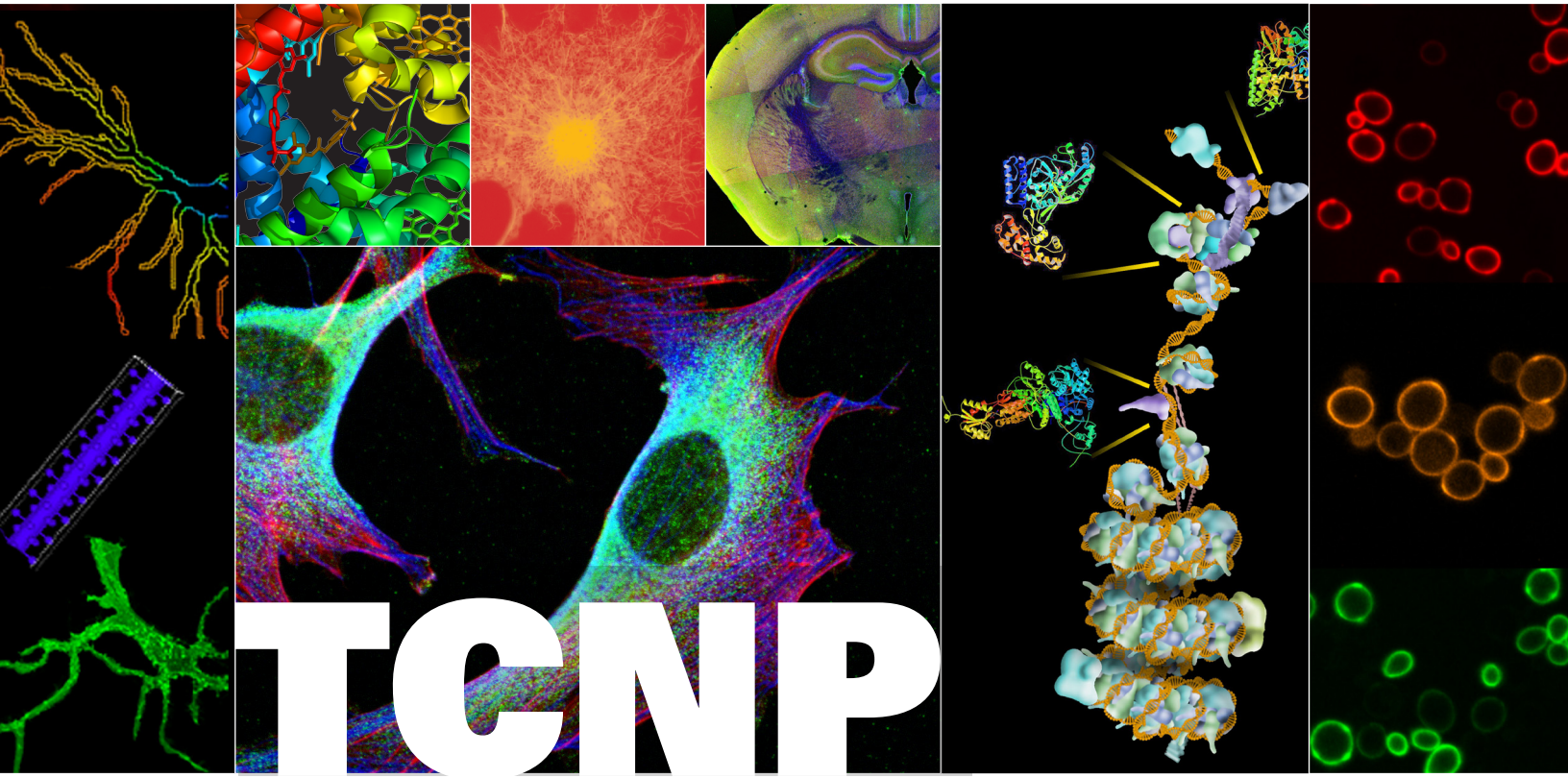


National Technology Centers for

Networks and Pathways

Technology and Resources for Proteomics of Dynamic Systems

NIH National Institutes of Health

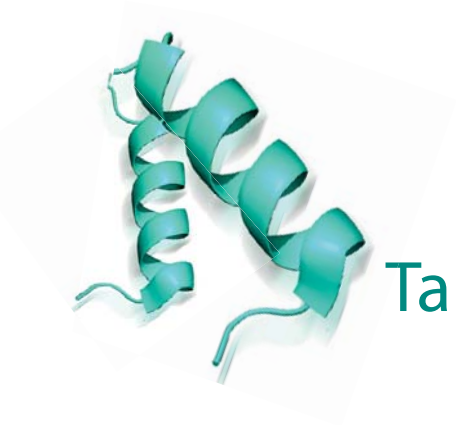
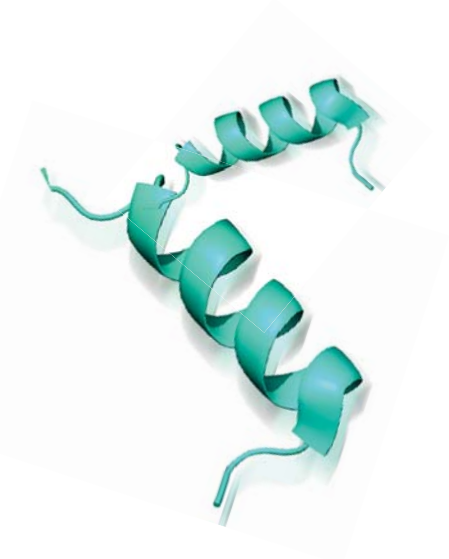


Table of Contents

An Introduction to the National Technology Centers for Networks and Pathways

The limitations of proteomics technologies often force investigators to treat dynamic systems artificially, as if they were static. As with early photography, current approaches to proteomics often impose long exposure times on the "subjects," and the "images" captured are vague and broadly defined, such as "normal vs. diseased," "the yeast interactome," or "the nuclear pore complex." Inadequate tools make us blind to the dynamics of biological systems. Even though we know they cannot be static, we must treat them as such for the time being. The result is that transient interactions or rapid changes in protein activity, location, or post-translational modification are just a blur—like a bird flying through the frame of a carefully composed, long-exposure photograph.



Complementary strategies, sometimes building on, sometimes quite independent of those used in conventional proteomics, are necessary to take science to the next level where the focus is on dynamic processes. The National Technology Centers for Networks and Pathways (TCNPs) are answering this challenge by creating technologies to measure and reveal the ever-changing nature of protein interactions, modifications, translocation, expression, and activity, and to do so at very high temporal, spatial, and quantitative resolution. The program has built a bridge between the quantitative and interaction domains, allowing us to break out of our artificially static view of complex systems.

