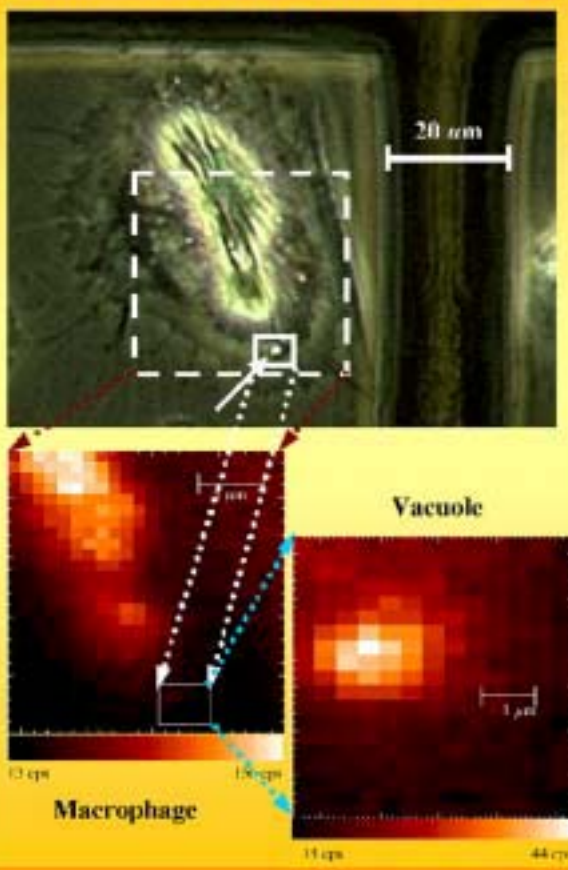


Fig. 2.3. X-ray fluorescence image of *Mycobacterium avium*, 24 h after infection of a fixed and air-dried macrophage. A map of the Cl distribution of both macrophage (30- by 30-μm field) and mycobacterium (3- by 3-μm field) is shown. [Source: Dirk Wagner, Kuzell Institute for Arthritis and Infectious Diseases; Jörg Maser, Barry Lai, and Zhonghou Cai, Argonne National Laboratory; and Luiz Bermudez, Kuzell Institute for Arthritis and Infectious Diseases]

Far-Field Imaging Methods Based on Intramolecular Vibrational Modes

Far-field vibrational-imaging methods will complement fluorescence- or absorption-based optical approaches for the study of individual cells and colonies. The use of synchrotron radiation and lasers provides diffraction-limited spatial resolution that could be improved further by image-enhancement algorithms. Vibrational signatures specific to molecular structure can provide direct chemical information without the use of dyes or other labels. In addition, vibrational imaging can be exploited to observe dynamic changes in living cells. Far-field infrared vibrational imaging can be developed to examine community signaling or colony and substrate chemistries at the scale of 10 μm and greater.

Collectively, vibrational-imaging techniques should be developed to analyze the microbial response, at the level of individual cells or communities, to perturbations such as changes in oxygen or other substrate levels (e.g., NH_3 and metals), pH or temperature changes, and other factors that will affect critical elements of the interaction of microbes with their environment. Because such hyperspectral-imaging methods can potentially yield large data sets including spectral, spatial, and temporal dimensions, imaging-analysis methods that can properly exploit and manage such data sets should also be developed.

In addition to their application to intrinsic molecular signatures, vibrational-imaging methods could be developed to take advantage of the introduction of vibrational "dyes." Thus, "tagging" the pathways or structures in microbes by the use of vibrational labels such as stable isotopes or ligands (e.g., CN and CO) could allow selective detection of specific molecular components within and among microbes.

Near-Field and Scanning-Probe Imaging Methods

Near-field microscopies typically are based on the spatial confinement of photons by probe microscopy structures. They offer attractive pathways to achieving resolution beyond the diffraction limit and potentially could combine topographical imaging with chemical sensitivity. Because of their dependence on probe microscopy techniques, near-field probes are usually limited to cell exteriors, but this could be overcome with newly developed photonic force microscopes that could "grab" artificially introduced labels and move them within the cell to probe the interior.

Chemically functionalized near-field tips also could be injected into the cell through the membrane and positioned to measure the uptake of chemicals through the cell membrane or their diffusion within the cell. The development of near-field imaging capabilities that could use narrow-band laser light to achieve ħ0 with apertures and eventually ħ00 without apertures with scanning probes in living cells and membranes is highly desirable.

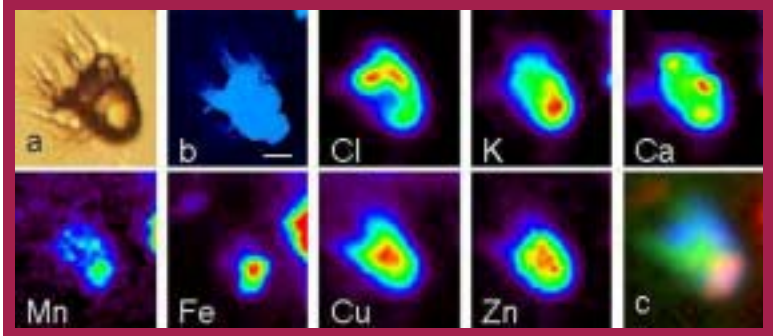


Fig. 2.4. (a) Soft X-ray micrograph and (b) X-ray fluorescence map of marine ciliate showing the cell's C content (scanning X-ray microscope, National Synchrotron Light Source at Brookhaven National Laboratory). [Scale bar (on b) = 6 Fm] Cl, K, Ca, Mn, Fe, Cu, Zn: X-ray fluorescence maps showing distribution [hard X-ray microprobe, Advanced Photon Source, sector 2, at Argonne National Laboratory (ANL)]. c: Superimposition of Cl, Fe, and Cu maps. Cl was chosen as a proxy for cellular biomass, while Fe and Cu represent two trace metals with notably different spatial distributions within this cell. [Source: Benjamin S. Twining, Stephen B. Baines, and Nicholas S. Fisher, Stony Brook University; Chris Jacobsen and Jörg Maser, Argonne National Laboratory]

"Classical" One-Photon Methods

All classical microscopy methods use continuous or low-power light sources. Contrast may be improved by physical methods such as phase or differential-image contrast. Specific and very useful information is provided by fluorescence methods that use intrinsic or added dyes. In general, fluorescence techniques and many fluorescent probe molecules are available, but improved labels are needed.

Obtaining chemical data (location of abundant species or important minor chemical compounds) is difficult with the above techniques. A direct means for obtaining molecular information is by vibrational resonances, either infrared or Raman. Even though improved detection technology will be useful, technologies for Raman and fluorescence measurements are approaching their theoretical limits, requiring the development of improved probes.

The now-accessible terahertz spectral range yields many vibrational excitations that can be explained by molecular dynamics calculations. High-power sources available at the national laboratories can be used for imaging applications, including detection of bioterrorism agents such as anthrax. Additional research is required to achieve the potential of this approach.

A special imaging category is scanning optical tomography, which requires a pulsed light source to provide depth profiling data. Although the laser must produce ultrashort light pulses, high power is not required, and obtaining 3D information useful on a tissue scale is possible, but not at the single-cell level.

Multiphoton Microscopy

The most promising avenue for obtaining new kinds of cellular information—and the area where funding at the national laboratory level can provide the greatest impact—is in multiphoton methods. A wide variety of characteristics can be studied, with the general requirement of an ultrafast laser technology. Multiphoton methods, which show real promise for providing chemically based imaging, offer the unique advantage of possibly generating several imaging signals simultaneously. All signals are generated at the laser focal point and can be separated to create a rich, perfectly registered, multidimensional-image data set containing a variety of chemical, physical, and structural data. Wider access to multiphoton methods will require improved (smaller, cheaper, and more reliable) laser technology and optical probes. Advanced imaging tools will help continue progress in solving a wide variety of biological problems.

Hole-Burning Microscopy

A hole-burning imaging (HBI) approach for probing local environments in microbial cells can be viewed as an optical analog of MRI. This high-sensitivity technique has the potential for detecting local effects on molecular probes. In principle, HBI has the spatial resolution of multiphoton microscopy methods (i.e., on the order of 100 nm). Whereas MRI is based on proton T_1-relaxation times, HBI is based, in part, on T_2*-pure optical dephasing times of organelle-specific dye molecules. The sensitivity of T_2* to subtle differences in molecular environment gives the technique its high specificity.

Dynamical Studies

Direct measurements of kinetics and dynamics in living cells could be obtained on a picosecond time scale by adapting pump-probe methods to obtain high spatial resolution. The nanosecond time scale can be probed by measurements of lifetimes and rotational correlation times that are sensitive to processes such as protein association to form complexes. Approaches of this nature have provided valuable information on a variety of biological systems, for example, retina function, photosynthesis, and energy conversion.

Electron Tomography

Electron tomography has the potential to reach the highest spatial resolution in imaging 3D structures of intact cells and mapping subcellular component locations. Theoretical analyses suggest achievable resolution of 2 to 5 nm with such specimens. At this resolution we should be able to identify such subcellular features as cytoskeletal filaments and nucleic acid condensates and such large macromolecular complexes as ribosomes, proteasomes, and polymerases. This could be achieved by matching the observed

densities with structures of corresponding molecules that have been solved by X-ray crystallography or other techniques.

Although being able to directly identify the larger protein complexes will provide a wealth of new information, selective labeling of components with electron-dense agents will enable localization of more components. The labeling strategies could include conventional antibody-targeted metal atom clusters. The use of quantum dots, for example, would allow correlative light microscopy, but the particle size for electron microscopy would be smaller than for current light microscopy.

One problem to be addressed is how to get the particles into the bacteria. Evidence suggests that complexing them with a naturally imported peptide such as transferrin may work. Other labeling approaches under development include modifying a particular gene to include a photoconvertable eosin or tetracysteine peptide that can be complexed with a heavy metal atom cluster.

Instrumentation development probably will be driven by the materials science community. Specimen preparation for frozen-hydrated, whole-mount samples is fairly well developed, but improving the methodology for frozen-sample sectioning would open up the technique to applications with multicellular samples. Finally, managing the image data sets and searching the reconstructed volume densities for patterns corresponding to known structures will require significant computational efforts.

3. Monoclonal and Heterogeneous Multicellular Systems, Cell-Cell Signaling, and Model Systems

SECTION 3

Goal 3 of the GTL program is to "characterize the functional repertoire of complex communities in their natural environments, at the molecular level." Even though the genomic sequence for several different organisms is known and we understand the general outlines of how some communities are composed, we lack an understanding of community function distribution among members and the molecular basis of different functions. This information is required to connect molecular expression data to the overall function of a community in a natural setting. To bridge this enormous gap of complexity, there is a need for simple, well-characterized model systems (i.e., laboratory, bench-scale, pure-culture, and mixed-culture cell systems) that capture essential functional aspects of natural communities (Fig. 3.1).

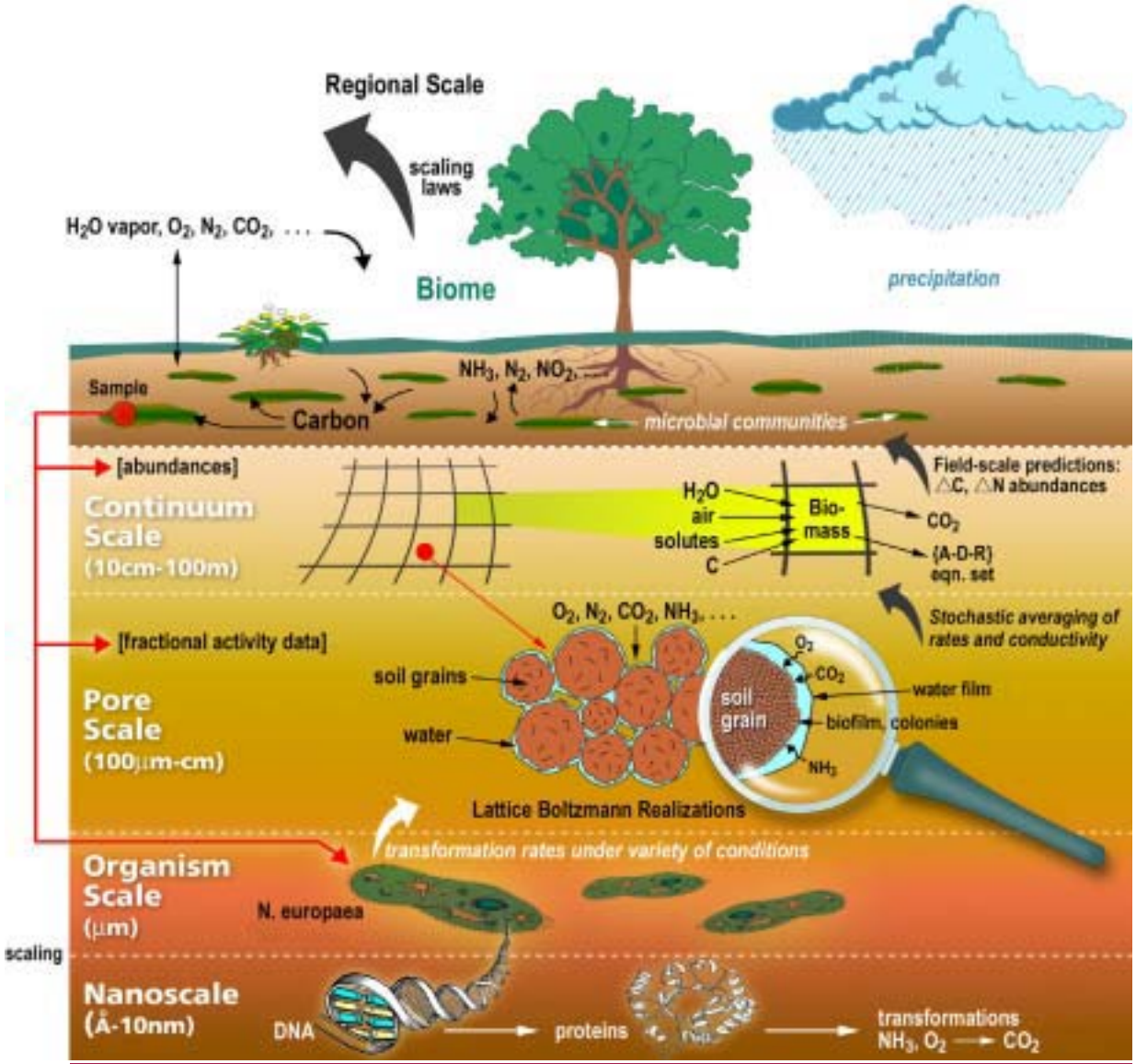


Fig. 3.1. Relevant level of complexity inherent in microbial soil ecosystem.
[Source: Pat J. Unkefer, Los Alamos National Laboratory]

Why Imaging is Needed

Imaging technologies are ideal for capturing the complexity of multicellular interactions because of their ability to acquire spectral, spatial, and temporal data over a wide range of scales. Because of their parallel and quantitative nature, imaging approaches are inherently high throughput (i.e., they can accommodate thousands to tens of thousands of simultaneous experiments). To realize this potential, model systems must be adequate for developing and validating imaging technologies. In turn, well-defined, multicellular model systems can reveal important principles that govern cell communities. Both monoclonal (pure-culture) and heterogeneous, multicellular (mixed-culture) systems need to be defined. Specifically, we must identify

- Structural and functional properties of multicellular systems,
- Imaging techniques needed to determine these properties, and
- Major technological challenges.

Understanding complex communities of cells, including how cells interact with their own and different species, is fundamental to the goals of the GTL program. Interactions include the exchange of information through signaling mechanisms involving chemical, biochemical, and ionic mediators. These processes, in turn, are directly impacted by the extracellular environment. Thus, understanding the interaction between various cells and the environment is essential for understanding complex communities.

Imaging technologies can provide this understanding; the challenge, however, is to image information flow among cells. First, we must understand the architecture of the molecular machines that mediate such flow. We must determine which cells are talking and which are listening. To do this, we must be able to measure signaling dynamics and image metabolites and regulatory molecules. In addition, we must be able to track metabolites and regulatory molecules spatially and temporally. Finally, we must connect this information to the consequent cellular responses (Fig. 3.2).

Imaging Issues

Several technological issues must be addressed, the first being those of spatial and temporal resolution. Microbes and organelles are close to the theoretical resolution of optical microscopy. Signaling molecules and metabolites are normally found at exceedingly low concentrations. Natural microbial communities are frequently found in optically opaque conditions or in the presence of interfering compounds. Experimental model systems can be constructed to prevent some of these problems, but they give rise to other issues. For example, high diffusivity in culture systems can affect sensitivity and spatial resolution. Then there are the fundamental problems related to measuring signal-dependent cellular responses and identifying unlabeled signaling molecules.

Improved Model Systems for Microbial Imaging

In the context of the GTL program, several scientific questions have been identified (see other sections) that are relevant to the goal of understanding the molecular basis of complex cell communities. Approaching these questions in

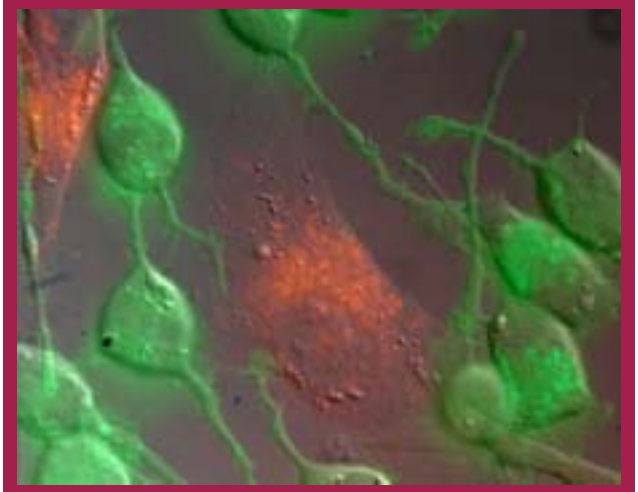


Fig. 3.2. Visualization of cell-cell communication in living cells. Mouse cells were engineered to express either a membrane-anchored growth factor tagged with a red fluorescent label or the complementary receptor tagged with a green fluorescent label. The cells were visualized by differential interference contrast optics. [Source: H. Steven Wiley, Pacific Northwest National Laboratory]

the context of natural communities is very difficult, however, and thus should be initially studied in well-defined model systems. In addition, a variety of powerful experimental imaging methodologies currently can be used only on model systems, and some cellular processes occur only in a multicellular context. We should use model systems that are simple enough to facilitate the development of imaging technologies but are still relevant to complex environments. The specific characteristics and components of an ideal or nearly ideal experimental system include

- Sequenced genomes,
- Genetically tractable genomes,
- Cultures grown and maintained under defined conditions,
- Culture state followed by measurable parameters, and
- Culture geometries suitable for multiple imaging technologies.

Specific scientific questions will, in large part, drive the choice of experimental system. For example, monoculture systems are most appropriate for investigating quorum factors, whereas mixed-cell systems will be required for investigating metabolic coupling, such as that observed in nitrifier (nitrate-converting), denitrifier (nitrate-removing), and syntrophic[1] systems. The challenges of monoculture vs mixed-culture systems are distinct, and both types of model systems should be developed.

Challenges in producing multidimensional model systems include culture stability and composition, analytical techniques appropriate for complex systems, and individual cell manipulation. We need to define how multispecies cultures evolve. How do we track all the parameters that change? Is there a way to create a virtual "chemostat" for multicellular systems? Both symbiotic and syntrophic systems are particularly attractive because of their mutual dependence, which will allow identification of temporal and spatial regimes where imaging can be productively used. Model systems should be standardized to facilitate information interchange.

[1]Syntrophic associations include two or more organisms for which intimate interaction is required for the complementary exploitation of resources.

Interrogation of Model Systems

Imaging technologies can provide very valuable information on the dynamics of cellular populations and the molecular processes occurring with the cells. To link molecular and cellular properties to overall community function, a number of parameters must be followed simultaneously. Imaging-based approaches should follow as many parameters as possible to create a high-throughput research capability using sample-preparation techniques that will not perturb the chemistry of the cells or their environment.

Identify Different Cell Types and "State"

Specific ligands and probes for cell labeling, both gene and protein level, are needed. We need to better define unique spectral signatures of specific cells and their metabolic state and associate cells with specific gene-expression patterns; thus specific marker genes should be identified and this information used to create "reporter" cells by inserting fusion proteins behind specific genetic elements. Better probes to directly monitor the metabolic state of cells are also needed. Advantage should be taken of endogenous cofactors that display absorbance spectra dependent on metabolic state. Visualization approaches should be developed to identify the activation state of specific signaling networks triggered in response to environmental perturbation.

Track Individual Cells in Time and Space

There is a strong need for both high-resolution and wide-field technologies. Being able to quickly switch between these two modes would be extremely valuable because it would allow tracking of both cell communities and individual cells. Following the trajectory of individual cells over a significant length of time, sometimes for many days, is often necessary. Useful model systems will include both 2D and 3D geometries; therefore, cell tracking should not be restricted to any particular one. Following individual cells for significant lengths of time also requires the development of stable cell-identification probes or establishment of specific spectral signatures.

Track, Quantify, and Identify Signaling Molecules in Time and Space

Another strong need is to define spectral signatures of important signaling molecules or to develop appropriate "tags." A valuable approach would be to make "detectors" or optical sentinels that can be placed in the extracellular environment. Other needs are specifically labeled signaling molecules and tagging approaches that allow correct transport with high signal-to-noise ratios. Obviously, these approaches should not interfere with function. Additional attention should be devoted to developing infrared, X-ray, nuclear magnetic resonance, mass spectroscopy, and Raman measurements of metabolites both within cells and in the extracellular environment. Nondestructive nanoscale probes could provide information without interfering with imaging technologies, and many of these probes could circumvent limitations of optical techniques.

Understand Cell-Environment Interactions

To understand multicellular systems, we also must investigate interactions between cells and their physical environment. Different physical and chemical interfaces will have distinct effects on individual cells within the community that can, in turn, influence their interactions. Effort should be focused on defining important environmental aspects that impact specific multicellular systems.

Many of the same techniques for quantifying metabolic products can be used to study the interaction of cells with their extracellular environment, but additional approaches will be required. Examples include (1) imaging technologies that can be "tuned" to specific classes of compounds in the extracellular environment, such as metal and organic compounds and reactive species and (2) specific probes for such identified compounds as organic molecules found in extracellular polymers, secreted proteins, nucleic acids, polysaccharides, metals and their states of oxidation, chlorinated aromatics, and aliphatics. In particular, we need to image species that are present, concentration gradients associated with chemical potential, and different pH states. Methods for visualizing macroscopic transport processes also should be developed.

Understanding what cells produce to modify their environment is important. Efforts should be made to determine all the proteins secreted by cells or actively shed from the cell surface, including both soluble and insoluble proteins. Because environmental modification will vary with time, labeling techniques are needed for examining rates of synthesis and deposition of proteins and other organic molecules. These techniques can include isotopic and fluorescence labeling and other suitable imaging approaches.

Required Instrumentation and Data Analysis

Imaging instrumentation for multicellular systems should have the following ideal characteristics:

- **Noninvasive.** Nondestructive, with high temporal and spatial resolution.
- **Highly sensitive.**
- **Multimodal.** Able to span multiple spatial scales and see population behavior (wide field of view), then zoom down to an appropriate resolution (e.g., molecular). For example, use one mode for identifying a particular cell and then use a high-resolution mode to specifically interrogate the cells or simultaneously gather multiple data streams for temporal and spatial correlations that can establish functional relationships.

Because of the complexity of cellular processes, establishing cause-effect relationships is often impossible without precise correlative data. Instruments should be accessible to many different groups of investigators. Accessibility can be guaranteed either by reserved time at user facilities for imaging or through the availability of onsite culture facilities. A strong need also exists for development of compact (bench-scale), affordable instrumentation for single- or multilaboratory use. Other needs are improved software for automated image acquisition, analysis, and customized statistical analysis and new algorithms for multivariate analysis and multimodal imaging.

4. Imaging Microbial Communities

SECTION 4

Microorganisms and the processes they catalyze are fundamentally important to DOE missions in energy and environment. Microorganisms rarely live alone in nature but rather in populations and communities of varying complexity. These communities drive the flow and cycling of energy and elements (biogeochemical cycles) in the biosphere and contain most of the biochemical diversity on earth.

A major goal of the GTL program is to develop a better understanding of the structure and function of microbial communities from their genome sequence–based information. Imaging can play a major role in testing hypotheses and predictions generated from this information at the community level. The understanding of community structure and function will enable the manipulation of microbial communities for the benefit of mankind. This includes meeting such DOE needs as environmental cleanup, carbon sequestration, global climate change, and clean energy production. The advancement of imaging for systems of this scale and complexity will also advance DOE's health-related research in low-dose radiation and radiation therapy.

Issues

DOE's Microbial Genome Program continues to produce whole-genome sequences that are being used to study the detailed biology of dozens of individual cells and populations of microorganisms of interest to DOE. The GTL program is providing extensive genome sequence–based data on microbial communities, but a major challenge is to link all this information to communities in nature. Structural and functional analysis will be needed because microbial environments can be extremely diverse and can influence the way communities function. Critical questions to be addressed include the following:

- How are the community and its physical environment arranged? Communities may be layered or composed of heterogeneous populations within biofilms. Physical and chemical properties in communities can vary dramatically over the scale of a few microns.
- What are the members, and where do they reside? What are their phylogenetic and functional properties?
- How are microbes interacting among themselves and with their environment? How does information (e.g., genes and signaling molecules) flow within the community? What are the ecological interactions such as predation and symbiosis among the various members?
- What is the flow of energy and elements through the communities, and how is this flow regulated?
- What are the intrinsic biological (genetic) and environmental factors that control the structure, stability, and functioning of communities?

Studying Microbial Communities

Most approaches currently used to study microbial communities are poorly suited for answering these questions. They provide, at best, a snapshot of a fraction of the community. Methods include culturing, bulk chemical, physical, and even molecular-based techniques that provide limited information on the spatial and temporal events critical to understanding community function. For example, information on gradients and spatial heterogeneity is lost in bulk analyses. The availability of whole-community genome sequences will provide an enormous amount of information about potential gene function and, therefore, community function. However, gene sequences will allow only function prediction, and such predictions must ultimately be tested

through in situ studies of communities. Microbial communities will require coupled biological, chemical, and physical analyses at multiple scales that may best be achieved through imaging.

Imaging, as it pertains to community analysis, provides spatial and temporal information on the chemical, physical, and biological characteristics of the community and its environment. This information can encompass length scales from the molecular to that visible with the human eye. Revealing this intricate structure provides insights into community function unavailable from bulk analyses. Further, imaging will provide the data required for modeling and testing predictions based on genome sequence–based information.

Challenges

Critical to addressing key scientific questions will be the novel application of existing and emerging technologies. Further, a number of scientific needs will require fundamentally new developments in imaging technology. The following are some of these needs:

- Spatial and temporal concentration measurements of signaling molecules and metabolites at the intra- and extracellular level and within populations and communities.
- Assessment of the role of extracellular polymers in community structure, function, and stability.
- Determination of macroscopic transport of water, solutes, and macromolecules and their relationship to microbial function.
- Detection and frequency of genetic exchange, recombination, and evolution within communities.
- Measurement of the elemental distribution, oxidation state of elements, and biomolecules within and among communities.
- Development and validation of models of community growth, function, and response to environment.
- Development of advanced chemical and biological probes, including engineered microorganisms, tagged biomolecules, and chemical sentinels that will help characterize microbial communities.
- Characterization of the physical and chemical properties of interfaces.
- Development of advanced tools for imaging characterization for use in the laboratory as well as in the field.

Specific examples of technical developments that would enable approaches to these challenges follow.

Specific Challenges

Biofilms and microbial communities can facilitate the formation of new solid phases (biomineralization) and interact with existing solids and with other microbes and soluble constituents in the environment. Determining the spatial distributions and gradients of elements and their redox state, as well as the distributions and gradients of molecular products (proteins and metabolites), can provide information required to better understand (1) how the microbial community and physical environment are arranged, (2) how microbes interact among themselves and with their community, and (3) the flow of energy and elements through the community and how that flow is regulated. Techniques are needed to image the complex 3D microbial communities and their environments and characterize the physical, chemical, and biological interactions occurring in them. Several currently available techniques need improvement. Some are minimally invasive or destructive, while others are noninvasive. Examples of such minimally perturbing techniques are given below.

Advanced Optical Methods

Optical spectroscopic methods can be used as tools for noninvasive characterization and monitoring of dynamic behavior. Measurements of absorption and in vivo fluorescence can be used to monitor the presence and relative concentration of optically active biochemical species. Light-scattering spectroscopy can probe the size distribution of structures in a community. Vibrational (infrared and Raman) spectroscopy is a technique for studying the composition of biological materials without perturbing or labeling the sample. Biological components (e.g., lipids, proteins, nucleic acids, and carbohydrates) and

biofilm and microbial surfaces (e.g., minerals and polymers) have unique vibrational spectra based on their chemical structures. The use of these methods will provide new information on the following:

- Large-scale (1- to 10-µm) biochemical organization.
- Composition and distribution of extracellular polymer matrices.
- Concentration and distribution of nutrients, metabolic proteins, signaling molecules, and other macromolecules.
- Interactions of biofilms and microbial communities with supporting surfaces.

Because vibrational spectromicroscopy is non-invasive, it can be performed on dynamic living systems in combination with other techniques. If synchrotron radiation is used as the photon source, a dynamic system can be studied directly on surfaces of geologic materials (Fig. 4.1).

Significant progress has already been made using confocal and two-photon fluorescence microscopy. The specificity of these techniques is provided by the exogenous chromophore. The resolution is on the order of a micron and is slightly higher for two-photon than for confocal microscopy. Delivering chromophores to remote regions is a particular challenge. Additionally, the identification of probes that maintain activity in diverse environments is required (Fig. 4.2).

All these techniques can be used in an imaging arrangement to monitor changes in community behavior in real time. Improvements are needed in such areas as spatial resolution, the ability to provide quantitative information, and data-acquisition speed. Additionally, advanced light microscopy techniques such as surface-plasmon resonance, surface-enhanced Raman spectroscopy, second-harmonic-generation imaging, optical-coherence tomography, and coherent antistokes Raman scattering microscopy can be developed for high-resolution 2D and 3D mapping, often with specificity to particular components associated with the imaging technique.

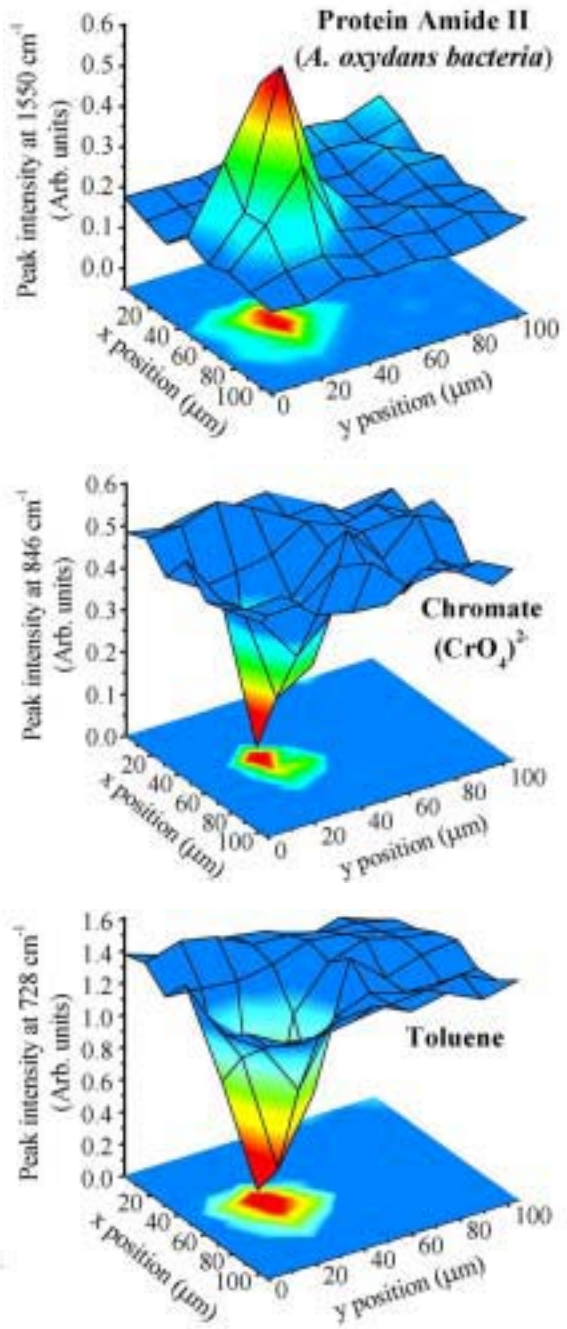


Fig. 4.1. Synchrotron infrared images showing a small colony of natural bacteria. A common organic contaminant, toluene (bottom), is used to accelerate the reduction of a carcinogenic form of chromium [chromate $(CrO_4)^{-2}$, middle] to its environmentally safe form. Bacteria are located via their spectral signature (top). [Source: H.-Y. Holman et al., "Real-Time Characterization of Biogeochemical Reduction of Cr (VI) on Basalt Surface by SR-FTIR Imaging," *Geomicrobiology Journal* **16**(4), 307–24 (©1999). Reproduced by permission of Taylor & Francis, Inc. (www.routledge-ny.com)]

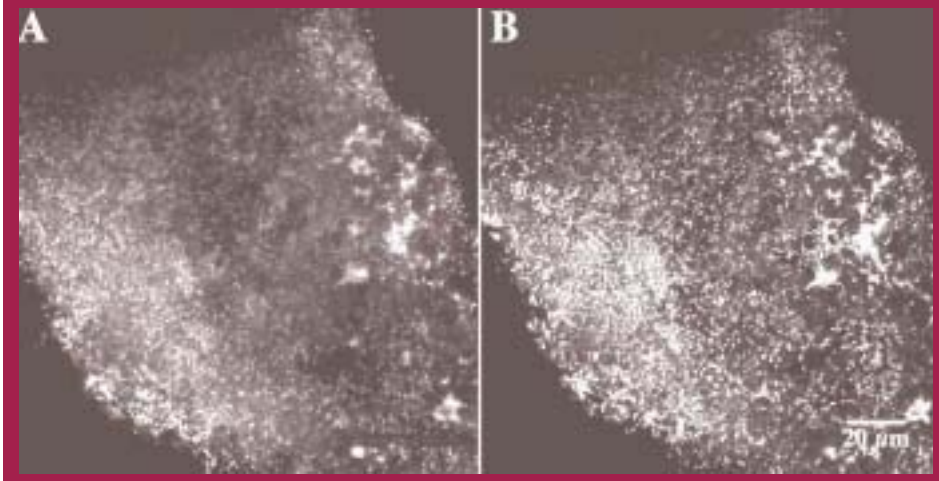


Fig. 4.2. : One-photon (left) and two-photon images of the same location and depth, showing a laboratory-grown biofilm stained with a nucleic acid–specific fluorochrome. Note the higher resolution achieved by two-photon excitation. [Source: Thomas R. Neu, UFZ (Centre for Environmental Research), Magdeburg, Germany]

Nuclear Magnetic Resonance (NMR)

Energetics and cell communication in a microbial community are governed by the flow of nutrient, waste, and signaling molecules to and from the cells. For numerical simulations, transport rates are often treated as constants even though they are known to vary with biofilm density. Imaging techniques are needed to image substrate and metabolite concentrations, biofilm density and permeability, and mass transport in biofilms and the surrounding (or supernatant) medium. Such techniques must be able to measure transport everywhere in the sample at the highest attainable resolution and under both quiescent and flowing conditions. They should be noninvasive and employ no added tracers. Furthermore, the measurements should be applicable to samples with special environmental requirements such as anaerobic systems. Finally, the techniques should allow imaging and comparison of the various parameters mentioned above on both a small sample (ideally a single microbe) and a larger microbial community and in their natural environments such as soils.

Nuclear magnetic resonance imaging (MRI) methods appear to meet many of these requirements. Existing MRI contrast techniques provide a spatial mapping of the sample: (1) structure via its hydrogen density; (2) molecular physical environment via NMR *spin-lattice* and *spin-spin* relaxation times; (3) chemical content via spectrally selective and chemical-shift imaging methods; and (4) fluid transport information via flow and diffusion imaging, all without opacity losses.

Fluid flow and diffusion imaging have been successfully applied to map convective and diffusive transport in biofilm and surrounding media. Improved NMR flow-mapping methods might integrate velocity and diffusion mapping, intracellular-extracellular water-exchange measurements, biofilm-density mapping, and metabolite-concentration mapping for comprehensive analysis of transport in planar biofilms and granular communities. MRI also can be combined in one instrument with optical confocal imaging to provide simultaneous data from both modalities.

Challenges for applications of this technique for in situ microbial studies include the development of novel MRI equipment optimized for studies of microbes in their natural environment; enhancement of NMR sensitivity; improvement of the spatial, spectral, and temporal resolution of the images; and development of multivariate data-processing methods.

Novel 3D Imaging Techniques

There are a variety of 3D imaging techniques, each with its own unique set of strengths and weaknesses (i.e., information type, contrast mechanisms, sample geometry requirements, spatial resolution, sensitivity, and probes). These techniques meet many but not all of the information requirements for a comprehensive characterization of microbial communities. Additional imaging techniques that provide complementary information—for example, acoustic imaging—are required to fill in the gaps. Further, resolution

enhancement (where feasible) is needed for some existing imaging methods to properly apply them to the fine scale of bacterial cells.

One approach is to identify and integrate complementary imaging modalities to attain a synergistic multimodal imaging capability. To support such combined imaging technologies, image-fusion methods for combining diverse image data with differing spatial resolutions are needed. An example would be optical fluorescent methods combined with MRI. A second approach is to study a specific microbial process with all available imaging techniques, combined with comprehensive data analysis using advanced computational techniques. This integrated method should improve understanding of the links between gene sequencing of individual bacterial cells and biofilm functioning in the environment.

Microsampling Techniques

In some cases, relying on direct determination of spatial and temporal distributions of chemical and physical properties will be necessary. Reasons include large background, lack of an appropriate contrast mechanism, lack of a probe or contrast agent, and unacceptably large perturbations to the system caused by the agent and low levels of species to be mapped. In these instances, instruments are required that can provide detailed, high-sensitivity maps of chemical species and physical properties.

Effective microsampling techniques must interface with analytical techniques in studying microbial communities. Microsampling can be used to precisely extract microorganisms or chemical samples for subsequent separation and analysis. Various analytical techniques can provide biosignatures. Advanced analytical tools that provide rapid chemical and biological information are needed. This combination of microsampling with analytical techniques will allow the imaged microbial behavior and functions to be related to specific metabolites, protein profiles, and genetic materials of microbial communities.

Advanced Chemical and Physical Sensors

Sensors, by definition, are instruments that detect and quantitate one or more agents or properties in the presence of the surroundings. To directly address the large issue of background, current-generation sensors either exploit a unique signal or characteristic or, more commonly, restrict the spatial region or separate the desired agent from its milieu.

Future sensors, to be of use in the characterization of microbial communities, must integrate the spatial segregation-separation function with the detection-quantitation function; therefore, a new generation of instrumentation must be invented to handle sophisticated sequential and parallel chemical manipulations simultaneously with ultrasmall sample quantities. Critical functions in microbial communities most likely will be implemented with signaling molecules that may be present at small copy numbers within any given cell. Thus, new-generation sensors must be capable of previously unavailable limits of detection. Furthermore, even though current-generation sensors address at most a few analytes, massively parallel measurement strategies must be invented to simultaneously measure a large number of different species with high spatial and temporal fidelity.

Finally, it will be important to understand how sensors affect the properties of the microbial community, especially physical sensors whose contact with the community may be part of the measurement operation. When measured against these requirements, the chemical and physical sensors of today are slow, insensitive, poorly integrated, and capable of very limited sample manipulation. A key goal will be to advance the state of the art among these figures of merit to make significant contributions to understanding microbial communities.

5. Computational Infrastructure for Imaging

There is a strong need for computational and informatics techniques that can meet the imaging challenges in data analysis, modeling, and distributed access and archiving. We expect that model systems and field studies will generate massive amounts of multi-spectral spatiotemporal image data at multiple scales. Figure 5.1 shows examples of different data modalities (imaging technologies) that may be used alone or as complements for investigating living systems at different scales.

A particularly valuable goal for the future is a computing environment that allows scientists to acquire data, archive a set of data, annotate and link the data from different modalities, find relevant computational tools on the Internet, compute various representations of these data for quantitative analysis, and compare observational data to simulated data when possible.

Issues

Significant barriers to the use of computational resources for imaging studies include the following:

- Useful computational tools in various forms in laboratories around the country that are not easily accessible by community researchers.
- Diverse set of algorithms and software tools for solving a specific problem that does not characterize reliability and performance behavior.

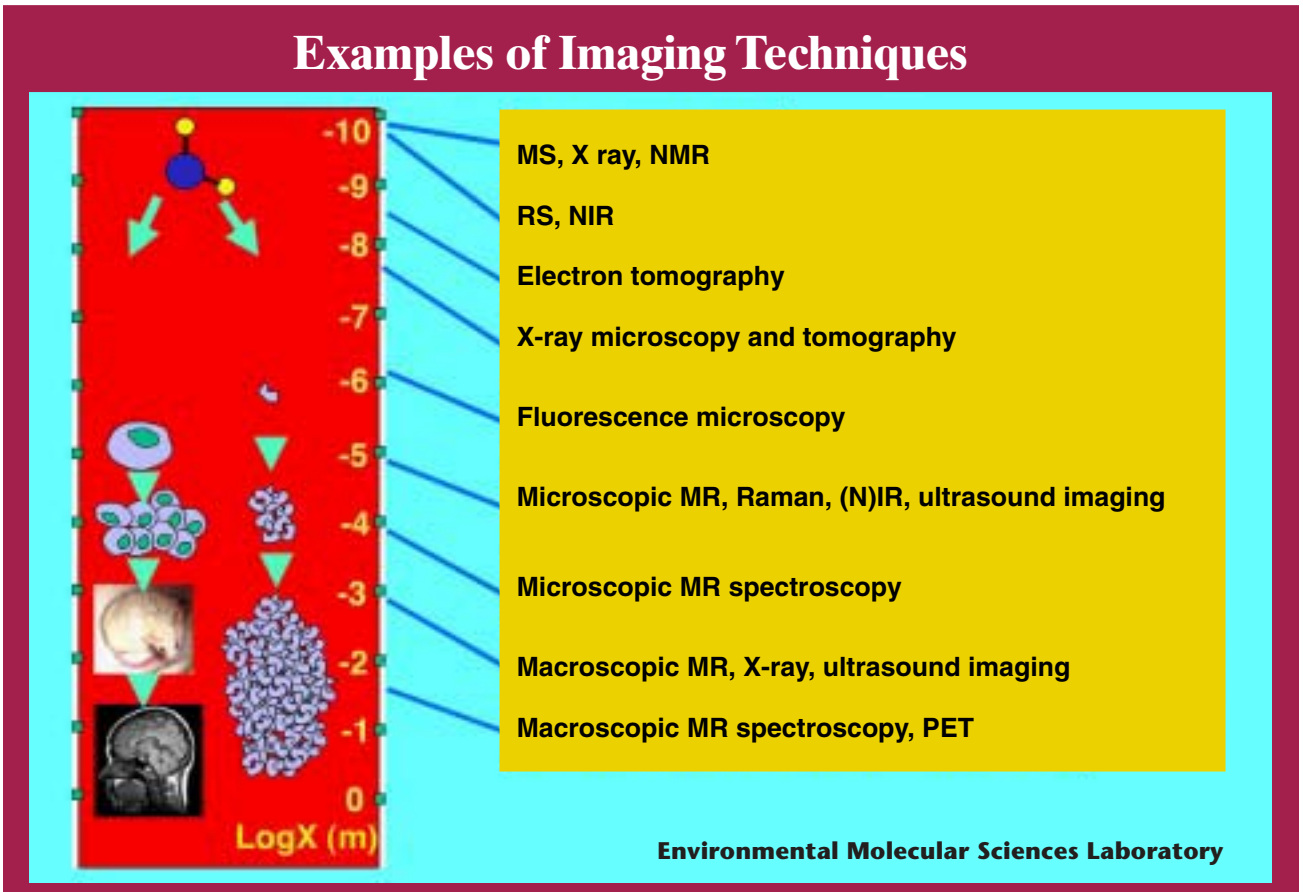


Fig. 5.1. Scales of imaging modalities. [Source: Robert A. Wind, Pacific Northwest National Laboratory]

- Lack of connectivity among instrumentation, computing infrastructure, and diverse databases.
- Lack of algorithms for modeling and inverse problem solving.
- Lack of user training in using advanced scientific tools.