A Closer Look at the Program

The National Technology Centers for Networks and Pathways (TCNP) program is a trans-NIH initiative supported through the NIH Common Fund. The program funds multidisciplinary research centers that work together as a consortium, with the common goal of creating new technologies to study the proteomics of dynamic systems at the highest possible level of quantitative, spatial, and temporal resolution. The centers seek to integrate the study of proteomics with cell biology, imaging, and

modeling to achieve a dynamic picture of cellular systems at the molecular level. These centers are unique among NIH-supported proteomics programs because they focus not just on protein expression, but also on protein interactions, and they emphasize imaging and quantitative methods.

TCNP funds principally support technological innovation. However, the centers are also a resource for NIH-supported investigators, committing substantial resources to collaboration with and training of biomedical researchers. The centers provide a broad range of reagents, vectors, cell lines, technologies, and expertise to hundreds of researchers every year.

In phase I of the program (FY 2004–2008), five National Technology Centers for Networks and Pathways were funded through cooperative agreements. Two centers were funded in September 2004, and three more in September 2005. Awards ranged between $12 million and $17 million over 5 years. After a midcourse review of the program in 2007 and 2008, an open re-competition was held in 2009 to support three centers for an additional 5 years (phase II, FY 2009–2013). The newly funded centers were all among the original five from phase I.

The National Institute of General Medical Sciences (NIGMS) administers these awards, and the TCNP Project Team manages them. Each center has an assigned NIH Science Officer, drawn from this Project Team. The Program Director for all of the centers is the Project Team Leader.

How the TCNPs Work

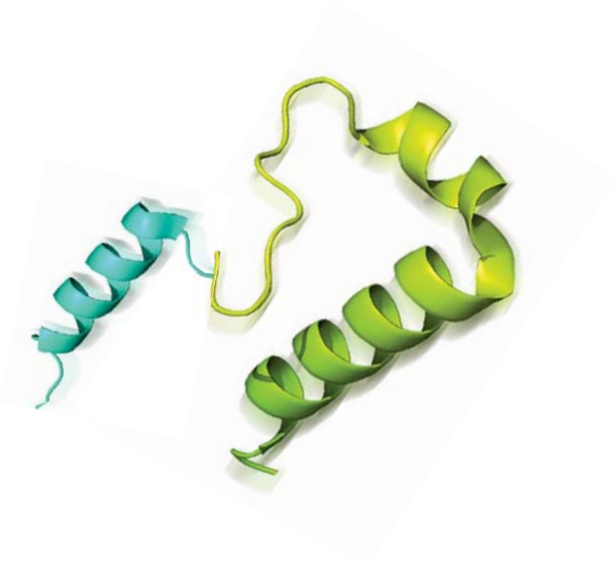
The TCNPs cooperate in a networked national effort to develop instrumentation, biophysical methods, reagents, and infrastructure for temporal and spatial characterization of complex biochemical pathways and interaction networks. They collaborate with biomedical researchers in a kind of "push-pull" relationship using technologic innovations to solve biomedical problems and advancing technology as they solve the problems. The centers are also tasked with providing broad access to the technologies, methods, and reagents they develop, and with providing appropriate interdisciplinary academic and peer training for biomedical researchers.

The scope of the challenge—revealing the dynamic interplay of the many factors involved in complex cellular systems and biological pathways—is very broad. Each center integrates multiple approaches to create a coherent biological, analytical, and informatics strategy, relying upon the strengths of different technologies and systems. An important feature of the TCNP program is its strongly cooperative and interactive nature, which allows the centers to complement one another.

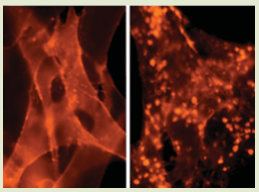
TCNPs Contribute Scientific Insights

In what follows we tell the story of each center and of the complex systems it is seeking to increase our understanding of. Each section presents an overview of the center, the technology and tools it is using to interrogate biological networks and pathways, a few research highlights, and the center's resources available to the wider scientific community.

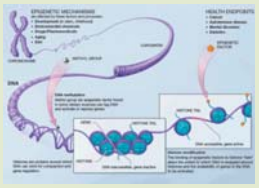
Scientists in the TCNP program have developed a broad range of technology platforms, methods, cells, and reagents that are now freely available. The centers continue to work with researchers who need access to the tools and expertise they have to offer. These tools will continue to be available, as support for this work transitions to other NIH programs, providing important insights into the networks and pathways that are important in human health and disease.



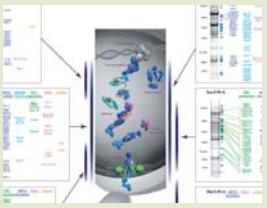
Centers



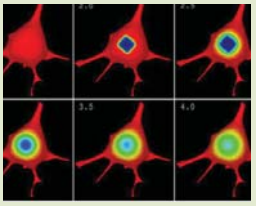
Molecular Biosensor and Imaging Center (MBIC)



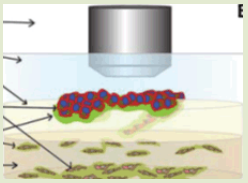
Networks & Pathways of Lysine Modification