In this context, we propose investments in several areas. Some of these research activities can be coordinated with the DOE Office of Science's Mathematical, Information, and Computational Sciences (MICS) office. MICS actively funds basic research in scientific data transport and analysis, information visualization, and computational physics; examples are adaptive mesh refinement with embedded boundaries and libraries based on general partial differential equations (PDE) for high-performance computing.

Technology Needs

Discussed below are technologies needed for the computational infrastructure for imaging, including data, compute, user, and infrastructure services.

Data Services

These services include data annotation, publishing, subscription, and retrieval. Data should have self-describing annotations, and users should be able to publish their data against an existing study or project topic. At the same time, users should be able to subscribe to the data and corresponding annotations generated by other researchers. There could be several mirror sites for archiving images, but users could have a private copy of their own data. The archive could be searched against user-defined annotations as well as computed features; for example, to collect all transmission electron microscopy (TEM) images of *Geobacter* in the GTL community or search for an example of biofilm at a specific state of its evolution.

Compute Services

These general computing services, including data preprocessing, data fusion, data and quantitative analysis, modeling and simulation, and visualization, should be available to users in a seamless fashion. The capabilities of each area should be integrated to allow users to move from one to the next without having to invest considerable time in learning a new method of doing the work.

Preprocessing

Image data are generated by various instruments used to study biological structures, from individual proteins to complex microbial communities. Up-front data processing is needed to improve data quality for subsequent visualization and analysis. In some cases, this preprocessing step will be specific to the type of imaging instrument used in the experiment. However, some preprocessing steps will be generally applicable to many modalities.

Filtering. Filtering algorithms are needed to eliminate or reduce noise (e.g., random and spatially correlated) or alter the frequency content of the image (e.g., smoothing, sharpening, and edge detection). In the case of image noise, different types are inherent to specific imaging instruments [e.g., speckle noise in magnetic resonance imaging (MRI) and X-ray tomography]. Filters that can be modeled after the physics of the instrument noise source have the best chance of being the most effective. Existing and new filtering algorithms should be extended to 3D so instruments capable of 3D imaging can benefit from this technology. In addition, improved 3D filters are needed that suppress or enhance morphological structures of a particular object (e.g., a cell) for subsequent segmentation and recognition purposes.

Data Reduction. A substantial imaging challenge in the GTL program is how to handle the vast amounts of image data generated during experiments. For example, a 3D multispectral image data set acquired at different times during a single experiment can be extremely large (tens of gigabytes), so the task of analyzing, storing, and sharing all this data is significant. One way to address the problem is to apply new and intelligent data-reduction technologies to retain only the data critical for subsequent fusion and analysis. Two areas that address this are image-data compression and dimensionality reduction:

- **Compression.** Technologies for image-data compression must be developed in collaboration with biologists to ensure that important

information is not compromised. New content-based data-compression technologies are needed that, for example, highly compress noncritical regions of the image using lossy compression algorithms but apply loss-less compression strategies to information-rich image regions. Fast up-front image-analysis algorithms are required to differentiate these critical image regions from noncritical regions. Other new and creative compression strategies that allow large image-data sets to be more easily stored and efficiently transmitted over the Internet are essential for accelerating GTL program science.

- **Dimensionality Reduction.** Another way to reduce the amount of image data is to convert to an alternate representation in which fewer dimensions are required to describe the important aspects. In a multispectral data set, new approaches might select only the spectral components containing information important for the application at hand. For a temporal study, this could involve change-detection algorithms that allow certain time sequences to be excluded if no significant biological activity is occurring. Other approaches would be to apply principal component–analysis strategies to the data so that only the major modes of variation would be captured, thereby reducing the data set's overall dimensions. The main goal here is to reduce the data set but ensure that it still captures the experiment's important essence.

- **Missing-Data Interpolation.** In many instances, the image-data set from a given instrument will contain incomplete or corrupt data. Examples of this follow:

  – Gaps in 3D data sets from limited-view angle tomography or MRI.

  – Missing or corrupt slices from a 3D data set (TEM, confocal).

  – Photobleaching and detector saturation in fluorescence microscopy.

Missing or corrupt data typically are found only after manual examination of image data and are then corrected using a combination of manual and semiautomated techniques. These gaps in the valid data set can create significant problems for automated data-analysis algorithms that are applied downstream for such purposes as object segmentation and quantification. New approaches must be developed that (1) automatically detect these missing or corrupt regions of data and (2) either flag data so that it is ignored in any downstream processing or accurately interpolate new data to replace gaps.

Data Fusion

Biological image data can be obtained from many sources or modalities such as Raman spectroscopy, confocal microscopy, and MRI. Using these different modalities, image data of the same subject can be obtained from multiple sensors. These data sets could have vastly different, often complementary, content. To fully exploit information distributed among data sets, techniques for data fusion will be necessary. To effectively combine individual data-set information, relationships between image-data aspects relating to the subject's physical attributes should be defined for the various image sources. Once these relationships are defined, they can then be used to map all images to a common representation (i.e., coregister the images to a common scale and orientation). Registration can be based on landmarks and include global or local warping of images obtained from different modalities. Global alignment often involves estimating parameters of an affine transform. Localized alignment corrects for localized distortion typical of many scientific instruments.

Data and Quantitative Analysis

This includes object detection, segmentation, motion tracking, feature-based query, matching and comparative analysis, multivariate data analysis, and classification. The problem's important aspect is the need to extract meaning from the image data. Many times this part is overlooked and left to users to determine their own approaches to get meaningful results. Data and quantitative analysis should be integrated in the overall system as well as in other areas. However, research often is still needed to develop these "tools" to help users quantify their data.

Object Detection. Counting is often necessary to detect the presence of objects of interest, record their position, and learn something about the underlying organization. Object detection can be based on simple correlation to a more elaborate model- and appearance-based approach.

Segmentation. A critical aspect of image analysis is segmentation (Fig. 5.2), which aims to delineate objects of interest and their individual compartments so that morphological features, protein expression, and physiological responses can be characterized on a cell-by-cell basis from each image and collection of such images. An important aspect of segmentation is its robustness and tuning for parameters. A new algorithm may need to be developed for a specific data set. Current computational techniques, however, appear suitable for a large class of fluorescent imaging microscopy. Interactive techniques (e.g., "snakes") will be provided for processing small sets of data.

Motion Tracking. Time-lapse microscopy has the potential of producing massive amounts of data (Fig. 5.3). A typical time-varying 2D experiment can generate 12 MB of data per minute at 30-s sampling intervals (1- by 1-k images at 12 bits per pixel at three different spectrums). Bacteria need to be detected and tracked. The ability to track a community of bacteria enables a more basic understanding of pattern formation and organization under a number of different conditions.

Feature-Based Query. Searching an image database against a reference texture and intensity often is desirable. For example, the fluorescence texture of a biofilm indicates its evolution and may be used as a template texture of interest to search for similar textures that may have other properties.

Matching and Comparative Analysis. Experiments are often designed to compare function, physiology, and behavior under different experimental conditions for same species, genetically modified species, or different species. Techniques must be developed for comparing experimental data in a systematic way, which may include sensitivity analysis.

Multivariate Data Analysis. As technology advances to interrogate complex biological systems, images often contain more than one variable of interest and are known as multivariate data. This is in contrast to univariate images in which only one variable, for instance a spatial map of Ca^{++} ion concentration, is obtained. Examples of multivariate data are multispectral and hyperspectral images such as Fourier transform infrared spectroscopy (called FTIR) and MRI. Data sets or images combined with temporal and environmental parameter information (i.e., temperature and treatment) also can be multivariate.

Multivariate data sets are often enormous, containing more than tens of thousands of pixels. A researcher interested in determining qualitative and quantitative relationships in the data can no longer simply look for the answers visually. Instead, sophisticated multivariate analysis algorithms are needed to extract useful information

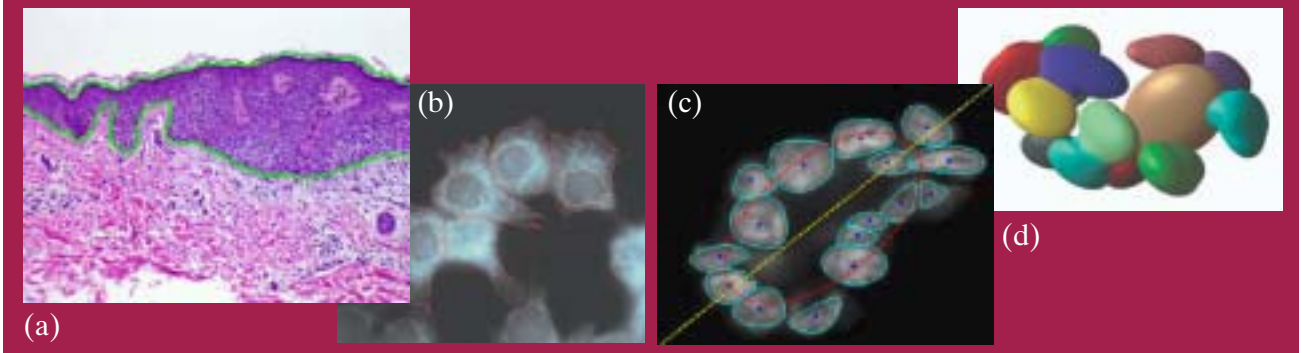


Fig. 5.2. Segmentation result for (a) basal cell carcinoma, (b) nuclei in cultured colonies, (c) multicompartment segmentation, and (d) 3D cultured colony imaged with confocal microscopy. [Source: Bahram Parvin, Lawrence Berkeley National Laboratory]

from multivariate image data. These algorithms use the variance in these huge data sets to determine relationships among variables and generate image-contrast and quantitative-concentration maps. Some very powerful multivariate algorithms such as principal-components analysis, factor analysis, least-squares regression, and multivariate curve resolution need to be applied, optimized, and automated for biological data and images. In addition, new, robust algorithms must be developed to more fully characterize the data in a quantitative fashion. These algorithms may include improved methods to remove signals from contaminants, background variation, and instrument or experimental artifacts, all of which are commonly found when cells and signaling events are probed in situ and in vivo.

Classification. Classifying groups of cells and structures, ideally in real time and in multicellular, heterogeneous systems is important. For example, there is a need to classify bacterial cells based on type while monitoring their behavior in culture. Although techniques such as hierarchical cluster analysis and discriminant analysis have been shown to provide classification of biological tissues and species, further advancement in this area is needed. Techniques available from other disciplines may be well suited for biological-data classification. These as well as novel algorithms should be investigated and adapted for real-time image classification of cellular species and events.

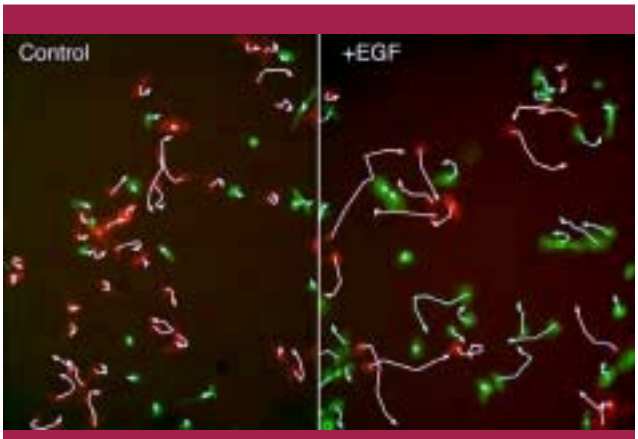


Fig. 5.3. Control cells (green) and engineered autocrine cells (orange) mixed overnight and followed for 4 h by two-color fluorescence time-lapse microscopy. Shown are the initial images overlaid with trajectories marked at 40-min intervals. Cells were either in the absence (left) or presence (right) of 2-nM exogenous growth factor (EGF). [Source: Reproduced from the *Journal of Cell Biology* (2001), **155**(7), 1123–28 by copyright permission of The Rockefeller University Press.]

Modeling and Simulation

Providing a modeling and simulation component to analyze, interpret, and predict biology data and experiments represents an approach to helping make biology a more quantitative science. Modeling is the mathematical definition and algorithmic implementation of a biological process. Simulation is a computer realization of a biology experiment. Taken together, modeling and simulation constitute an approach to doing "virtual biology" experiments. A modeling and simulation capability can take many different forms, from simple reaction network models to full temporal and spatially resolved models.

Image data from biology experiments are useful for the modeling and simulation community as a source of data for parameter estimation (e.g., chemical reaction rates and diffusion coefficients), initial and boundary conditions (e.g., velocity and temperature), geometric feature data, model-verification data, and simulation-validation data. Different types of image data from a single biology system may be combined to form the basis of a biology modeling and simulation "virtual experiment." Some examples follow:

- Nuclear magnetic resonance (NMR) and spectral imaging techniques show chemical reaction rate data.
- NMR data can provide distributed maps of component velocity and diffusion coefficients.
- Geometric surfaces defined by feature extraction and reconstruction algorithms from volume image data form a source of geometry on which volume meshes can be used for discrete time- and space-dependent integration methods including finite elements, volume, and difference algorithms.

The sources of this type of image data include confocal microscopy, NMR, X ray, and electron tomography.

Figure 5.4 represents the series of steps to transform a 3D digital volume image (made up of a stack of 2D TEM images) into a computed 3D geometric mesh.

Visualization

Improved data- and information-visualization techniques will be necessary to assist researchers in their overall understanding of the biological system (Fig. 5.5). A number of systems already can accommodate 3D volume data with reasonable speeds. Data transfer among visualization methods will be important to help researchers present their ideas.

Where appropriate, visualization systems should be used and seamlessly incorporated into the computational environment. Although volume rendering is a standard tool for viewing the quality of raw 3D data, surface rendering often is used with segmentation to evaluate the quality of automated analysis. Triangulation is appropriate for visualization, but it is not the tool of choice for 3D geometric analysis of surfaces.

User Services

This is a general classification of services for software training, design, and availability that need to be provided at all levels of implementation. These services include documentation and tutorials, open-source software, self-describing data, location of application-specific data-analysis tools, and remote instrumentation (Figs. 5.6 and 5.7).

Documentation and Tutorials

All services should be documented with tutorials and examples to enable users to directly access a resource. Workshops must be established to learn more about user needs. Experience indicates that one-to-one interaction is often necessary for successful application of a particular tool. Part of the workshop series should include classes with homework and exercises that will be modeled after the current coursework taught for comparative genomics.

Open Source

All computational tools developed for GTL should be open source to enable users to process their data on the desktop.

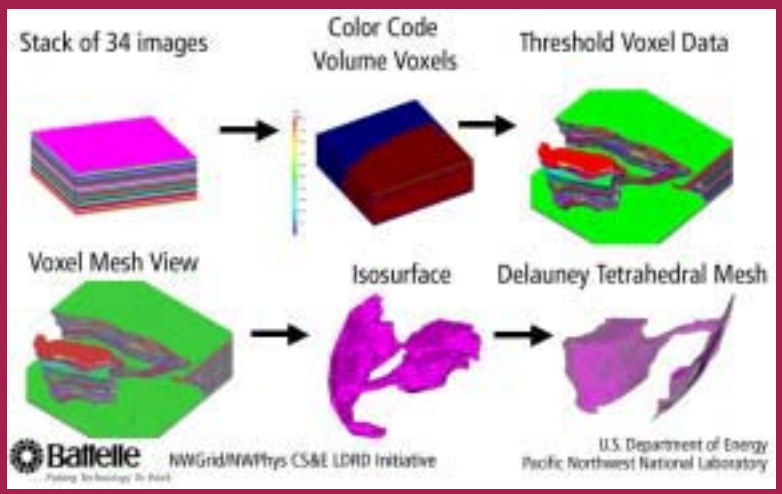


Fig. 5.4. Computational biology 3D reconstruction of a mitochondrion crista. [Source: Harold E. Trease, Pacific Northwest National Laboratory, University of Connecticut Health Center data]

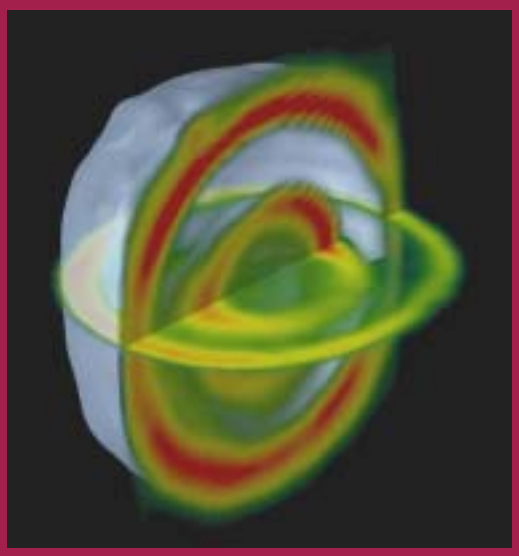


Fig. 5.5. Example of 3D volume visualization of a tomographic magnetic resonance imaging data set of a *Xenopus laevis* frog oocyte after heat shock. [Source: Robert Wind, Pacific Northwest National Laboratory]

Self-Describing Data

Data should be annotated with a markup language so Web pages can be generated automatically from user-defined descriptions. The markup language will simplify exchange and distribution of data among multiple institutions.

Locating Application-Specific Data-Analysis Tools

Stepping through a large data set and processing it for quantitative analysis is often necessary. An example is quantifying protein expression in cells that reside at a specific location in tissues. A typical study may include several hundred images with 5 to 100 cells per image. The user should be able to locate an appropriate analysis tool or set of tools based on this level of problem statement.

Remote Instrumentation

Remote instrumentation satisfies two requirements: (1) it enables researchers to share a unique instrument for data collection, and (2) it couples data acquisition with computation and data services to streamline the process.

Infrastructure Services and Computing Resources

The proliferation of imaging methods and analytical approaches has created two significant challenges:

1. Current approaches for analyzing and interpreting image data are still inadequate to meet the special challenges associated with the complex multisensor data sets typical of advanced applications.

2. The numerous analytical-software tools resulting from this research generally are not compatible with each other, a situation that hinders collaboration and more rapid progress—

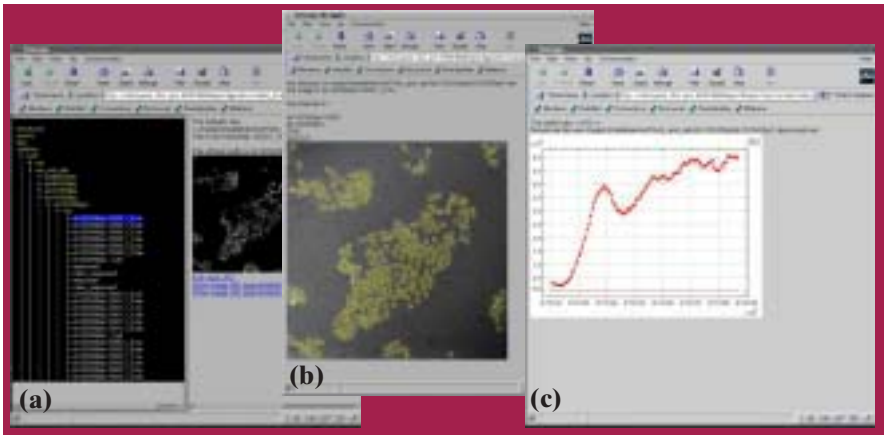


Fig. 5.6. Automated uptake analysis for a population of cells: (a) image and corresponding data model, (b) active page with segmented result, and (c) time-series response of a selected cell from (b). [Source: Bahram Parvin, Lawrence Berkeley National Laboratory]

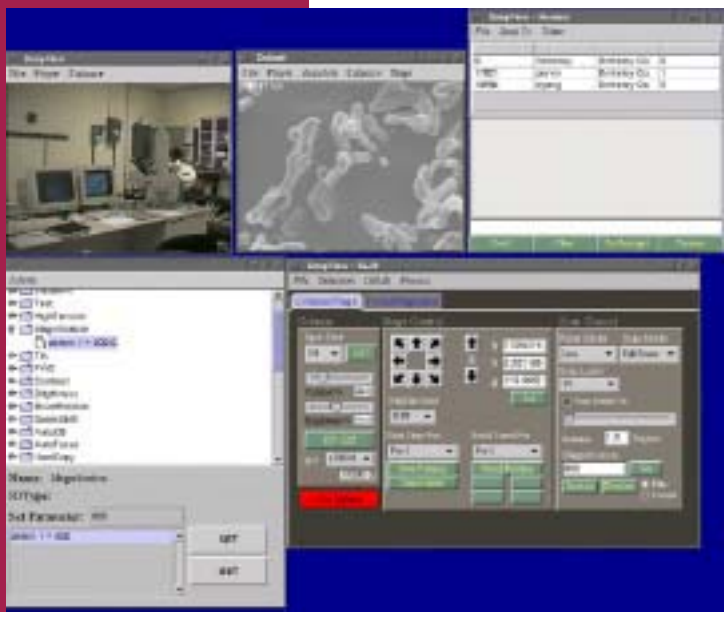


Fig. 5.7. An example of remote instrumentation for shared collaborative use—the DeepView architecture for remote microscopy funded by MICS. DeepView provides a self-describing framework for defining scientific instruments. This example shows DeepView's front end to a scanning electron microscope at Oak Ridge National Laboratory. [Source: Bahram Parvin, Lawrence Berkeley National Laboratory]

especially across domain boundaries, where radical scientific advances often occur.

Development of a flexible, extensible "architecture" for image-processing and -analysis tools is needed to help researchers collaborate more effectively and gain greater insight into the imagery they study. This architecture would provide a framework of standards, protocols, and basic utility functions into which new and innovative image-analysis methods could easily be placed and compared with other techniques. The architecture will allow access to heterogeneous databases, leverage current efforts funded under the Scientific Discovery through Advanced Computing (called SciDAC) program, and provide scalable evolution for meeting new scientific requirements (Fig. 5.8).

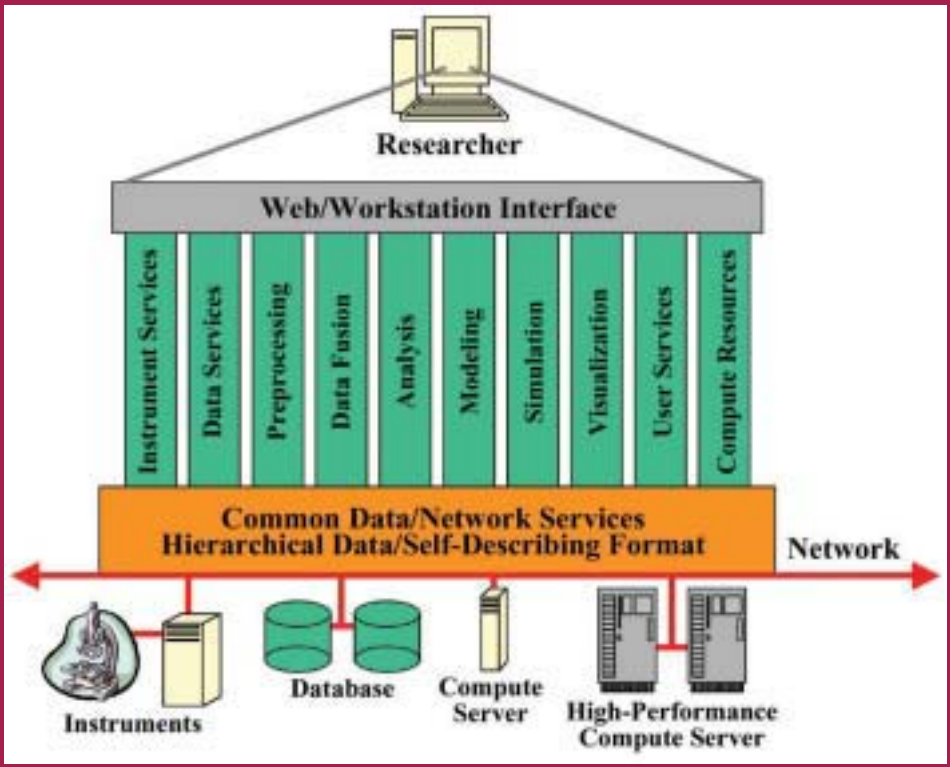


Fig. 5.8. Conceptual vision of the infrastructure for imaging. [Source: Donald R. Jones, Pacific Northwest National Laboratory]

6. Probe Development for Advanced Imaging Methods

Key to all aspects of the GTL program is development of improved and new probes. Fluorescent probes are key reagents for extracting information from cells and tissues. Their use has expanded greatly over 30 years, and a billion-dollar industry is based on fluorescent detection. Approaches for probe development, current status, and needs for new probes are discussed below. Fluorescence technologies contain the most mature examples of probe development at this time, but new probes are needed for other imaging technologies, including infrared (IR), Raman, surface-enhanced Raman particles (SERS), coherent antistokes Raman scattering (CARS), X ray, and magnetic resonance.

The ideal probe for monitoring molecular machines and biological functions has several characteristics. First, it should provide a robust signal that is easily observable above any background present in the sample. Second, it should be unaltered by the means through which it is observed and by the sample into which it is introduced. The probe should be highly selective in recognizing the intended target and have a high mobility to the regions of interest. It should be optically stable, allowing the dynamic processes associated with molecular machines to be monitored over long time scales. Finally, it should be small enough to be easily incorporated without perturbing the system being studied.

Having defined the ideal probe, we must recognize the trade-offs with current probes. For example, organic chromophores are small and can be easily conjugated to a number of biological molecules. However, even the most robust organic chromophores will emit photons for only a few seconds before photobleaching when saturated. At the single-molecule level, this severely limits the amount of time that dynamic processes can be monitored. On the other hand, plasmon-resonant particles (see below) are extremely bright and do not photobleach; however, their large size (40 to 100 nm) prevents their use in a number of systems.

In addition to optical probes, new density probes are needed for use as fiducials in tomography and as high-resolution labels for individual protein subunits in complexes being characterized or located by transmission electron microscopy (TEM) or X-ray microscopy. These should have high density, chemical specificity, and minimal nonspecific binding.

Applications for Fluorescent Probes

This section addresses the wide range of current and possible applications of fluorescent probes (Fig. 6.1).

Labels and Indicators

Labeling

- DNA and RNA for fluorescence in situ hybridization (FISH), antibodies for immunofluorescence, ligands for drug discovery, and intracellular proteins for studying cell function.

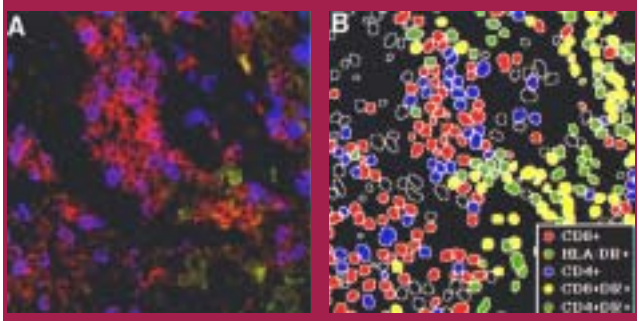


Fig. 6.1. FISH using multiple fluorescently labeled DNA probes. [Source: A. I. Dow et al., "Automatic Multiparameter Fluorescence Imaging for Determining Lymphocyte Phenotype and Activation Status in Melanoma Tissue Sections," *Cytometry* **25**, 71–81. ©1996, Wiley-Liss, Inc. Reproduced with permission of John Wiley & Sons, Inc.]

- Immunofluorescence, diagnostics, gene expression, and DNA sequencing.

Live-cell indicators

- **Ion concentrations.** pH and calcium are the main targets. Most probes for other ions need improvement. Long-wavelength and wavelength-shift probes for ions and other physiological parameters are needed (Fig. 6.2).

Gene expression: Beta lactamase, fluorescent proteins, and fluorescein digalactoside

- **Membrane potential.** Some membrane-permeant probes respond too slowly to determine vesicle and cell potentials. Rapid-responding probes permit detection of action potentials in neurons. Current probes show relatively small changes in fluorescence response to changes in membrane potential. Improving the fluorescence response would make these probes more useful. Probes that become brighter with integrated neural activity are needed.

Cell and vesicle tracking

There is a need for near-IR fluorescent tracking dyes that can be used deeper in complex multicellular systems as well as protein-protein interaction and conformational changes. This exciting new area is beginning to grow with the use of multicolor fluorescent proteins that can be incorporated into the design of biosensors for kinases and other regulatory protein activities. There is a strong need for chemical and genetic methods of introducing multiple fluorescent labels into specific sites on specific proteins so that new cellular sensors can be constructed (Fig. 6.4).

- **Cell- and mitochondrial-energetics indicators.** Other than NADH (nicotinamide adenine dinucleotide), FAD (flavin adenine dinucleotide), and membrane-potential probes, there are no good indicators of energization. Such indicators would be valuable for studying mammalian cells and populations of microorganisms. The need for improved redox potential indicators should be included in this category.
- **Enzyme-activity indicators.** Few probes can indicate the wide range of enzyme activities in cells.
- **Protein-concentration sensors.** Targetable biosensors are needed to indicate the concentration of different regulatory proteins in cells and in the interstitial spaces between cells.

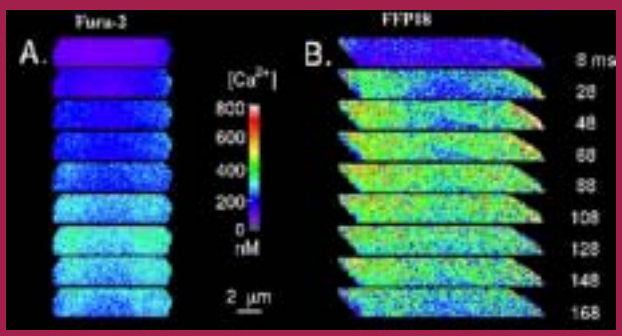
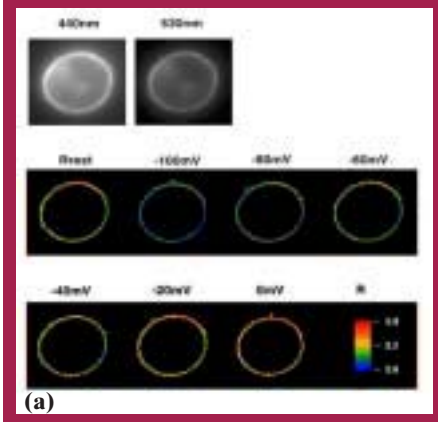


Fig. 6.2. Intracellular calcium in smooth muscle cells imaged using a cytosolic fluorescent calcium indicator Fura-2 (A) and the near-membrane calcium indicator FFP18 (B). Cells were depolarized at time = 0, and images were recorded every 20 ms. FFP18 reveals localized calcium transients that are much larger and faster than those seen with Fura-2. [Source: E.F. Etter et al., "Near-Membrane [Ca^{2+}] Transients Resolved Using the Ca^{2+} Indicator FFP18," *PNAS* **93**, 5368–73 (1996). ©1996 National Academy of Sciences, U.S.A.]



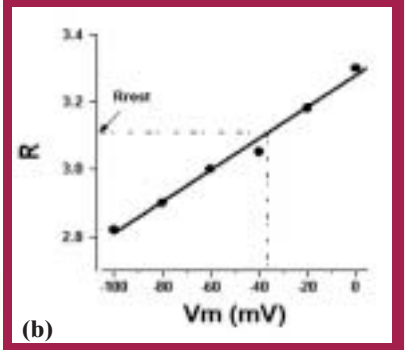


Fig. 6.3. (a) Fluorescence of Di-8-ANEPPS in lipid vesicles from excitation at two alternate wavelengths (440 nm and 530 nm). (b) The ratio of emission intensity is linearly proportional to membrane potential. [Source: J. Zhang et al., "Membrane Electric Properties by Combined Patch Clamp and Fluorescence Ratio Imaging in Single Neurons," *Biophysical Journal* **74**(1), 48–53 (1998)]

- **Additional cell and tissue properties.** Probes in need of improvements include those for lipid-protein interaction, nitric oxide, pressure-strain, oxygen, and cell health.

- **Membrane dynamics.** Improved probes are required to study dynamics of lipid-protein interactions in membranes (Fig. 6.5).

- **Single-molecule studies.** Ways to tether and orient the probe at particular sites are needed.

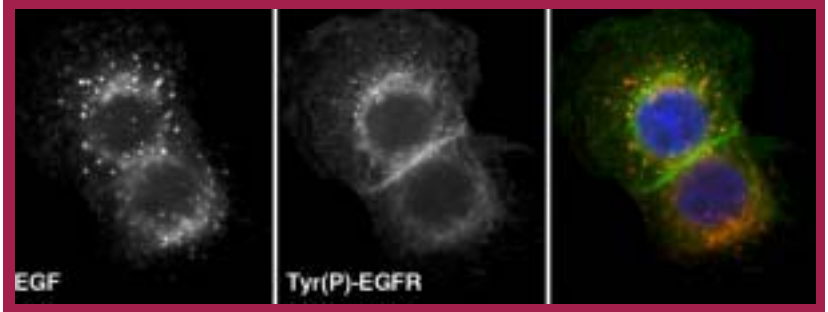


Fig. 6.4. Localizing sites of receptor activation by a juxtacrine ligand. Cells expressing EGF receptors (EGFR) were engineered to also express membrane-anchored EGF. The cells were fixed and stained with an Alexa488-labeled antibody to EGF (left panel) and affinity-purified antibodies to the activated EGFR (middle panel). Strongest receptor activation was seen at sites of cell-cell contact and transport vesicle-containing ligand. [Source: H. Steven Wiley, Pacific Northwest National Laboratory]

Philosophy of Probe Development

A need for specific biological information drives probe development, which entails an understanding of cellular and biochemical systems, instrumentation, and chemistry. The probe chemist uses the theoretical principles of fluorescence and chemistry to design a core chromophore-fluorophore, the core light-absorbing fluorescent component of the probe. The design also must incorporate chemical or genetic modifications that will permit the core fluorophore to function as a probe. At the same time, the probe chemist must be knowledgeable about instrumentation that will be used to extract the probe's information in the biological environment so that it functions optimally. The probe is then ready for testing and use.

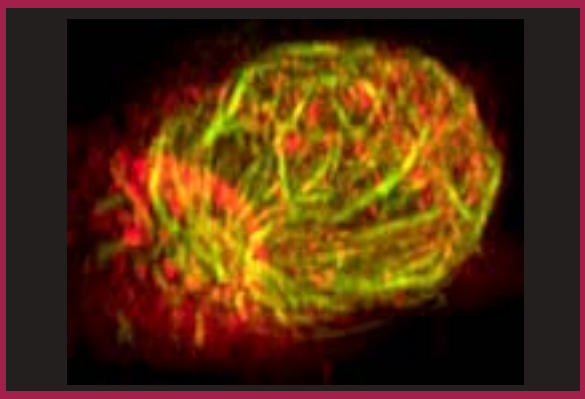


Fig. 6.5. A cytotoxic T cell (CTL) engaging an antigenic target immunostained for tubulin (green) and LFA-1 (red). The target cell in this image is not visible because its microtubules were depolymerized before mixing with CTLs. This image shows the dynamic clustering of LFA-1 into a ring at the target contact site. [Source: Reprinted from J. R. Kuhn and M. Poenie, "Dynamic Polarization of the Microtubule Cytoskeleton During CTL-Mediated Killing," *Immunity* **16**, 111–21 (©2002), with permission from Elsevier Science]

Fluorescent Probe Design: Integration of Fluorescence Principles

The power of fluorescence detection lies in the various fluorescence "tricks" that can be used to obtain molecular information. The probe designer and the applications and imaging scientists must be aware of the principles below so probes are properly designed and used (Figs. 6.6 and 6.7).

- Simple signal detection from probes in cells and tissues
- Single-molecule detection
- Multicolor detection
- Energy transfer
- Polarization
- Fluctuation spectroscopy
- Surface-resonance effects
- Time-delay fluorescence detection
- Combinations of the above

Fig. 6.6. Multicolor staining for different proteins (CD8, CD4, and HLA) allowing different cells to be distinguished in a mixed population. [Source: Thomas Ried, National Institutes of Health]

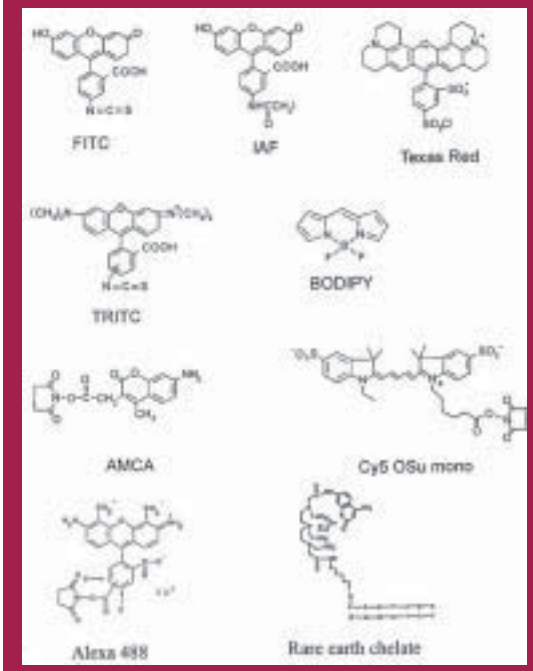


Fig. 6.7. Structures of a number of common fluorophores. [Source: Alan Waggoner, Carnegie Mellon University]

Fluorescent Probe Design: Design of the Core Fluorophore

Brightness

Brightness of the fluorophore is essential to overcome background and noise. Generally, the brightness is the numerical product of the light-absorbing power (extinction or cross section) and the quantum yield. We already have some visible and UV fluorophores that are within a factor of 2 to 3 of their maximum potential.

However, we need more of these bright probes that span the range of useful wavelengths and whose properties (e.g., photostability and environmental sensitivity) are advantageous for the development of sensitive indicators (see Table 6.1). One of the important features of fluorescent probes is their brightness, which is a function of their extinction coefficient and quantum efficiency, shown in the table for some commonly used probes. As can be seen, a number of these probes are near their theoretical maximum. Other problems needing to be addressed include probe-probe quenching and the presence of other quenching mechanisms. Development of fluorophores with large two-photon absorption cross sections is important but has not received much attention.

Wavelength

Probe absorption must be matched to excitation wavelengths. Overlap of fluorescence peaks in multicolor experiments is limiting. Although quantum dots have narrow emission peaks, smaller organic dyes with narrow peaks also are needed.

Photostability

Photostability is important in most applications. Bright dyes that quickly fade when illuminated are inherently less useful. For many high-throughput applications, the photostability of current probes is adequate. For single-molecule detection and optical sectioning of cells and tissues, however, improvements in photostability are required. Photostable near-IR fluorophores are particularly needed to extend the range of multicolor fluorescence detection and to provide probes for improved deep imaging in complex multicellular systems (Figs. 6.8–6.12).

This parameter is important in probe design to optimize excitation and light collection while permitting multicolor detection. Different applications have different requirements.

Fluorophores with excited state lifetimes in the 10- to 100-ns range would be useful for autofluorescence and background reduction yet would produce high replication rates for signal acquisition.

Table 6.1. Brightness of Fluorescent Labels

Dye	Extinction	Quantum Yield	Brightness (kilobrights)	Percent of theoretical maximum
Fluorescein	70 k	0.70	49	70%
Alexa	80 k	0.80	56	80%
Cy3	150 k	0.04	6	4%
Rcy3	150 k	0.80	120	80%
Cy5	250 k	0.300	75	30%

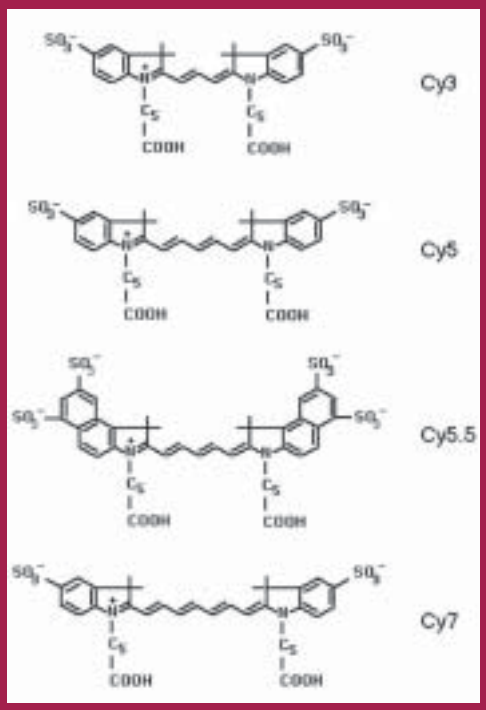


Fig 6.8. Series of cyanine dyes with an increasing number of conjugated double bonds. As can be seen in the next figure, excitation and emission wavelengths increase with the number of conjugated double bonds. [Source: Alan Waggoner, Carnegie Mellon University]

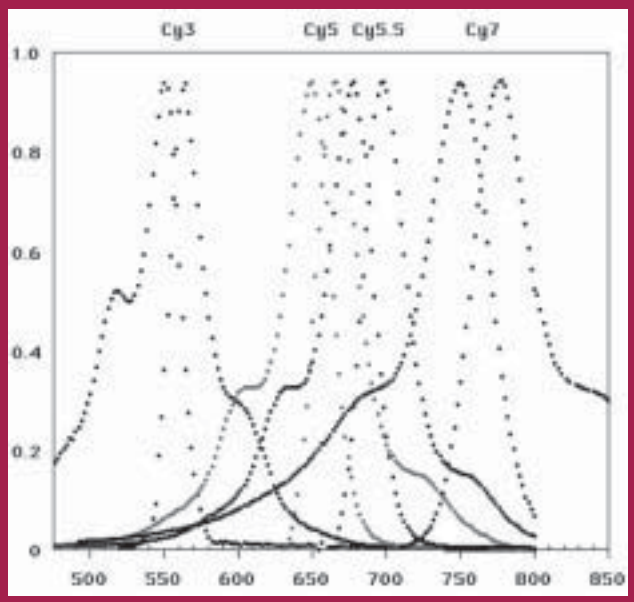


Fig. 6.9. Excitation and emission spectra shown for Cy3 through Cy7. [Source: Alan Waggoner, Carnegie Mellon University]

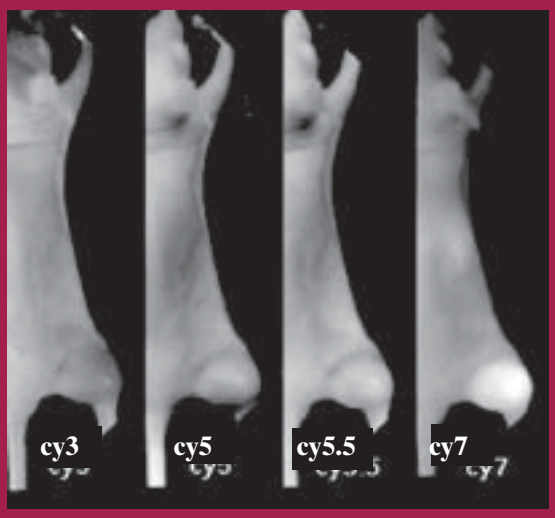


Fig. 6.10. Subsurface cancer shown as more easily detected by the antibody-conjugated Cy7 dye than other Cy dyes because the emission properties better match the tissue's optical window. [Source: B. Ballou et al., "Cyanine Fluorochrome-Labeled Antibodies In Vivo: Assessment of Tumor Imaging Using Cy3, Cy5, Cy5.5, and Cy7," *Cancer Detection and Prevention* **22**, 251–57 (©1998). Reprinted with permission from Elsevier Science]