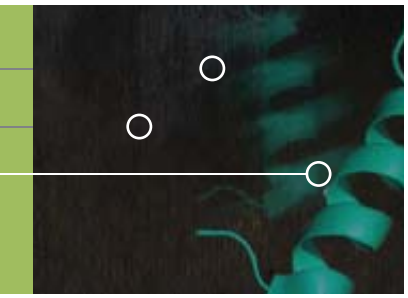
National Center for Dynamic Interactome Research (NCDIR)



R.D. Berlin Center for Cell Analysis and Modeling (CCAM)



Center on Proteolytic Pathways (CPP)



Molecular Biosensor and Imaging Center

Fluorescent Probes and Imaging for Networks and Pathways
Carnegie Mellon University
Principal Investigator: Alan Waggoner, Ph.D.
www.mbic.cmu.edu

Introduction

Biosensors:
A device that monitors and transmits information about a life process; especially: a device consisting of a biological component (as an enzyme or bacterium) that reacts with a target substance and a signal-generating electrochemical component that detects the resulting products or by-products.

Within any human cell, hundreds, if not thousands, of proteins interact in complex networks when the cell is called upon to yield a necessary product, such as an enzyme or hormone, or to respond to an environmental stress. Understanding the elements and dynamics of these networks in real time—that is, while they are happening — could provide the keys for understanding how cells malfunction in disease and for developing therapies to repair or bypass that dysfunction by selective and sensitive modulation of protein function. This challenge requires that scientists in academia and the pharmaceutical industry be able to quantify the activities of key proteins of the pathways in living cells. Yet, until now, there have been few probes or biosensors that offer this capability, even though the instru-mentation is widely available to detect signals from such probes.

The Molecular Biosensor and Imaging Center (MBIC)

The Molecular Biosensor and Imaging Center (MBIC), one of the NIGMS-administered National Technology Centers for Networks and Pathways, is developing a powerful toolbox of novel fluorescent biosensors to study in detail how proteins interact and change—in the 3-dimensional space of a living cell and in real time. Headquartered at Carnegie Mellon University, MBIC is a partnership between that university and the University of Pittsburgh. The multidisciplinary MBIC research team, led by principal investigator Alan Waggoner, comprises biologists, computer scientists, biophysicists, electrical engineers, synthetic chemists, chemical engineers, and biomedical engineers. The powerful molecular biosensors being developed at MBIC are being used to map the many cellular networks implicated in health and disease. In the future, such biosensors may be used in hospital- and office-based diagnostic medicine.

The sensor units are generated by combining genetically engineered, target-binding proteins with non-fluorescing fluorogen dyes that light up only when the target molecule

Fluorogen Dyes:
Fluorogenic refers to a process in which fluorescence is generated. A fluorogen is typically an organic dye that is "dark," meaning that it absorbs light, undergoes an excitation, but does not re-emit that energy in the form of light, rather it transfers that energy into vibrational or torsional motion.

Interstitial Spaces:
The space between cells or tissues in the body.

Phosphorylation: pg.3
A process in which a phosphate group is added to a molecule, such as a sugar or a protein. Phosphorylation is one mechanism by which a protein's function can be turned on or off.

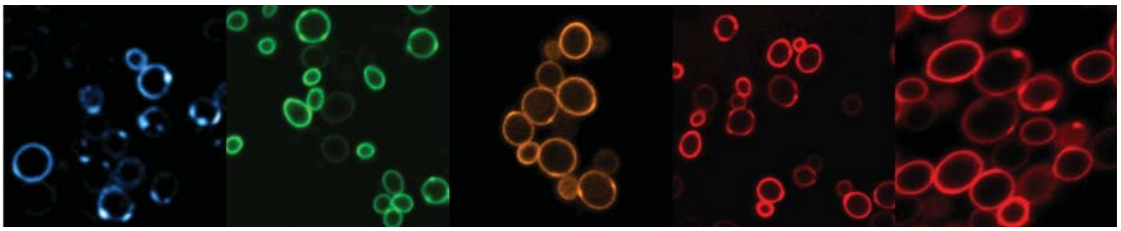
Peptides: pg.3
Molecules that contain two or more amino acids (the molecules that join together to form proteins).

Recombinant: pg.3
Recombinant proteins result when DNA sequences from one organizm are expressed in a living cell of another (host) organism.

is bound, thereby eliminating the problem of background fluorescence that plagues many fluorescence-based cellular imaging technologies. The sensor units are then incorporated into intracellular sensors, sensor particles, or optical fiber sensors for interstitial spaces in tissues, sensors on chips for *in vitro* assays, and sensors for high-throughput automated assays used in drug discovery. Figure 1 shows the range of fluorescent colors that can be generated when fluoromodules constructed with different fluorogen dyes are activated by binding to their cellular targets.

Figure 1.

Fluorogen dyes bound to engineered target-binding proteins fluoresce after the protein components bind to their molecular targets on yeast cell surfaces: OTB (blue), TO1 (green), CN-DIR (orange), MG (scarlet), and DIR (deep red)



Kimberly J. Zanotti, Gloria L. Silva, Yehuda Creeger, Kelly L. Robertson, Alan S. Waggoner, Peter B. Berget and Bruce A. Armitage. Blue fluorescent dye-protein complexes based on fluorogenic cyanine dyes and single chain antibody fragments. (2011) Org. Biomol. Chem., 9, 1012.

Szent-Gyorgyi C, Schmidt BF, Creeger Y, et al. Fluorogen-activating single-chain antibodies for imaging cell surface proteins. Nat Biotechnol. 2008;26:235–40.

The beauty of the biosensor system lies in its flexibility: by using a modular approach, researchers can create sensor units that can be readily adapted for imaging different important biological regulatory molecules --targets -- in the space within and around living biological cells. By using a combination of biosensors with different fluorogen dyes, it is possible to visualize several different cellular processes at the same time. Not only does MBIC have an extensive toolkit for constructing biosensors, but also it provides reearchers with a wide array of powerful optical detection and imaging informatics technologies, allowing them to detect and visualize molecular interactions and modifications occuring

within living cells.

This new paradigm offers the potential to create hundreds of new probes that could provide much of the valuable pathway activity information sought today. The biosensors developed by MBIC researchers are being used in a host of studies, for example, to quantify the surface density and trafficking of cell membrane receptors, and to detect protein modifications, such as phosphorylation, that regulate pathways and networks.

Highlights

Fluoregen-activating peptides

MBIC scientists have developed fluorogen-activating peptides (FAPs) that can be used to monitor in real time the biological activities of individual proteins and other biological molecules within living cells. A FAP is a specialized single-chain antibody (scFv), a recombinant fragment of full-size human antibody proteins.

The FAP attaches to a cellular protein of interest. When the fluorogen binds to the FAP, fluorescent light is emitted, alerting researchers to the protein's location and activity. In collaboration with the Scripps Research Institute, MBIC scientists use x-ray crystallography to determine the 3-dimensional structure of FAPs when they are complexed with their corresponding fluorogen (Figure 2). Such detailed structural knowledge allows FAPS to be improved and new biosensors to be developed.

Investigators

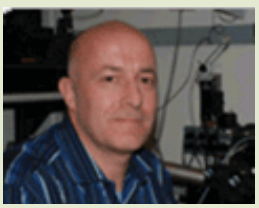
Alan S. Waggoner, Ph.D., professor of biological sciences and director of MBIC, is the center's principal investigator. In 1999, he established the MBIC at Carnegie Mellon to focus on development of fluorescence detection tools for biomedical research and drug discovery.

Bruce Armitage, Ph.D., professor of chemistry, is the co-director of the Center for Nucleic Acids Science and Technology at Carnegie Mellon. His primary contribution to the MBIC has been the development of new fluoromodules, including both synthesis of new fluorogenic dyes and selection of scFv proteins that bind and activate the fluorescence of the dyes.

Marcel Bruchez, Ph.D., associate professor of chemistry and biological sciences/technology is the associate director and program manager of MBIC. For MBIC, Dr. Bruchez is applying fluorescent microscopy and fluorescent probes to the study of how protein synthesis, trafficking, and degradation are coordinated and regulated in biological systems.