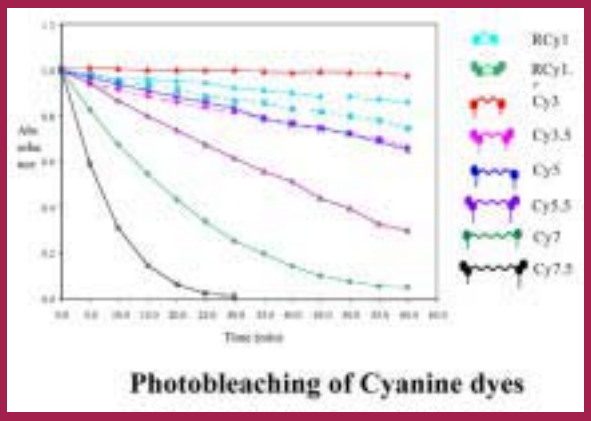


Fig. 6.11. Photostability of a series of cyanine dyes. [Source: Alan Waggoner, Carnegie Mellon University]

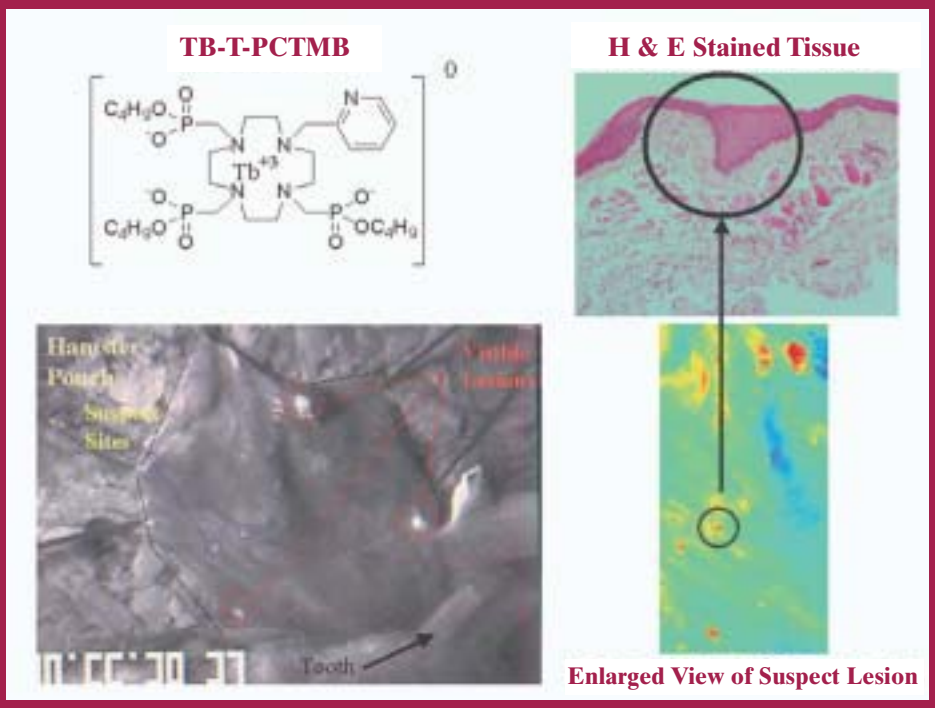


Fig. 6.12. Large spectral shifts (~260 nm) possible with lanthanide chelates because absorption is due to antennae attached to the macrocycle holding the metal, while metal ions provide fluorescence. Images of the hamster cheek pouch show that the pyridyl butyl–ester complex of Tb^{3+} has the potential to detect early-stage cancer lesions on tissue surface using simple light sources and direct visualization. [Source: Unpublished results from the Bornhop Research Group at Texas Tech University]

Environmental Sensitivity

Fluorescent labels are optimal if they are environmentally insensitive. Physiological indicators often require sensitivity to charge distributions, viscosity, solvent polarity, and conformation of their local environments. There may be opportunities for improvements in these areas.

Chemical Stability

Some fluorescent dyes, particularly at long wavelengths, are too unstable for many practical applications. Improvements are needed.

Solubility and Size

The smaller and less perturbing the fluorescent probe, the better. Solubility in aqueous systems is a challenge in core fluorophore development because of the usual requirement for conjugated ring systems. Also, fluorophores in the visible and near IR have a minimum size to absorb light at these wavelengths. Development of useful up-converting fluorophores may be possible.

Fluorescent Probe Design: Adaptation of Fluorophore to Create a Probe

Targeting

Genetic methods for creating fluorescent labels and sensors within cells need continued pursuit. Few current probes can be considered as true site-specific labels. Other methods need to be developed to introduce organic and inorganic fluorescent probes into multicellular systems. Loading probes into bacterial cells is particularly problematic (Fig. 6.13).

Figures 6.14 and 6.15 compare penetration of probes in living vs fixed multicellular systems. Targeting of biosensors to quantify cytokines and growth factors in interstitial spaces is needed. Peptide-based delivery systems and other chemical tricks sometimes prove useful for targeting.

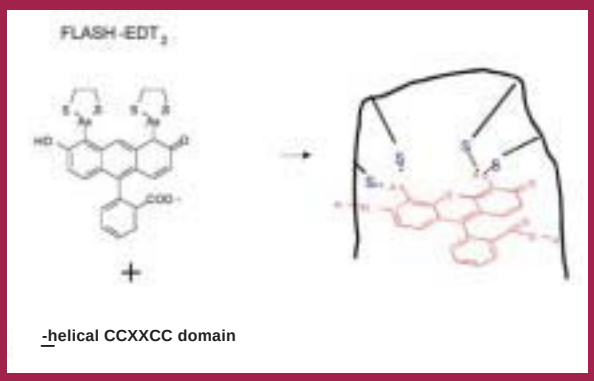


Fig. 6.13. The fluorescein arsenide FLASH-EDT2 specifically labeling proteins containing the tetracysteine sequence CCXXCC. This provides a way to introduce a site-specific fluorescent label in proteins. [Redrawn with permission from B. A. Griffin, Stephan R. Adams, and Roger Y. Tsein, "Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells," *Science* **281**(5374), 269–72. ©1998 American Association for the Advancement of Science]

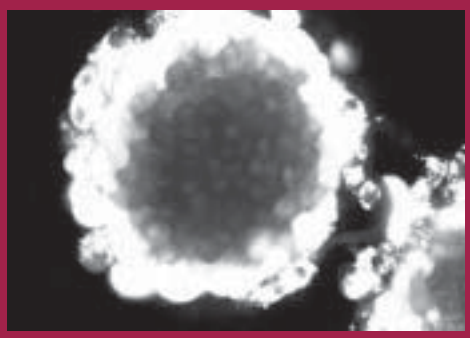


Fig. 6.14. A single optical section taken by two-photon excitation imaging through an intact pancreatic islet (~100-μm diameter) loaded with Fura-2/AM. Only the outer two cell layers incorporate the probe, while the cells in the center remain unlabeled. [Source: Dave Piston, Vanderbilt University]

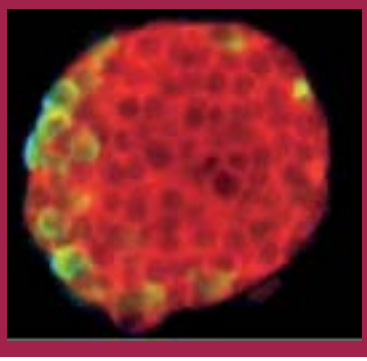


Fig. 6.15. A single optical section of a fixed and permeabalized pancreatic islet. This islet has been immunolabeled with antibodies to insulin (red), which labels ß cells, and glucagon (green), which labels _cells. Uniform penetration of these dyes is possible in fixed samples. [Source: Dave Piston, Vanderbilt University]

Reporter Functionality – Sensitivity to Environment

Addition of chemical groups that confer specificity to environmentally sensitive fluorophores is needed to create probes for new targets, whether they are ions, enzymes, or new chemical environments of cells.

Fluorophore Modifications

Modifications of fluorophores change solubility properties, reduce nonspecific binding, and affect targeting, penetration, and clearance of probes in cells and tissues or multicellular systems. These properties must be addressed in fine-tuning the properties of a probe instrument (Fig. 6.16).

- Considerations for fluorescent probes
- Bright, multicolor, solid-state lasers over spectrum
- IR-detector improvements as new near-IR probes are developed
- Picosecond and nanosecond excitation-detection engines as new probes appear in the 10- to 100-ns lifetime range
- Rapid, high-rejection multicolor optical filtering for multicolor fluorescence detection
- Long working distance—high NA objectives
- Environmental systems for cells and tissues
- Cheap, easy-to-use, fast, powerful software systems

Probes for Technologies Other Than Fluorescence

Plasmon-Resonant Particles

Plasmon-resonant particles (PRP) are metal nanocrystals, typically gold or silver, capable of supporting a plasmon resonance (Fig. 6.17). This feature causes them to scatter light with a narrow distribution of wavelengths. The color of the scattered light is dependent on particle size. PRPs are a million times brighter than fluorescent chromophores, and, because the optical signal results from scattering rather than an absorption and emission process, they are not susceptible to photobleaching, which makes them useful for monitoring dynamics over long time scales (hours). Their large size, however, restricts their use to monitoring dynamics in vitro.

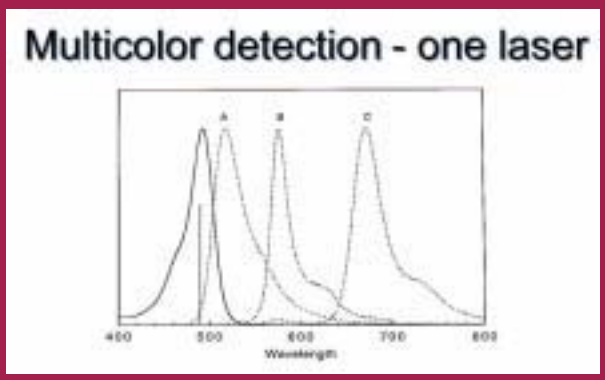


Fig. 6.16. Dyes that are excited at a single source but have emission at different, easily resolved wavelengths. An advantageous three-dye system is depicted here. [Source: Alan Waggoner, Carnegie Mellon University]

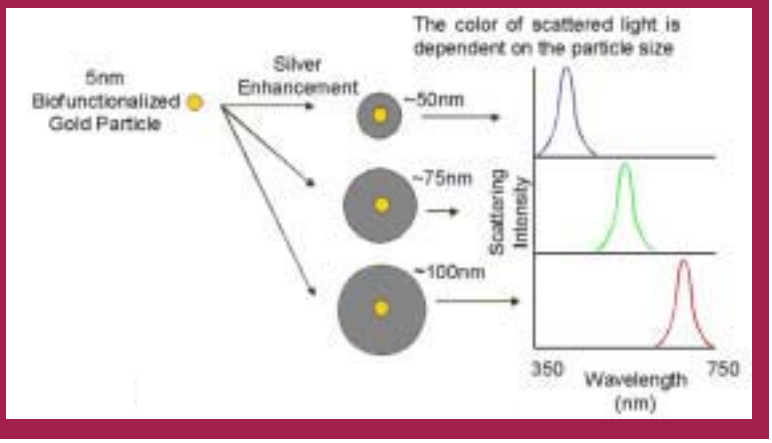


Fig. 6.17. Plasmon-resonant particles useful for monitoring dynamics over long time scales. [Source: Alan Waggoner, Carnegie Mellon University]

Rare-Earth Complexes

Rare-earth complexes have extremely long excited-state lifetimes (on the order of microseconds). This property makes them useful for time-gated imaging of biological samples. Using time-gated imaging eliminates the large background from autofluorescence and allows single molecules to be observed.

Quantum Dots

Quantum dots are small nanocrystals of semiconducting materials that display luminescent properties similar to organic chromophores but are more robust and provide longer-lived probes for monitoring dynamics over longer time scales. While quantum dots have sparked a great deal of interest and research, their implementation in the biological arena has been quite limited because of the difficulty in maintaining their robust optical properties when probes are biofunctionalized. New approaches are needed for biofunctionalizing quantum dots and novel materials for quantum dots.

Surface-Enhanced Raman Particles

SERS provide a robust signal and are easily biofunctionalized. However, their similarity in size to PRPs limits their applications to in vitro studies.

Dye-Loaded Silica Particles

Silica particles loaded with organic chromophores are robust probes. Because the chromophore is sequestered from the environment, the structures are much more resistant to photobleaching. They are somewhat smaller than the metal nanoparticles used for PRP or SERS probes; however, they are still on the order of 10 nm, which again limits their use in vivo. To avoid having the tags interfere with protein function, particles must be developed in the nanometer to subnanometer range. Other approaches, such as the trapping of fluorochromes in molecular cages or dendrimers, may enable production of much smaller tags exhibiting the same robustness as the silica particles.

Probes for Second-Harmonic Generation

The incorporation of small particles such as polystyrene beads into a system provides a method for multiple contrast mechanisms within the same specimen. For example, one protein could be labeled with a fluorescent dye suitable for use with multiphoton spectroscopy, and another protein could be labeled with a polystyrene bead. The multiphoton fluorescence would provide visualization of one protein while the second harmonic signal generated from the beads allows the second protein to be localized.

Electron-Dense Probes

X-ray and electron microscopy require probes in the size range of 0.5 to 10 nm and composed of heavy atoms that provide high scattering contrast. Particles near molecular scale can be used to identify specific sites within molecular complexes or provide fiducial markers for particle alignment (Fig. 6.18).

Currently, the best labels are small gold clusters containing either 11 or 75 gold atoms surrounded by an organic shell with a single specific binding site for any of a variety of functional groups. These enable the localization of structures in the 2-nm range. Higher specificity can be obtained using linkers such as antibody fragments to permit localization in the 4-nm range. Larger colloidal gold particles coupled to primary or secondary antibodies are useful for identifying specific structures in a larger context. Smaller clusters, such as tetrairidium, can be used in conjunction with image averaging to give resolution in the subnanometer range.

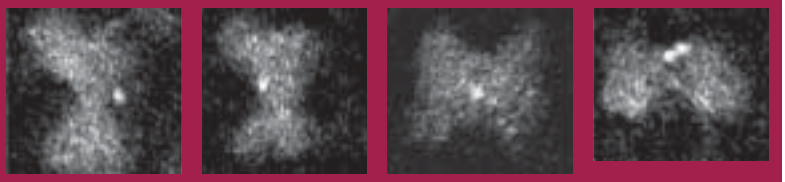


Fig. 6.18. STEM images of freeze-dried unstained phosphorylase kinase that has undergone exchange. [Source: K. W. Traxler et al., "Direct visualization of the calmodulin subunit of phosphorylase kinase via electron microscopy following subunit exchange," *Journal of Structural Biology* **135**(3), 231–38 (2001)]

One useful aspect of electron-dense probes is the inclusion of a second contrast mechanism such as fluorescence so that the same specimen can be studied by both optical and electron microscopy. This requires careful attention to quenching by the metal particle, which can be used to advantage in producing molecular beacons. Gold-cluster contrast can be enhanced for optical microscopy by catalyzing deposition of additional metal.

Future needs include the development of additional clusters of different sizes or shapes that can be distinguished from each other and used in studies requiring multiple labels. Clusters with different metal cores would also be distinguishable in multiple labeling experiments. Another issue that must be addressed in the development of these labels is the reduction or minimization of nonspecific binding.

Summary of Needs for Probe Development

- **Probes with more photostability and less toxicity, especially IR**
- **Probes for extreme environments that tolerate a wider range of pH and temperature**
- **IR probes for more color imaging and for imaging deeper into colonies, tissues, and complex multicellular systems (up to 2 mm)**
- **Narrower emission bands for organic dyes of multicolor analysis**
- **Better coatings for quantum dots**
- **Dyes with lifetimes in the 10- to 100-nsec range**
- **More genetic fluorescent probes**
- **Greater capability for site-specific fluorophore modification of proteins in cells**
- **Improved fluorescent probe-delivery systems for cells and penetration into colonies, biofilms, and tissues**
- **Creative probes for obtaining new kinds of cellular information**
- **Multisignature probes for imaging technologies other than fluorescence**
- **Ways to tether and orient probes on substrate**
- **Training support for development of probe-chemist manpower**

Appendices

Appendix A: Workshop Attendees 51

Appendix B: Workshop Agenda 55

Appendix C: Imaging Methodologies for Genomes to Life Research 59

Appendix D: National Laboratory Capabilities and Imaging Technologies 69

Appendix A: Workshop Attendees

Imaging Workshop Attendees

Workshop Co-Chairs

Steve Colson, Pacific Northwest National Laboratory
Damir Sudar, Lawrence Berkeley National Laboratory

Workshop Planners and Participants

*Steering Committee Member

Ames Laboratory

Ryszard Jankowiak

Argonne National Laboratory

*Ken Kemner
Lee Makowski
Jörg Maser
*Mike Miller

Brookhaven National Laboratory

*Carl Anderson (Executive Summary Writing Team Co-Leader)
Lisa Miller
*Joe Wall

Carnegie Mellon University

Alan Waggoner (Section 6 Writing Team Leader)

Cornell University

Watt Webb

East Carolina University

John Sutherland

Lawrence Berkeley National Laboratory

Rod Balhorn (Section 1 Writing Team Leader)
*Mark Biggin
Thomas Budinger
Kenneth Downing
Hoi-Ying Holman
Jian Jin
Carolyn Larabell
*Bahram Parvin (Section 5 Writing Team Leader)

Lawrence Livermore National Laboratory

Stavros Demos
Chris Hollars
Thomas Huser
*Steve Lane
Judy Quong
Chad Talley

Los Alamos National Laboratory

Robert Donahoe
Larry Hersman, (Section 3 Writing Team Leader)
*Jim Jett
Judith Mourant
*Cliff Unkefer

National Water Research Institute (Canada)

John R. Lawrence

Oak Ridge National Laboratory

*Mitch Doktycz (Section 4 Writing Team Leader)
Shaun Gleason

Pacific Northwest National Laboratory

*Jim Fredrickson
Julie Gephart
Gary Holtom
*Don R. Jones
Paul Majors
Doug Ray
(Executive Summary Writing Team Co-Leader)
Harold Trease
*H. Steven Wiley
Robert Wind

Sandia National Laboratories

Alan Burns
David Chandler
Jerilyn Timlin
Texas Tech University
Darryl Bornhop

Thomas Jefferson National Accelerator Facility

*Stan Majewski
*Gwyn Williams
(Section 2 Writing Team Leader)
*Andrew Weisenberger

University of Connecticut Health Center
Leslie Loew

University of Illinois
Paul Bohn

University of Iowa
Everett Greenberg

University of Southern California
Ken Nealson

University of Texas
Martin Poenie

Vanderbilt University
David Piston

DOE/BER/ASCR Advisors

ASCR
Gary Johnson

BER
Dean Cole
Marv Frazier
Mike Viola
Daniel Drell
John Houghton
Prem Srivastava
Peter Kirchner
Roland Hirsch
David Thomassen

Consultant
Mike Knotek

Technical Support

ORISE
Mikki Dawn
Deborah Garland

PNNL
Julie Gephart

Appendix B: Workshop Agenda

Agenda for Imaging Workshop

Monday, April 15 Arrivals

Tuesday, April 16

8:00 a.m. **Welcome and Charge to the Workshop Attendees**
Marvin Frazier/Michael Viola

Product: A draft workshop report will be prepared at the meeting that will speak to a wide audience and provide technical input to DOE/OBER.

8:30 a.m. **Review of the Schedule and Other Administrative Details**
Steve Colson/Damir Sudar

Executive Summary Writing Team Formation and Charge

Computational Infrastructure Team Formation and Charge

9:00 a.m. **Plenary Lectures** (projecting a broad vision for the future needs and value of imaging to GTL research)

Biology Perspective 1: Ken Nealson (http://wrigley.usc.edu)

Biology Perspective 2: Peter Greenberg

Technology Perspective: Watt W. Webb

Noon: Lunch (catered)

1 p.m. **Technical Session 1: Molecular Engines - Protein Complexes**

Overview Presentation: Biology Perspective

Overview Presentation: Technology Perspective

Brief Presentations: A few one-slide presentations on biology needs and Technical Approaches

Guided Discussion

Molecular Engines Writing Team Formation and Charge

3:30 p.m. **Writing Session 1**

Rooms available for each writing team. (Teams can start organizing information before related technical sessions.)

6:00 p.m. Evening Reception

Wednesday, April 17

8 a.m. **Technical Session 2: Intracellular and Cellular Structure, Function, and Processes**

Overview Presentation: Biology Perspective

Overview Presentation: Technology Perspective

Brief Presentations: A few one-slide presentations on biology needs and Technical Approaches

Guided Discussion

Writing Team Formation and Charge

10:00 a.m. **Technical Session 3: Multicellular: Monoclonal and Heterogeneous Multicellular Systems, Cell-Cell Signaling, and Model Systems**

Overview Presentation: Biology Perspective

Overview Presentation: Technology Perspective

Brief Presentations: A few one-slide presentations on biology needs and Technical Approaches

Guided Discussion

Writing Team Formation and Charge

Noon: Lunch (catered)

1:00 p.m. **Technical Session 4: Cells In Situ and In Vivo: Bacteria in the Natural Environment, Micro Environment, and In Vivo Systems**

Overview Presentation: Biology Perspective

Overview Presentation: Technology Perspective

Brief Presentations: A few one-slide presentations on Biology Needs and Technical Approaches

Guided Discussion

Writing Team Formation and Charge

3:00 p.m. **Writing Session 2: Parallel Report Writing Team Sessions**

4:30 p.m. **Evening Meal**

6:30 p.m. **Writing Session 3: Parallel Report Writing Team Sessions**

Rooms available for each writing team to start at their convenience.

Thursday, April 18

8:00 a.m. **Report and Guided Discussion of Each Report Element**

Executive Summary

Molecular Engines

Intracellular and Cellular

Multicellular

Cells In Vivo

Image Data Computational Infrastructure

Probe Development

11:00 a.m. **Development of Completion Plan for Report**

Capture of Cross-Cutting Technologies, Methods, and Needs

Identification of Duplications

Refinement of Writing Plan

Assignments to Obtain Information and Figures not Available at Workshop

Noon: Lunch (not provided by meeting)

Those not participating in report writing may depart at this time.

1:00 p.m. **Writing Session 4: Parallel Report Writing Team Sessions**

(Bag lunches provided if the team wants to work during lunch.)

4:30 p.m. **Evening Meal (not provided by meeting)**

6:30 p.m. **Wrap-Up Writing and Capture of Follow-Up Actions**

Friday, April 19

Final Departures

Appendix C: Imaging Methodologies for Genomes to Life Research

Appendix C: Imaging Methodologies for Genomes to Life Research

Protein Identification and Characterization

A complete understanding of molecular machines will require information covering a wide range of dimensions and time scales. Such structural data as molecular dimensions, subunit structure, and overall shape typically are obtained with high-resolution techniques on a scale ranging from angstroms to a few nanometers. Information on dynamics and function of molecular machines, however, is usually obtained on larger spatial and shorter time scales that interrogate a molecular machine within its immediate environment. A variety of techniques can be used to access the structural and dynamic information of protein complexes. However, there is much room for improvement in several areas, including the possibility for exploitation of new phenomena.

Short-Wavelength Techniques

X-ray crystallography reveals atomic-resolution structures of subunits and small machines such as ribosomes.

X-ray microscopy is promising for the future but is not yet at a resolution to give internal details of complexes. The recent solution of the whole ribosome structure demonstrates that X-ray techniques can be extended to larger structures, but the process requires considerable effort.

Electron microscopy (EM) is well established for viewing complexes at 0.5- to 2.0-nm resolution. Particles typically are viewed by tomography either in 2D arrays or as isolated objects. If the sample is freeze-dried, quantitative measurement of electron scattering can give direct mass distribution useful for determining subunit stoichiometry or faithfulness of reassembly.

High-resolution techniques such as **transmission electron microscopy (TEM)** must be carried out on thin specimens placed in a vacuum chamber and held stationary. **Flash-freezing** is a well-developed technology providing high-resolution 3D structures of isolated complexes and 2D arrays.

Scanning transmission electron microscopy (STEM) gives information complementary to TEM techniques by furnishing a direct link to the biochemistry. Use of metal clusters such as undecagold as labels for specific constituents makes STEM a valuable method for localizing gene products within complexes.

The resolution limit of **electron tomography** is not the microscope's resolution but the number of obtainable distinct views of the cell. Because every cell is different, an entire tilt series must be carried out on a single cell. Methods for lowering radiation damage are needed so larger numbers of images can be obtained on a single specimen. This may be possible using higher-voltage electrons (>500 keV) or by imaging specimens maintained at liquid helium temperatures.

Intact molecular machines composed of multiple proteins often are too large or too flexible to make a crystallographic approach feasible. This challenge is best suited to EM. Three-dimensional reconstructions can be calculated from images of a molecular machine taken at many different angles.

Optical Microscopy

Laser-based optical probe techniques show great promise for real-time monitoring of the motions of molecular machines. Evolving techniques are under the diffraction limits of focused light and will allow monitoring of individual molecules in vivo and in vitro. **Multiphoton and multicolor techniques** hold promise for high selectivity and simultaneous detection of multiple species. Advances in detector technology are needed to permit the real-time imaging of multiple species, making possible the simultaneous monitoring of chemical composition and molecular motion.

Multiplexed real-time measurements will uncover details of the mechanisms for building and operating molecular motors.

The following are specific needs:

- **Fluorescence-based microscopes** that provide high-speed, multicolor detection for study of the time dependence of the co-localization (complex formation) dynamics of complexes composed of multiple proteins.

- **Optical microscopes** that use nonlinear optical methods [i.e., **two-photon fluorescence, coherent antistokes Raman scattering (CARS), second-harmonic generation (SHG), third-harmonic generation (THG),** and **sum-frequency generation (SFG)**] singly or in combination to obtain data about the structure and dynamics of molecular machines with minimal requirement for addition of probes.

- Single-molecule, distance-sensitive optical methods [i.e., **fluorescence resonant energy transfer (FRET)** and **lifetime time-polarization** evolution methods) to obtain real-time data on structural fluctuations within molecular machines under controlled conditions.

- **Surface-enhanced Raman scattering (called SERS)** to provide a fingerprint spectrum of the structure and to show correlated spectral fluctuations with molecular structure. The use of surface-enhanced approaches and predetermined spectral features correlated with the movement of interest should enable millisecond time scales over extended periods.

Scanning Probe Microscopy

Scanning probe microscopy is becoming a true workhorse for determining biological structure and function. It can be performed in situ (on live organisms, if necessary) without staining. This technique potentially provides for nanometer-scale topographic resolution. **Atomic force microscopy (AFM)** can be combined with and enhance **simultaneous fluorescence imaging**. It also can be used to image individual proteins, protein complexes, and molecular arrays (e.g., membrane proteins located in membranes). The phase-sensitive tapping mode furnishes information about chemical changes and "softness" (compliance). Functionalized tips may be used to generate force vs displacement curves, providing the ability to obtain single-molecule bond-strength information. This capability lends itself to the characterization of protein-protein or protein-ligand interactions and tension-induced protein unfolding.

Diffraction Methods

Information about size and shape of purified structures randomly oriented in solution can be obtained by X-ray or neutron scattering. **Neutron diffraction** is particularly useful for defining the volume of water excluding space through the use of scattering in varying ratios of H_2O and D_2O. Variation of the H_2O-D_2O ratio can be used to contrast deuterated proteins, making possible the study of ligand binding, including conformational changes induced by the interaction. Wide-angle solution scattering also has the potential to provide scattering fingerprints for determination of the structural class of a protein in solution. For instance, this technique can readily distinguish between an immunoglobulin fold and a beta-barrel.

Structure Determination of Single-Protein Molecules or Complexes

The following are needs for determining the structure of single-protein molecules and complexes.

In the future we anticipate a need for laboratory-scale X-ray diffraction and electron scattering sources that will allow individual scientists to have dedicated machines for protein and protein-machine structure determination.

Rapid methods are needed for structure determination of single-protein molecules or complexes. Current methods, which are essential to identifying protein function, take years to provide a finished structure. **Nuclear magnetic resonance (NMR)** structure determination currently is limited to proteins of 40 to 60 kD, and X-ray diffraction methods require that the proteins be crystallized. Crystallization is the slowest step in the process, often taking many years of concerted effort. In addition, many proteins never crystallize; among these are the membrane proteins. Most molecular signaling involves traversing the membrane, and over 30% of all proteins are known to be tightly associated with membranes.

New biochemical and biophysical approaches will be required to identify the molecular structures of these proteins.

New developments are leading to the production of a "next-generation" beamline such as the **Linac Coherent Light Source** at Stanford University. This technology offers the promise of obtaining high-resolution (2 Å or less) diffraction data from single-protein molecules, protein complexes of any size, viruses, and even cells without crystallization, using a coherent X-ray source that is ten times brighter than any today. To aid development of this technology, which may become available as early as 2007, supporting efforts will be needed specifically for single-molecule diffraction techniques using this beamline. These techniques involve sample preparation and introduction, design of rapid-readout detection arrays, and development of computer algorithms for combining and interpreting diffraction data from tens of thousands of molecules.

Imaging Biological Functions at the Subcellular, Cellular, and Multicellular Scales

Short-Wavelength Techniques

X-Ray Microscopy

X-ray microscopy today uses electromagnetic radiation with photon energies in the 250-eV to 30-keV range with corresponding wavelengths from 45 nm to 0.4 Å. Because of this short wavelength, suboptical spatial resolution can be obtained. Several different contrast mechanisms can be used to form an image, thereby making X-ray microscopes useful for probing different properties of matter. High-resolution X-ray microscopes use Fresnel zone plates as focusing and imaging optics, therefore becoming the "enabling technology" in achieving suboptical spatial resolution.

X-ray microscopes operate second- and third-generation synchrotron sources because of the need for high-photon density in the specimen plane, high brilliance for all probe-based systems, and energy tunability for applications involving spectroscopy. Development of laboratory X-ray microscopes is under way, but limited flux density from laboratory sources will most likely limit the spatial resolution in these systems to the level of 100 nm or slightly below.

Two classes of X-ray microscopes are used in both the soft and hard X-ray range. These are **transmission X-ray microscopes (TXM)**, which record a full image of the specimen in absorption and phase contrast; and **scanning X-ray microprobes,** which focus the coherent part of the monochromatized X-ray beam into a diffraction-limited spot and use secondary signals such as X-ray fluorescence or photoelectrons for contrast formation.

Specimen preparation for all X-ray microscopies is relatively straightforward because no vacuum is required and unsectioned specimens up to hundreds of micrometers thick can be penetrated. Xrays are ionizing radiation, however, and radiation damage is of significant concern. **Cryomethods**, therefore, have to be used at higher spatial resolution to achieve structural preservation. Cryomethods are a standard—if challenging—technique in the soft X-ray range and will be required in hard X-ray microscopes as spatial resolution improves.

Soft X-Ray Microscopy. Soft X-ray microscopy typically covers the photon energy range from 250 eV to 1 keV, with particular emphasis on the "water-window" spectral range between the electron banding energy of oxygen (Eb = 543 eV) and carbon (Eb = 284 eV). Here, water layers of 10 mm can be penetrated, and good contrast of the object's carbon distribution is obtained. **Absorption spectroscopy** at the K-absorption edges of C, N, and O has been performed; some applications at higher energies involve L-edge spectroscopy of transition metals such as Fe. X-ray absorption maps and spectra can be quantified using the published elemental absorption length.

TXM and **scanning transmission X-ray microscopes (STXM)** are used in the soft X-ray region. TXM can operate at an incoherent bending magnet source of synchrotrons, while STXM requires the high brilliance of an insertion device to provide spatially coherent X rays of sufficient flux. TXMs are operated at the Advanced Light

Source (ALS) at Lawrence Berkeley National Laboratory (LBNL), BESSY in Germany, and Aarhus in Denmark. Because of their relatively simple setup, they funish the highest spatial resolution of any X-ray microscope to date, currently around 20 nm. Immunogold labeling is a way to specifically tag certain proteins to investigate cellular functions. Tomography has provided 3D reconstructions of whole, hydrated cells. STXMs operated at the National Synchrotron Light Source at Brookhaven National Laboratory and at ALS are unique tools for spectroscopy such as C-XANES. They also are capable of using secondary signals such as photoemission.

Hard X-Ray Microscopy. Although soft X-ray microscopes have reached a significant level of maturity, high-resolution **hard X-ray microscopes** using zone-plate optics have only recently begun operation. Their application to biological and environmental sciences is yet more novel, and significant R&D is under way to develop their potential. Hard X-ray microscopy covers the energy range from ~1 to 30 keV; in this energy range, **X-ray fluorescence** becomes a very efficient contrast mechanism, and the sensitivity that can be achieved by fluorescence mapping becomes significantly larger than what absorption contrast imaging can provide. By using K and L states, all elements in the periodic system above Na can be directly excited, including heavy elements like Hg, Pb, and U. X-ray fluorescence is a fully quantitative method.

Because X-ray fluorescence requires excitation by a focused beam, **scanning X-ray microprobes (SXMP)** are used primarily in the hard X-ray region. As in the soft X-ray range, these systems require a coherent hard X-ray source, which in turn needs the brilliance of an insertion device at a third-generation synchrotron and high electron energies of 6 to 8 GeV in the storage ring. Only three such facilities exist in the world: the Advanced Photon Source at ANL, European Synchrotron Radiation Facility (ESRF) in France, and Spring8 in Japan. A TXM for the energy range around 8 keV is set up at ESRF. As transmission contrast becomes smaller with increasing X-ray energy, a TXM for the hard X-ray range can no longer be operated by a bending magnet, instead requiring the full photon flux of an insertion device at a high-energy synchrotron.

Very high trace metal sensitivity, on the order of 10 to 17g for transition metals, can be routinely achieved in an SXMP. Therefore, hard X-ray mapping provides a unique and powerful way to map trace-element distributions in cells and microbes. Energy-dispersive detectors allow collection of X-ray spectra of all elements below the excitation energy, so whole sequences of elements (e.g., P, S, Cl, Fe, Ni, Co, Cu, and Zn) are mapped simultaneously. X-ray spectroscopy can in principle be performed for any element heavier than Na. Particularly for Ca and heavier elements, spectroscopy at trace sensitivity can be carried out. This allows, for example, collection of XANES spectra of Cr in microbial specimens or distribution mapping of Cr subspecies in the specimen. The spatial resolution of hard X-ray microprobes is limited by the technology available for the manufacture of Fresnel zone plates. Current technologies limit the spatial resolution achievable with reasonable efficiency to ~100 nm. A significant investment is required to improve the technology to push the spatial resolution to 30 nm or better.

Applications. X-ray fluorescence mapping can be applied to trace metals such as Fe, Cu, Zn, As, Se, Cd, Hg, Pb, or U in organisms like microbes or eukaryotic cells, while simultaneously collecting distribution from P, S, Cl, K, and Ca.

Challenges. The critical element of any X-ray microscope is the zone plate used for imaging or focusing. The outermost zone width and the aspect ratio define the spatial resolution and diffraction efficiency that can be achieved. A major investment in zone-plate technology is therefore needed to improve resolution. Achieving an outermost zone of 10 nm or less, while maintaining a thickness on the range of 200 nm for soft X-rays and >>1 μm for hard X-rays, would enable a major breakthrough in all applications. Another challenge is X-ray tomography. In particular in the hard X-ray region, 3D imaging is not feasible at dwell times of a second per pixel. A major improvement could be achieved by maximizing the solid angle of detection using novel detectors and improving the brilliance of hard X-ray storage rings.

Electron Tomography

Electron tomography has the potential to reach the highest spatial resolution in imaging 3D structures of intact cells and allow mapping of the locations of subcellular components. Specimens can be prepared in the frozen-hydrated state, preserving the native conformation of the cell and its contents. Intermediate-voltage electron microscopes, operating at 300 to 400 kV, allow imaging of samples several hundred nanometers thick, so we can examine elongated bacteria such as *Magnetospirillum* and *Caulobacter* intact. Theoretical analyses suggest that we should be able to reach a resolution of 2 to 5 nm with such specimens. At this resolution, identification of subcellular features such as cytoskeletal filaments and nucleic acid condensates should be possible. Also, large macromolecular complexes such as ribosomes, proteasomes, and polymerases could be identified by matching the observed densities with structures of corresponding molecules that have been solved by X-ray crystallography or other techniques.

Although the ability to identify the larger protein complexes directly will provide a wealth of new information, selective labeling of components with electron-dense agents will enable localization of more components. Labeling strategies could include conventional antibody-targeted metal atom clusters. The use of quantum dots, for example, would allow correlative light microscopy, but the particle size for EM would be smaller than for current light microscopy. A problem yet to be addressed is how to get the particles into the bacteria. Evidence suggests that complexing them with a naturally imported peptide such as transferrin may work. Other labeling approaches under development include modification of a particular gene to include a photoconvertible eosin or tetracysteine peptide that can be complexed with a heavy-metal atom cluster.

Instrumentation developments are needed before the technique becomes routinely applicable. Microscopes operating at even higher voltage would be helpful in obtaining higher resolution with thicker specimens. Such developments probably will be driven mainly by the materials science community. The performance of charged coupled display (CCD) cameras used to record images on the microscopes degrades at higher voltages, so technology must be developed to compensate for this loss in performance. The microscope's specimen stages need to be improved so they operate reliably with liquid helium cooling, eucentric rotation around the tilt axis, and rotation of the grid within the holder. Automation of the microscope operation for recording the series of images is still under development.

Specimen preparation for frozen-hydrated, whole-mount samples is fairly well developed, but improvement of the methodology for sectioning frozen samples would open up the technique to applications with multicellular samples. Finally, managing the image data sets and searching the reconstructed volume densities for patterns corresponding to known structures will be significant computational efforts.

Optical Microscopy Methods

Many imaging methods using visible or **near-infrared (called NIR)** light are a part of optical microscopy. We loosely define wavelength boundaries as the optical transmission windows of cells and common glasses (~300 nm) through the range of efficient detectors (~1700 nm). These methods vary in difficulty, from inexpensive and routine tools to major interdisciplinary efforts requiring investments at the national-laboratory scale. The accessible-distance scale ranges from the "optical limit" of about half the wavelength to the macro scale, or, in other words, from the bacterial and cellular organelle scale to sample sizes limited only by optical transparency.

Image production results from a wide variety of physical and chemical properties, providing rich information content. In favorable circumstances, the detection limit may be a single fluorescent molecule. To define a role for GTL in meeting the needs of the biological community, we may classify optical-imaging methods according to difficulty and by application. This classification also is intended to show where multidisciplinary teams of physicists, chemists, and biologists should be encouraged to work together to produce the capabilities required for the next generation of methods and instruments.

Classical One-Photon Methods

All "classical" microscopy methods use continuous or low-power light sources. The most routine methods use white-light illumination in either bright- or dark-field configurations, with the more-usual transmitted light or the less-common front-side or epllumination. Contrast may be improved by a number of physical methods including **phase contrast** or **differential image contrast (called DIC)**. This is the most straightforward of microscopy techniques and is usually a part of the everyday laboratory environment. Detector technology advanced recently by introduction of **direct electronic detection** (CCDs) to replace film, but further advances are possible in sensitivity as well as in spatial and temporal resolution.

Specific and very useful information is provided by fluorescence methods, using intrinsic or added dyes. A variety of imaging systems are available, using wide-field or confocal methods. The dyes may have either general or highly specific affinities for biological constituents of interest either within cells or in the extracellular matrix. Fluorescent dyes can be tailored to provide environmental data such as membrane potential, pH, or polarizability of the local environment. Intensity information can be useful, particularly with ratio-detection methods, but content is improved considerably by making images based on fluorescence lifetime and spectral shifts.

Although many probe molecules are available, creating improved labels is an important effort. Better photostability, improved ability to reach targets within or on live cells, lessened photo-induced toxicity, and the ability to quantify chemical messengers or environmental characteristics around the cells or within the tissue require continuing research. In general, fluorescence techniques are routinely available and do not require a national commitment.

Obtaining chemical data (location of abundant species or important minor chemical compounds) is difficult with the above techniques. A direct means for obtaining molecular information is by vibrational resonances, either by IR or by Raman. IR methods are described in the "Infrared Imaging Methods Based on Intramolecular Vibrational Modes" section below.

Raman measurements can be performed with the same equipment as fluorescence. While sensitivity is a problem, full 3D imaging can be obtained with confocal methods. In special cases, the target molecule may be detected at improved sensitivity by resonance Raman techniques. Although improved detection technologies will be useful, those for Raman and fluorescence measurements are approaching their theoretical limits, leaving better probes as the leading development opportunity.

A special imaging category is **scanning optical tomography**, which requires a pulsed light source to provide depth profiling information. While the laser must produce ultrashort light pulses, high power is not required, and obtaining 3D information useful on a tissue scale but not on a single-cell dimension is possible.

Multiphoton Microscopy

The most promising avenue for obtaining new kinds of information—and the area where funding at the national-laboratory level can provide the greatest impact—is in multiphoton methods. A wide variety of characteristics may be studied, with the general requirement of an ultrafast laser technology. Very often, an experiment may be designed to provide numerous data channels for the production of multidimensional images, with a number of useful ways to understand the complex biological sample.

Two-photon fluorescence (TPF) microscopy is the most common application. A very intense, long-wavelength laser pulse can excite fluorescent molecules at the focal region, even in the transparency window. Typically a titanium-sapphire laser producing a train of pulses in the range of 700 to 1000 nm is used to excite a wide range of species, including NADH or other intrinsic chromophores and a variety of blue-, green-, or red-emitting fluorescent probes. Both organic and inorganic species are routine. The technique can involve temporal and spectral resolution, polarization information, and (in some cases) Raman and other scattering signals. Optical penetration is much improved, and the method is intrinsically suited to 3D imaging because the signal is produced only at the focus of the laser beam. Signal detection is simplified since a confocal aperture is not required, and the

signal may be collected by epidetection. The primary development effort required for wider access of TPF methods is improved (i.e., smaller, cheaper, and more reliable) laser technology and improved optical probes.

Direct access to fluorescent amino acids within proteins is possible by **THG** (chromophores at wavelengths <300 nm).

A variety of physical characteristics are accessible by related methods. **SHG** produces a monochromatic beam at half the wavelength of the incident beam. It is a coherent process and consequently sensitive to physical properties, particularly the organization of interfaces within the distance scale of the wavelength of light. Only asymmetric structures are visible as a consequence of optical-selection rules (most cell membranes produce a negligible signal). Some ordered macromolecules such as collagen produce a very large SHG signature. Useful physical information is available by comparing the transmitted, or forward, SHG with the backscattered signal, which is insensitive to phase and consequently produces very thin structures.

Another physical property may be imaged by THG, which requires lasers with wavelengths longer than 1200 nm to provide a signal transmitted in the optical window. Such ultrafast laser sources are not routinely available.

The nonlinear microscopy analog of Raman imaging is **CARS**. Various CARS applications (i.e., epi, trans, pol, time-delayed, and BOXCars) provide particular imaging advantages, and new methods for taking advantage of them are under development.

Multifunctional Microscopy. A unique advantage of multiphoton microscopy is the possibility of combining various imaging technologies. All the signals are generated simultaneously at the focal point of the laser, often available in both the transmitted and epi directions. Fortunately, the signals may be separated and detected in separate channels. Particularly useful combinations are **CARS** and **TPF**, or **TPF/SHG/THG,** although strategies for collecting **CARS/TPF/SHG,** in both epi and transmitted directions, are available. A rich, perfectly registered, multidimensional image data set that contains a variety of chemical, physical, and structural data may be produced.

Challenges and Needed Technologies for Multiphoton Imaging. Clearly, the laser systems for **CARS** techniques are too complex for routine operation. Technologies for improving the lasers and detection systems, needed for widespread adoption of new multiphoton imaging methods, are possible but not commercially available. Conditions for optimizing laser sources to produce the best images with tolerable signal levels will require continuing effort. Providing advanced imaging tools for routine biological research and as related imaging tools at other national-laboratory imaging facilities will be helpful for continuing progress in a wide variety of biological problems.

Infrared Imaging Methods Based on Intramolecular Vibrational Modes

Far-field infrared vibrational-imaging methods will provide an important complement to fluorescence- or absorption-based optical approaches for the study of living individual cells and colonies. The use of synchrotron radiation and lasers provides diffraction-limited spatial resolution, which could be improved further by the use of image-enhancement algorithms. Vibrational signatures are specific to molecular structure and, as such, can produce direct chemical information without the necessary use of dyes or other labels. Because of the nondestructive nature of single-photon mid-IR probes, vibrational imaging can be exploited to observe dynamic changes in living cells.

Far-field infrared vibrational imaging can be developed to examine community signaling or colony and substrate chemistries at the scale of a few microns and above. Collectively, vibrational-imaging techniques should be developed to analyze microbial response at the level of individual cells or communities to perturbations such as changes in oxygen or other substrate levels (e.g., NH_3 and metals), pH or temperature changes, and other factors that will affect critical elements of the interaction of microbes with their environment. Because such hyperspectral imaging methods can potentially yield large data sets including spectral, spatial, and temporal

dimensions, imaging-analysis methods that can properly exploit and manage such data sets should also be developed.

In addition to their application to intrinsic molecular signatures, vibrational-imaging methods could be developed to take advantage of the introduction of molecules that have unique vibrational signatures.

Thus, "tagging" pathways or structures in microbes by the use of vibrational labels such as stable isotopes or ligands (e.g., CN and CO) could allow selective detection of specific molecular components within and among microbes.

Near-Field and Scanning Probe Imaging Methods

Near-field microscopies typically are based on the spatial confinement of photons by probe-microscopy structures. They offer attractive pathways to achieving resolution beyond the diffraction limit and could potentially combine topographical imaging with chemical sensitivity. Due to their dependence on probe-microscopy techniques, near-field probes usually are limited to cell exteriors, but this restriction could be overcome with newly developed photonic force microscopes that could "grab" artificially introduced labels within a cell or cell constituent and move them to probe the cell interior.

Chemically functionalized near-field tips also could be injected through the cell membrane and positioned within the cell to measure the uptake of chemicals through the cell membrane or their diffusion within the cell.

The development of near-field imaging capabilities using narrow-band laser light to achieve 8/10 using apertures is highly desirable. Later, 8/100 could be developed using apertureless systems with scanning probes within living cells and membranes.

Magnetic Resonance Methods

Magnetic resonance imaging (MRI) and (localized) **magnetic resonance spectroscopy (MRS)** have matured as powerful methods for studying biological objects. These methods are noninvasive and nondestructive, so live objects can be examined and an object can be studied repeatedly to follow possible changes. Moreover, in principle, the object need not be disturbed because no contrast agents are required.

With **MRI**, usually the water-density distribution is imaged, and this method also can be used to study flow and diffusion of water and other fluids in the object. With **MRS** the presence, concentrations, and spatial distributions of small–molecular-weight chemical compounds such as cell metabolites can be examined; this can assist, for example, in gaining an improved understanding of metabolic activities in the object.

Specialized instrumentation and technologies have been developed to make **MRI** and **MRS** applicable in both large- and small-scale objects, and the techniques have been applied to study objects as small as a single eukaryotic cell and as large as human patients. Also, **MRI** instrumentation has been developed that can be used in the field to study, for example, the presence of oil and the propagation of groundwater. Moreover, **MRI** has been used to study the growth of biofilms and their distribution in porous and nonporous bioreactors, and **MRS** of protons and other magnetic nuclei has been used to investigate metabolic pathways and their kinetics in microbial systems. Therefore, **MRI and MRS** are potential key techniques for examining prokaryotic cell systems of different sizes, in both the laboratory and the field.

Challenges. The main challenge of **MRI and MRS** is to improve the techniques' notoriously low sensitivity-per-unit time, which severely limits the spatial resolution of the images. The best (isotropic) resolution now obtainable in a water image of a small (1 mm or less) cell system is about 10 μm. This resolution is reduced to ~100 μm when the metabolites, which have a much lower concentration than water, are mapped. In larger samples the resolution is further reduced roughly proportionally to the sample size. This means that only relatively large microbial communities can be investigated and that the NMR sensitivity needs to be multiplied at least three times to increase the image resolution to ~1 μm, the minimum resolution needed to image a single microbe.

Possible ways to increase MR sensitivity are to (1) use ultrahigh magnetic fields and construct ultrasmall MR microcoils, (2) cool the MR coil and receiver system, (3) enhance the MR signal via special techniques such as dynamic nuclear polarization, (4) enhance the MR image resolution by simultaneously imaging a system with another imaging technique having a larger spatial resolution, and (5) use a combination of all the above. Moreover, ultralarge magnetic gradients could be used to avoid diffusion losses in the high-resolution images. Enhancing the sensitivity also means enhancing the temporal resolution.

Another important challenge is to improve MR information for metabolic studies. Examples of such improvements are the development of techniques to enhance spectral resolution, separate intra- and intercellular metabolites involved in the metabolic processes, and use spin-labeled chemicals to follow the spatial and temporal evolution of metabolic changes with multinuclear MRS such as 13C and 31P MRS.

Appendix D: National Laboratory Capabilities and Imaging Technologies

Appendix D: National Laboratory Capabilities and Imaging Technologies

This section is a summary of imaging capabilities and technologies available at DOE national laboratories and is not intended to be a comprehensive list. Contact names and Web sites are provided for more information on each laboratory's capabilities.

Argonne National Laboratory (ANL)

Contacts: Lee Makowski (lmakowski@anl.gov) and Jörg Maser (maser@aps.anl.gov)
Web Site: www.bio.anl.gov/resources/

X-Ray Absorption Spectroscopy

This technique provides up to 0.01-Å resolution related to the radial distances from an absorbing atom to its neighbors. The local chemical probe provides information from ~0.01 to 5.0 Å in length. Additionally, it can provide information on chemical speciation and valence state for the absorbing atom. Depending on the absorbing atom's concentration in the sample and the design of the experiment, information can be obtained at the subsecond (sometimes at the microsecond) time frame.

X-Ray Fluorescence Imaging

This technique provides information related to the spatial distribution (150-nm resolution) of elements (atomic number > ~15) in a sample; imaging is not necessarily molecular unless a specific element is associated with that molecule. Time resolution depends on the ability to flash-freeze the sample each time a 1-h measurement is made.

Microfocus Secondary Neutral Mass Spectrometer

ANL also has been developing a microfocus secondary neutral mass spectrometer called SPIRIT that is capable of desorbing molecules using either a 0.25-μm laser pulse or a 50-nm–diameter ion beam. SPIRIT photoionizes the desorbed molecules using a VUV laser source for single-photon ionization and then mass-analyzes the photoions in a 4-m time-of-flight (TOF) mass spectrometer. Molecules can be imaged with an extremely high useful yield.

Brookhaven National Laboratory (BNL)

Electron Microscopy (EM)

Contact: Joseph Wall (wall@bnl.gov), Biology Department

Scanning Transmission Electron Microscopy (STEM)

Web Site: http://bnlstb.bio.bnl.gov/biodocs/stem/stem.html

STEMs, unique electron microscopes built at BNL, provide quantitative microscopy to biologists. More than just a picture, the digital STEM image provides data used in mass analysis performed by PCMASS, a computer program written by a STEM facility scientist. Individual molecules or assemblies of molecules of up to several billion daltons can be "weighed." One of the few electron microscopes in the world that can image single heavy atoms, STEM is suitable for visualizing tiny gold clusters attached as tags for biological macromolecules.

An NIH Biotechnology Resource since 1977, BNL-STEM1 conducts more than 50 active projects a year with scientists worldwide, using mass mapping and heavy-atom labeling for studies of native and reassembled complexes in the molecular-weight range of 0.1 to 1000 MDa. STEM1 is a custom-built instrument operating at 40 keV and –160EC with 0.25-nm spot size and efficient dark-field detectors capable of visualizing single heavy atoms and unstained DNA.

BNL-STEM3 is similar to STEM1 but operates at 70 keV with addition of an electron energy-loss spectrometer for elemental analysis. Operational since 2000, it is not a user facility.

Cryoelectron Microscopy (Cryo-EM)

BNL has a JEOL 2010F cryo-EM equipped with a 200-KeV field emission gun, a high-tilt stage, cryostage, 2-K by 2-K charged coupled display (CCD) camera, and STEM capability. Its planned use is for studies of membrane proteins in 2D arrays and for single-particle analysis by tomography; it also should be useful for STEM studies of frozen-hydrated specimens.

Atomic Force Microscope (AFM) and Scanning Tunneling Microscope (STM) in Air and Liquid

Contact: Stanislaus Wong (sswong@bnl.gov), Materials Sciences Department

BNL's state-of-the-art Nanoscope IIIa Multimode AFM allows the measurement of topographic data, viscoelastic properties, roughness, and morphological inhomogeneities as well as chemically sensitive phase information on the surfaces of samples (<5 nm routinely doable), often in simultaneous data-collection modes. Force measurements have a sensitivity of about 20 pN, and dynamic processes such as crystal growth can be measured.

Performing STM measurements is also possible with this instrument, which is capable of obtaining atomic resolution. Several probes are currently available, including conventional silicon nitride and silicon tips. A nanomanipulation setup is available for fabricating carbon nanotube tips important for high-resolution imaging.

Advanced Electron Microscopy Facility

Contact: Yimei Zhu (zhu@bnl.gov), Materials Sciences Department

BNL's Advanced Electron Beam Facility has several instruments, including the newly installed 300-kV JEOL 3000F transmission electron microscope (TEM), currently the best of its kind in the United States. It has a very bright field-emission source with a probe size down to <0.4 nm, a high-resolution pole-piece (point-to-point resolution <0.16 nm and lattice resolution <0.1 nm) with a permissible sample tilt of ±30E. It is equipped with an integrated Gatan energy filter, an electron energy-loss spectrometer, and an energy-dispersive X-ray spectrometer for high-quality imaging, diffraction, and spectroscopy.

The instrument also has analog- and digital-scanning units for spectroscopic imaging and chemical mapping for both light and heavy elements at a spatial resolution better than 1 nm and an annular dark field detector for Z-contrast imaging. An attached electron holography unit is designated to retrieve phase-sensitive magnetic and electrostatic information directly from the samples, while the heating and cooling stages (temperature range 15 to 1300°K) and versatile data-acquisition assemblage (Fuji Imaging Plate System and Gatan CCD and video cameras) allow quantitative, dynamic in situ experiments with a temporal resolution of 1/30 s.

The instrument is dedicated to materials science research for revealing crystal, electronic, and magnetic structures at nanoscale in solid state. It also is suitable for medical and biological research. We have demonstrated the instrument's ability to image boron distribution at nanometer resolution using the boron K-edge of biological samples, including tobacco mosaic virus.

Magnetic Resonance Imaging (MRI) with Contrast Agents

Contact: Charles Springer (cspringer@bnl.gov), Chemistry Department

In the context of MRI, the term "molecular" imaging means the use of paramagnetic contrast reagents (CR) targeted to specific sites on enzymes, receptors, and polynucleotides. Unlike the case for radiotracers, in MRI the CR is detected indirectly via its effect on the $^{1}H_2O$ nuclear magnetic resonance (NMR) signal measured. This generally requires some degree of molecular access of water to the CR. Among other features, it allows the design of CRs that are "activatable" only upon interaction with the target. Thus, they are somewhat analogous to the

"molecular beacon" fluorescent probes for optical imaging but can in principle be detected from deep within any living organism's tissue.

The BNL Chemistry MRI group has an international reputation for elucidating the physicochemical mechanisms operative for MR CRs in vivo. For example, BNL recently discovered that the threshold concentration for detection of CR is decreased as the magnetic field increases. This is very important in the design of CRs for crucial targets present only at very low concentrations in vivo. A BNL 4-T (Tesla) instrument for human and animal subjects soon will have a 7-T instrument for rodents, allowing the imaging of human and animal subjects with spatial resolutions easily better than 1 mm and 300 μm, respectively.

Infrared (IR) Microspectroscopy

Contact: Lisa Miller (lmiller@bnl.gov), National Synchrotron Light Source (NSLS)
Web Site: http://nslsweb.nsls.bnl.gov/infrared/u10b/

Synchrotron IR light is an ideal source for spectrospectroscopy because of the combination of its high brightness (i.e., flux density) and broadband nature. Through a 10-mm pinhole, a synchrotron source is 100 to 1000 times brighter than a conventional global source, and the broadband nature of the source is important for performing spectroscopy.

The U10B beamline at NSLS is developed for combined optical and fluorescence microscopy and IR microspectroscopy to allow chemical imaging of materials. The beamline is equipped with a Thermo Nicolet Continuum IR microscope and Magna 860 step-scan Fourier Transform Infrared (FTIR). This configuration allows simultaneous sample viewing and IR data collection, including digital video recording through 15× and 32× Schwarzchild IR objectives. Using various available beamsplitters (CaF2, KBr, and Si) and detectors (MCT-A, MCT-B, Cu-doped Ge, and bolometer), IR spectra can be collected from 50 to 5000 cm^{-1}.

Ongoing applications of synchrotron IR microspectroscopy include examination of mechanisms of heavy-metal accumulation in plants; thin metal–oxide layers and corrosion; minerals and archaeology; protein and mineral content in bone during osteoporosis and osteoarthritis; misfolded protein structure in Alzheimer's disease and scrapie; enzyme structure and biopolymerization; chemical changes in single cells during various stages of apoptosis; and submillisecond folding of proteins such as cytochrome c.

Microspectroscopy and Imaging at NSLS

A unique strength of NSLS imaging capabilities is the broad range of wavelengths available from the VUV and X-ray rings. This broad spectral range allows for imaging from the far IR through the hard X-ray regimes. A number of beamlines on both the VUV and X-ray rings are used for a wide range of imaging and microprobe techniques (* indicates currently under development): diffraction-enhanced imaging (X15A); hard X-ray fluorescence microprobe and imaging (X13B* and X26A); soft X-ray microprobe and imaging (X1A); microtomography (X2B and X27A); X-ray microdiffraction (X26A, X17B and C, and X13B*); infrared microspectroscopy and imaging (U2A, U2B, and U10B); and extreme UV (called EUV) lithography (U13UC).

Broad scientific fields of research take advantage of NSLS imaging capabilities. To date, the largest user communities in imaging come from the environmental, geological, and biomedical fields; a number of users, however, also come from the materials and polymer science communities. Only a fraction of current users take advantage of the multiple imaging and microprobe techniques at NSLS.

Lawrence Berkeley National Laboratory (LBNL)

Contact: Bahram Parvin (parvin@media.lbl.gov)
Web Site: http://vision.lbl.gov/index.htm

Instrumentation

Synchrotron radiation-based infrared (SRIR) spectromicroscopy enables the nondestructive in vitro monitoring of specific chemical information in different compartments of individual cells without treating cells with exogenous dyes, fluorescent labels, stains, or destructive techniques. Each cell in a population thus can be studied sequentially over a period from subseconds to

months. By monitoring cells individually or within a population, detecting and probing the heterogeneity of disease onset and responses to various treatments or insults is possible. The spatial resolution is 3 to 10 μm. The development of SRIR spectromicroscopy will result in a broadly applicable and powerful research tool for the scientific community.

Wide-field deconvolution and structured light microscopy allows high-resolution, 3D fluorescence imaging. Its major strengths are (1) high sensitivity across the visible spectrum, (2) ease of recording multiple-wavelength information with minimal displacement, and (3) higher spatial resolution than conventional light microscopes (about 0.2 μm in X and Y, 0.5 μm in Z). Current camera configuration allows a maximal real-time imaging mode of 17 fps but can be upgraded with available cameras. Through collaboration with the laboratory of John Sedat at the University of California, San Francisco, plans are to eventually incorporate a structured illumination mode, which is reported to allow 0.1-μm resolution in X, Y, and Z. Structured illumination provides the basis for rapid reconstruction of 3D images, which is necessary for construction of rapid responses.

Cryo-EM provides very high resolution EM imaging and tomography but is less temporal.

LBNL also has the capacity to do X-ray cryotomography as well as multiphoton, multispectral imaging. Consequently, the laboratory can correlate live-cell light and high-resolution X-ray microscopy analyses on the same specimen. A soft X-ray microscope at 300 to 900 eV with 25-nm spatial resolution produces images of 1000 by 1000 pixels, 1000 photons per pixel recorded in 3 s at 517 eV, 0.2% BW.

Informatics

A collection of unique algorithms for capturing 2D and 3D subcellular structures has been developed. Labeled structures provide the basis for protein expression for large-scale image analysis. These algorithms have been applied to fixed cells imaged with fluorescence confocal microscopy.

Novel visualization, morphing, and registration techniques have been developed to track evolution of cultured colonies and their corresponding properties in time.

The visual serving optical microscopy (VSOM) system quantifies physiological responses as a function of external exposures and uses them to elicit a specific response by manipulating external exposures. Strictly for live-cell studies, VSOM can monitor hundreds of cells simultaneously for statistical validation, tracking each cell spatially and measuring uptakes and retention factors for each desirable component. Plans are under way to use these physiological responses to estimate parameters such as kinetic orders and rate constants.

An imaging bioinformatics system can track distant features and correlate events and responses across genetically altered cell lines or the same cell line at different times. A data model derived from multiple experimental protocols couples experimental annotations to images and their derived features for hypothesis testing and information retrieval. The system, named BioSig, is coupled to VSOM and protein-expression studies in fixed cells.

DeepView is an extensible open-system architecture for coupling imaging instruments with computational infrastructure in distant online collaborations. The system, which provides for both synchronous and asynchronous collaboration, has been integrated with five electron and optical microscopes.

Lawrence Livermore National Laboratory (LLNL)

Contacts: Rod Balhorn (balhorn2@llnl.gov), Stavros Demos (demos1@llnl.gov), Chad Talley (talley1@llnl.gov), Tom Huser (huser1@llnl.gov), and Steve Lane (slane@llnl.gov)

Optical Microscopy

LLNL has high-speed, 3D, multiwavelength (one 4-wavelength 3D stack every 0.5 to 1 s) optical microscopy capabilities for imaging chromosomal motion, signal transduction, and time-structure analysis (250-nm resolution). The

optical microscope, contained in a temperature-controlled environment ideal for live-cell imaging, allows for automatic tracking of live samples during data collection [J. Trebes, LLNL, and J. Sedat, University of California, San Francisco (UCSF)].

Electron Microscopy

LLNL can perform EM of large molecular complexes (via single-particle reconstruction or EM tomography) (0.5- to 1.0-nm resolution) and structural analysis of large molecular assemblies that are impossible to crystallize (J. Trebes, LLNL, and J. Sedat, UCSF).

X-Ray Microscopy

X-ray microscopy of in vitro cellular structures is available through synchrotron- and laser-based X-ray microscopes for high-resolution, 3D imaging of subcellular structures in both frozen and hydrated states (30-nm resolution) (J. Trebes, LLNL).

Hyperspectral Imaging and Microscopy

Hyperspectral imaging and microscopy of cells via polarization, light scattering, and autofluorescence (250-nm resolution) are available for imaging tissue and single cells using sequential, multiple-contrast techniques within a matter of seconds (S. Demos, LLNL).

Atomic Force Microscopy

AFM is available for imaging single virus particles, proteins, and biomolecular complexes (1- to 3-nm resolution) (J. De Yoreo, C. Orme, A. Noy, and R. Balhorn, LLNL). High-resolution force microscopy involving carbon-nanotube tips can resolve individual domains in proteins and is easily used in vitro. Smart-probe AFM using chemically modified tips can monitor changes in cell surface or receptors upon stimuli (10-nm resolution) (R. Balhorn, A. Noy, and T. Huser, LLNL).

Single-Cell Microscopy

Capabilities at LLNL include single-cell Raman microscopy to obtain information about composition, structural changes, and changes upon stimuli via either single laser source or coherent anti-Stokes Raman spectroscopy (CARS) imaging (250-nm resolution) (R. Balhorn, T. Huser, and S. Lane, LLNL). Single-cell Raman imaging based on confocal microscopy is available with 300-nm spatial resolution. This instrument can generate images of single living cells at specific Raman lines with CARS or conventional micro-Raman scattering and can obtain full Raman spectral imaging of subcellular compartments.

Single-molecule imaging techniques include fluorescence inside cells via genes with green fluorescent protein (GFP) tag to watch expression under different conditions or morphogenesis, surface-enhanced Raman spectroscopy (250-nm resolution) (R. Balhorn, A. Noy, C. Hollars, C. Talley, S. Lane, and T. Huser, LLNL), and fluorescence and surface-enhanced Raman microscopy and spectroscopy based on a confocal microscope with a spatial resolution of 300 nm.

Other Capabilities

- Ion-probe microscopies of living cells (PIXE, 1000-nm resolution), STIMS, nanosecondary ion mass spectrometry (SIMS) (50-nm resolution), and single-cell X-ray absorption near-edge (called XANES) that provide material-dependent 2D (PIXE) and 3D SIMS maps of samples (R. Balhorn).

- Near-field scanning microscopy (50- to 100-nm resolution) for simultaneous topographic and subwavelength imaging (T. Huser, C. Hollars, and C. Talley).

- Single DNA-manipulation and video-imaging system (250-nm resolution) with an optical trap that manipulates single DNA molecules in multisolution flow chambers (R. Balhorn).

- Advanced fluorescence microscopy (250 nm) with fluorescence lifetime imaging; two-photon excitation imaging; and multiwavelength video imaging of single molecules in flow channels (T. Huser, C. Hollars, C. Talley, O. Bakajin, and S. Lane).

- Novel labels for imaging based on semiconductor quantum dots and gold and silver light-scattering nanoparticles (S. Lane, T. Huser, G. Galli, and G. Fox).

Los Alamos National Laboratory (LANL)

Contacts: Robert Donohoe (rjd@lanl.gov), Larry Hersman (hersman@lanl.gov), and Judith Mourant (jmourant@lanl.gov)

Molecular-Imaging Capabilities

- **IR focal plane array (MCT) microscope (Spectra Tech)** coupled to a step-scan FTIR for continuous-wave (CW) and time-resolved molecular imaging (a full interferogram is created at each pixel). The spatial resolution currently is twice the diffraction limit (~10 μm), and coupling of the step-scan bench to a pump-and-probe trigger experiment can yield microsecond time resolutions. Current work aims to improve the S/N by several methods but has not yet explored tuning-range (called TR) capabilities, which have been used with the step-scan bench on other occasions.

- **Epi-CARS and two-photon fluorescence microscope** based on two synched mode-locked titanium:sapphire lasers. Spatial resolution is roughly 200 nm, and we have used this instrument to examine the uptake of fluorescent beads and GFP-labeled bacteria by white cells. The independent tuning of the two lasers allows full "fingerprint" vibrational CARS imaging, although we must first overcome difficulties in the synchronization apparatus. Conventional CW lasers also are available for typical fluorescence microscopic measurements.

- **Near-field fluorescence imaging.** This capability is using a wave guide to collect near-field fluorescence from quantum dots. The spatial resolution is roughly 50 nm. There is no time-resolution capability as yet.

- **Conventional Raman microscope** using a Zeiss inverted-axis microscope coupled to CW argon and titanium:sapphire lasers. Detection of hyper-spectral images is based on holographic single-stage spectrographs with CCD cameras. Spatial resolution is roughly diffraction limit (500 to 300 nm), and images take from minutes to hours to collect.

- **Development of radiopharmaceuticals** for imaging of target organs (~3-mm spatial resolution) using positron emission tomography (called PET) technologies.

- **Single-molecule detection.** This is not so much an imaging capability per se as a capability that uses spatial and temporal resolution to detect single molecules and monitor their dynamics, including fluorescence decay.

- **Data analysis.** Genetic learning algorithms and multiscale, multiresolution algorithms (often based on wavelet transformations) are capable of improved feature extraction compared with PCA and SVD methods.

Oak Ridge National Laboratory (ORNL)

Contact: Mitch Doktycz (doktyczmj@ornl.gov)

Imaging Technologies

Molecular-Recognition Force Microscopy (MRFM)

MRFM provides molecular-level spatial resolution of proteins present on cell surfaces. In addition, this technique can be combined with optical techniques to characterize intracellular components.

Topographic and Electrostatic Imaging of Single Molecules

This technique can measure the topographic and electrostatic properties of single molecules in real time. As an example, photovoltages from single-molecule Photosystem I reaction centers have been measured both on atomically flat gold surfaces and in room-temperature aqueous environment. The resolution of the electrostatic maps enabled us to measure the fine structure of the ferredoxin docking site at the reducing end of Photosystem I.

Nonlinear Photonic Scanning Microscopy (NPSM)

NPSM is used with simultaneous fluorescence spectroscopy of single molecules having a lateral resolution of 2.9 nm and a normal resolution of

0.1 nm. Using two-photon fluorophores, the nonlinear tunneling provides signals acquired without the usual problems of fluorophore decay.

Scanning Tunneling Microscopy (STM)

ORNL has developed a photon STM that can provide optical images and localized spectroscopy (in air and fluid) of biomaterials with a spatial resolution of 40 nm.

Atomic Force Microscopy

The spatial resolution ranges from subnanometer to a few nanometers for AFM-based techniques.

Novel Scanning Microscopes

ORNL has modified conventional scanning probe microscopes (STM and AFM) to yield structural, physical, and chemical information not readily available in conventional scanning microscopes. The spatial resolution ranges from subnanometer to tens of nanometers depending on the basic physics involved. In addition, most of these techniques can be used under solution. For example, microcantilever tips modified with biological probes can be used to locate the receptor sites on cell surfaces with nanometer precision (chemical force imaging). Combining chemical force information with phase-enhanced topography can create a chemomechanical topography of cell surfaces and biomolecules.

Optical Nanosensors for Single-Cell Analysis

Nanobiosensors are based on submicron optical fibers for in situ intracellular measurements of single-cell, antibody-based nanoprobes with spatial resolution of 30 nm (probe-tip diameter) and temporal resolution of nanoseconds (fluorescence lifetime).

Multispectral Imaging (MSI)

MSI is accomplished using an acousto-optic tunable filter (AOTF) coupled to an optical microscope. This imaging modality can visualize molecules based on many different spectroscopic techniques (e.g., fluorescence, phosphorescence, reflection, and Raman spectroscopy), thus providing real-time, in vivo information on electronic as well as vibrational information for any type of molecule. It has a spatial resolution of 250 μm and a temporal resolution of nanoseconds or shorter.

Near-Field Scanning Optical Microscopy (NSOM) and Atomic Force Miscroscopy

This technique is capable of monitoring fluorescence or Raman signals as well as topography. These two modes will provide information about the electronic, vibrational, and rotational structures of the molecules. NSOM has a spatial resolution of ~50 nm, and AFM has nanometer-to-subnanometer resolution.

Two-Photon Fluorescence Imaging (TPFI)

TPFI is capable of monitoring the fluorescence signal from 3D objects with a theoretical spatial resolution smaller than the diffraction limit of light. This is particularly useful for in vivo measurements of the electronic and molecular properties of molecules that have been immobilized in a 3D matrix. The system will provide resolution capable of approaching the molecular level, and its temporal resolution will be on the nanosecond regime or smaller.

UV Resonance Raman Microscope

This system can image molecules based on their vibrational properties. In addition to conventional Raman microscope features, the UV resonance Raman microscope is particularly well suited for stimulating the resonance Raman vibrational modes of macromolecules such as proteins. This feature enables highly sensitive detection while minimizing fluorescence background. An intensified CCD offers 5-ms acquisition times in the continuous-running mode. Special gated-mode imaging can be achieved with acquisition times as low as 5 ns.

Surface-Enhanced Raman Scattering (SERS) Microscopy

The Raman-based microscope can be used with SERS-inducing media for dramatically improved sensitivity. The SERS process typically can enhance Raman scattering signals 106 to 108 times. Temporal, spectral, and optical resolutions of the Raman microscope are not significantly affected by the SERS effect.

Surface-Enhanced Raman Scattering Nanoprobes

This system offers the spectral and chemical selectivity of SERS with highly localized and position-controlled detection. This regime

includes an inverted microscope equipped with probe micromanipulators, an AOTF and a CCD. Optical fibers can be produced with extremely small tips (e.g., 100 nm) through special tapering processes implemented in our laboratory. The intensified CCD offers 5-ms acquisition times in the continuous-running mode. Special gated-mode imaging can be achieved with acquisition times as low as 5 ns.

Confocal Raman Imaging

This scanning-based mode of imaging is appropriate for applications requiring resolution at the diffraction-limited level but for which a high level of temporal resolution is not a key factor. Furthermore, samples should be distributed in a planar format. The pinhole-based optical system virtually eliminates optical aberrations induced by lens imperfections, yielding optical resolution at the diffraction limit.

Pacific Northwest National Laboratory (PNNL)

Contacts: Steve Colson (steven.colson@pnl.gov), Robert Wind (Robert.Wind@pnl.gov), Steve Wiley (steven.wiley@pnl.gov), Don Jones (dr.jones@pnl.gov), Harold Trease (het@pnl.gov), and Paul Majors (paul.majors@pnl.gov)

Cellular Observatory Instrumentation

PNNL's cellular observatory houses a collection of unique instruments:

- **CARS and two-photon confocal combined microscope.** For visualizing molecules based on their vibrational properties and simultaneously by their fluorescent properties, CARS can visualize selected molecular species, such as lipids and deuterated compounds. Spatial resolution is 0.3 μm.

- **AFM-enhanced fluorescence-imaging microscopy.** This novel nanoscale-characterization combined microscope uses the optical or AFM approach and a high-sensitivity far-field microscope to provide unperturbed measurements of reaction rates at sites that have been characterized by AFM imaging.

- **Single-molecule spectroscopy and imaging microscopy.** This instrument does simultaneous structural and spectroscopic analyses of single biomolecular complexes and their reaction (interaction and association) rates, providing insight into the relationship between structure and function of cell-signaling proteins and enzymes. Information collected from these approaches is typically lost or hidden in measurements using current conventional ensemble-averaged methodologies.

- **Single-molecule patch-clamp–optical confocal imaging microscopy.** This combines a confocal scanning linear and nonlinear optical microscope with state-of-the-art patch-clamp technologies. The instrumentation significantly enhances the diagnostic and investigative capabilities of both methods in characterizing ion channel and receptor dynamics and mechanism in a living cell.

- **High-speed confocal microscope for 3D fluorescence resonance energy transfer (called FRET) in living cells.** The microscope combines a Nipkow-disk–based confocal scanning head with dual Gen III–intensified CCD cameras and a fast AOTF coupled to a high-powered argon: krypton (called Ar-Kr) laser system. High-resolution images from both cameras (1000 × 1000 × 12 bits) are simultaneously captured and spooled to high-speed RAID drives at 30 fps. XYZ scanning allows 3D images to be rapidly acquired, and custom software provides for FRET and ratiometric analyses to be performed in near-real time.

Magnetic Resonance Microscopy

Several MRI capabilities are available at PNNL. These microimaging systems include

- Varian Unity + 7T vertical wide bore (89 mm)
- Varian Unity + 12T vertical wide bore (89 mm)
- Vertical Bruker Biospin 12T wide bore (89 MM)

All these spectrometers are equipped with a variety of imaging probes optimized for MR microscopy of samples varying in size from 200 μm to 4 cm, probes for (respiratory-gated) in

vivo MRI and spectroscopy of live mice and other mammals of a similar size, and a combined confocal and magnetic resonance microscope, as described in the section regarding the cellular observatory.

Several probes are equipped with sample chambers that can be made part of a perfusion system, so that live biological cell systems and the dynamics of cellular processes can be studied.

Multisensor Image Registration

Novel mathematical methods and algorithms are being developed for spatially coregistering 2D and 3D images of cellular and tissue structures and dynamic biological processes. These are being adapted from methods originally developed for registration of multisensor satellite imagery. Image data are obtained by several imaging modalities including confocal microscopy, NMR, and CARS.

Feature Extraction and Object Classification

PNNL is developing a suite of statistically based methods as well as analysis tools to detect locations of change and extract feature in fused, multisensor imagery. Two new methods (inverse kriging and conditional spectral mixing) are applied to imagery to improve the spatial resolution of image features. Statistical techniques are then used to identify and classify features or objects within the image. Methods for detecting and quantifying areas of change between successive images also are being developed. Taken together, these tools enhance analyses with improved capabilities to identify and classify features and monitor changes in multisensor images.

Three-Dimensional Visualization and Analysis Interaction Systems

Innovative 3D visualization and human-interaction technologies are being developed that greatly enhance the ability of scientists to interact with and understand biomedical imagery. New 3D display and "augmented reality" visualization technologies enable scientists to view imagery in 3D in an interactive manner. The core of this technology is the "Human-Interaction Workspace," a tabletop display coupled with algorithms that enable human gestures to control the image display. The intent of this technology is to enable scientists to gain deeper and more rapid insight into complex biological processes.

Parallel Computing Environment for Imaging Science

Decreasing the time to solve many digital image-processing tasks using parallel computing methods can also decrease the number of tasks. Our current research is to develop a working environment that is scalable, portable, and with near real-time image processing, steering, and dynamic load balancing. We have been able to achieve increased performance by processing a 300-MB image file in less than 6 s. Additionally, the system can display 9.2 Megapixels (24-bit) with full real-time interaction.

Image-Based 3D Reconstruction and Modeling

NWGrid, a 3D mesh-generation system under development, allows users to perform image segmentation, mesh generation, and mesh optimization on volume digital-image data. These parallel algorithms do 3D image reconstruction and feature extraction to produce geometry from which computable volume meshes are generated. We have successfully modeled a lung and a rat's respiratory system.

Architecture for Cross-Domain Image Analysis

Although image analysis is a key element of a wide range of research domains, there are no standards for interoperability among domains. Each domain tends to use a different set of software tools and formats. Even within a given domain, there often are a variety of image formats and tools, many of which do not "play well together." We have designed an extensible open architecture that supports interoperability among a broad range of tools and formats. The architecture includes configuration and deployment management that can extend beyond a single enterprise to support remote collaborators as well.

Novel Interactions and Visual Analytics

The effective exploitation of complex, multidimensional image data is a nontrivial challenge. We are developing visualization tools and new interactive methods to support the collaborative analysis of image data. This includes such techniques as the use of augmented reality (AR) to support 3D interactions with data, as well as remote collaboration via AR-based telerepresentation of collaborators, and direct user interaction with data using hands instead of a mouse.

Segmenting 3D Volumetric Image Data

The digital data viewer (DDV) provides an interactive interface for segmenting, comparing, and overlaying 3D volumetric-image data. The data may be a 3D stack of 2D slices (e.g., with MRI scans, confocal images, or TEM images) or from any source of volumetric data. Segmentation rules are propagated through volume regions. Algorithms for image enhancement, smoothing, and correction of distortion errors are included. Fast, robust isosurface and surface-smoothing algorithms are also part of DDV.

Sandia National Laboratories (SNL)

Hyperspectral Imaging

Contact: David Haaland (dmhaala@sandia.gov)

SNL has designed and built a hyperspectral-imaging system for DNA microarrays, cells, and other biological systems via fluorescence emission excited in the visible region with 10-μm spatial resolution. The spatial resolution can readily be extended to the diffraction limit. At least three laser wavelengths are available for excitation. The system uses a sensitive electron-multiplying CCD detector and therefore is ideal for low-light imaging applications. Because an entire emission spectrum is collected rather than a discrete band of light, multiple and even overlapping emissions can be monitored simultaneously and separated from one another. When combined with our multivariate curve-resolution software, pure fluorescence spectra of emitting species can be discovered and quantitative spatial maps of the fluorescing species can be obtained, even when no standards are available.

Multivariate Image Processing

Contact: David Haaland (dmhaala@sandia.gov)

SNL's team of researchers has developed and improved multivariate methods for quantitative and qualitative analyses. These methods, which are applicable to a wide variety of spectral data, allow the extraction of more information than is possible with conventional image analysis. Recent work with multivariate curve resolution (MCR) has shown the techniques to be very powerful in extracting highly overlapped spectral signatures from hyperspectral images of microarray gene-expression data and yeast cells expressing GFP. MCR accurately identifies the spectral species and determines the concentration of each species in every pixel even when no standards are available, thus improving quantitative accuracy and providing correction for impurity and background-emission sources.

Macro- and Microscopic FTIR Imaging

Contact: Christine M. Wehlburg (cmwehlb@sandia.gov)

SNL's Biorad Stingray FTIR Hyperspectral Imager uses a 64- by 64-pixel array detector to capture hyperspectral image cubes of either a macro (4- by 4-mm) or micro (400- by 400-μm) sample area over the frequency range of 950 to 3800 cm^{-1}, which is the typical IR "fingerprint" region of the mid-IR spectral region. The spatial resolution of the 64- by 64-array detector is ~60 μm in macro mode and approaches the IR diffraction limit of ~6 μm resolution in the micro mode. Together with sophisticated multivariate calibration and classification methods, we have addressed a variety of materials analysis, aging, and characterization problems.

Atomic Force Microscopy

Contact: Alan Burns (aburns@sandia.gov)

SNL has two AFM instruments dedicated to imaging biological interfaces such as membranes in fluid environments. The first is a Digital Instruments Multimode with assorted scanners that have a range up to 120 μm. The second instrument, which has an inverted microscope base for simultaneous fluorescence and AFM

imaging, uses a piezo plate for sample scanning. Near-field techniques are employed to push the optical resolution to <100 nm. For both instruments, the AFM topographic resolution is ±1 nm lateral and 0.1 nm vertical.

Electron Microscopy

Transmission Electron Microscopy

Contacts: Tom Headley (tjheadl@sandia.gov) and Paul Kotula (pgkotul@sandia.gov)

SNL has several TEM systems for materials characterization. The Philips CM30 provides low Z, full spectral (elemental) imaging of energy-dispersive spectrometer (EDS) and EELS signals. It operates at 300 kV with 0.14-nm lattice imaging and 0.23-nm point-to-point resolution on electron-transparent samples. SNL also has an FEG-TEM providing phase imaging via spectral imaging of EDS and EELS signals at interfaces. This instrument also operates at 300 kV with 0.19 pt-to-pt resolution.

Scanning Electron Microscopy (SEM)

Contact: Joe Michael (jrmicha@sandia.gov)

SNL can provide high-resolution imaging via several different SEM systems. SEM capabilities have been used to characterize the microstructure, elemental composition, and morphology of bugs, dust pollen, ash, and particles. Capabilities include SE 1.5 nm at 30 kV, BSE 10 nm at 5 kV, and FIB/SEM with dual column (FEG column resolution, 2.0 nm; LMIS ion column resolution, 5.0 nm); sample chamber permits viewing low vacuum for charging or wet samples, large specimen capacity (150 × 150 × 150 mm and <5kg), and automated image analysis. Data and images are available via the Internet to the customer's desktop.

Time-of-Flight Secondary Ion Mass Spectrometry (SIMS)

Contact: Tony Ohlhausen (jaohlha@sandia.gov)

TOF-SIMS is a surface analytical technique capable of identifying molecular and elemental species at the PPM-PPB level. Because it is a mass spectral technique, TOF-SIMS is well suited for organic identification based on the unique spectral "fingerprint" caused by fragmentation of the analyzed species. It is capable of distinguishing elemental isotopes and detecting all elements on the periodic table. The TOF-SIMS instrument at Sandia can perform depth profiles at a submonolayer depth resolution and ion imaging at 0.25-μm lateral resolution. Typical uses of TOF-SIMS include monolayer and particle analysis, molecular-weight distribution analysis of polymers, contamination analysis, and aging studies of surfaces. Typical substrates include metal coupons, silicon wafers, paper, polymer films, ceramics, and biological specimens.

Optical Microscopy

In addition to the specialized techniques listed above, SNL has multiple traditional optical microscopes available for imaging microbes, cells, and tissues with various modes of contrast. These include fluorescence emission, confocal fluorescence emission, and phase contrast.

Thomas Jefferson National Accelerator Facility

Contacts: Gwyn Williams (gwyn@mailaps.org) and Stan Majewski (majewski@jlab.org)

X-Ray Microscopy

This subcellular imaging method provides elemental image maps of carbon, oxygen, calcium, and other elements in specific chemical states. The spatial resolution is 300 nm, and the time resolution is 1 min.

Far-Field IR Microspectroscopy

This subcellular novel imaging method uses specific protein, lipid, mineral, and other specific vibrational modes as a contrast mechanism. The technique is enabled through the use of bright synchrotron radiation. Spatial resolution is 3 μm. Temporal resolution is typically 1 s, but special pump-probe techniques can provide temporal resolution to subpicosecond time scales.

Near-Field IR Microscopy

This imaging method uses specific protein, lipid, and other specific vibrational modes as a contrast mechanism. The technique is enabled by a unique, tunable free-electron laser. Spatial resolution is 10 to 100 nm, and temporal resolution is as high as 1 ps.

Pulsed High-Contrast X-Ray Imaging

Femtosecond X-ray pulses are used to provide phase and absorption contrast in X rays of tissue. Spatial resolution is 1 mm. Temporal resolution is subpicosecond.

Terahertz Imaging and Spectroscopy

Submillimeter waves are used to recognize specific proteins as an image-contrast mechanism. Examples are basal-cell carcinoma screening and early dental-cavity detection. Spatial resolution is 0.3 mm, and temporal resolution is 1 s; with pump-probe switching, however, dynamics at the picosecond time scale can be imaged.

X-Ray Diffraction (XRD)

Jefferson Lab provides XRD of macromolecular samples such as proteins, with a spatial resolution of 2 Å and temporal resolution of 10 s.

Magnetic Resonance Imaging

Jefferson Lab has the capability to produce unique MRI imaging molecules, an example being the use of the free-electron laser to encapsulate gadolinium in fullerenes as well as to functionalize the cage complex with a water-soluble organic group. Radioactive elements such as holmium also can be encapsulated as tracers or anticancer agents.

Large-Scale Computing and Data-Handling Capabilities

These capabilities, developed for nuclear physics, can be readily adapted. They include high-performance systems for data acquisition and control; petabyte-scale systems for data storage and management; geographically distributed software infrastructure for data analysis (advanced data grid); and secure, high-speed networking. Such capabilities and corresponding expertise at Jefferson Lab could be leveraged to benefit an advanced medical-imaging program that would include centralized data storage (disk and tape), Web-based data serving with strong authentication, and assistance in designing and deploying a fully distributed infrastructure for data ingestion and export.

Genomes to Life Program and Facilities Planning Workshops

A series of program-planning workshops has been held to help plan and coordinate Genomes to Life. Meeting reports are placed on the Web as soon as they become available. To learn more about the program, please see the Web site (DOEGenomesToLife.org).

Web site for workshop reports: DOEGenomesToLife.org/pubs.html

2000

October 29–November 1	Genomes to Life Roadmap Planning, San Diego

2001

January 25–27	Genomes to Life Roadmap Planning, Germantown, Md.
June 23	Role of Biotechnology in Mitigating Greenhouse Gas Concentrations, Arlington, Va.
August 7–8	Computational Biology, Germantown, Md.
September 6–7	Computational and Systems Biology, Washington, D.C.
September 9–10	Science Mission Payoffs, Washington, D.C.
October 24–25	Energy and Climate Mission Payoffs, Chicago
December 10–11	Technology Assessment for Mass Spectrometry, Washington, D.C.

2002

January 22–23	Computational Infrastructure, Gaithersburg, Md.
March 6–7	Computer Science, Gaithersburg, Md.
March 18–19	Mathematics, Gaithersburg, Md.
April 16–18	Imaging, Charlotte, N.C.
April 16–19	Computing Strategies, Oak Ridge, Tenn.
June 19–20	Facilities Planning, San Francisco
August 16–17	Facilities Planning, Chicago
October 14–15	Facilities Planning, Gaithersburg, Md.

U.S. Department of Energy
Office of Biological and Environmental Research (SC-72)
Office of Advanced Scientific Computing Research (SC-30)
Germantown Building
1000 Independence Ave., NW
Washington, DC 20585-1290