Gordon Rule, Ph.D., professor in the department of biological sciences at Carnegie Mellon, is applying biophysical techniques to study protein dynamics and protein-ligand interactions.

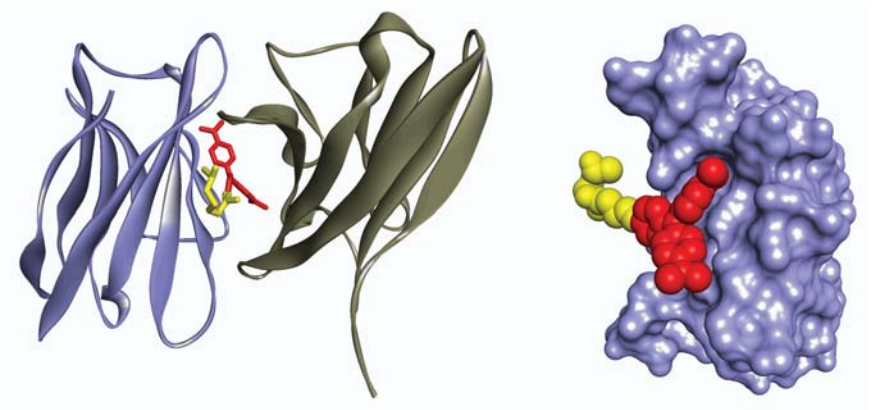
Brigitte F. Schmidt, Ph.D., senior research chemist for MBIC, designs biological probes endowed with specific properties. She is a member of the core group that developed FAB-binding fluorogenic dyes.



Simon C. Watkins, Ph.D., is a professor of cell biology and physiology, and immunology at the University of Pittsburgh and the founder of the Center for Biologic Imaging at that university. He is elucidating the mechanisms of communication between cells of the immune system using molecular and optical imaging tools.

Figure 2.

The ribbon molecule at left depicts a FAP bound to a malachite green fluorogen (red) coupled to a polymer linker (yellow). In this diagram, the biosensor module is sandwiched between two antibody variable domains. The space-filling model on the right shows a cut-away view of the biosensor and a single antibody domain.



Kimberly J. Zanotti, Gloria L. Silva, Yehuda Creeger, Kelly L. Robertson, Alan S. Waggoner, Peter B. Berget and Bruce A. Armitage. Blue fluorescent dye-protein complexes based on fluorogenic cyanine dyes and single chain antibody fragments. (2011) Org. Biomol. Chem., 9, 1012.

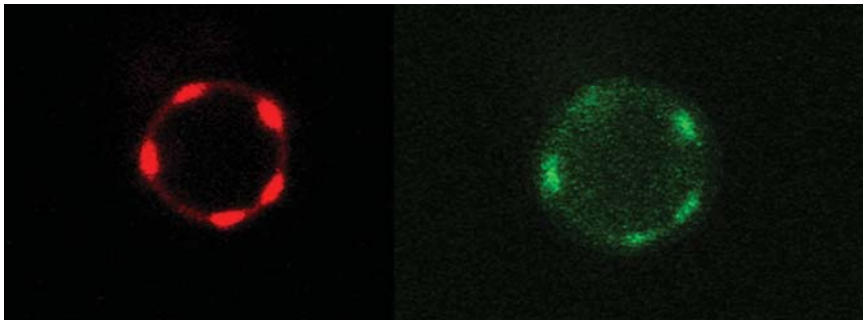
Szent-Gyorgyi C, Schmidt BF, Creeger Y, et al. Fluorogen-activating single-chain antibodies for imaging cell surface proteins. Nat Biotechnol. 2008;26:235–40.

Researchers can use FAPs to see where the target protein is within the space of the cell and observe color changes when it becomes fluorescent (see Figure 3). Color changes reflect changes in the local environment of the protein, thereby allowing immediate measurement of the biological activity of proteins and biomolecules that are near each other. The FAP's fluorescence can be turned on and off by adding or removing the fluorogen, a characteristic that makes the fluoromodules more useful than other fluorescent proteins.

Variable Domains:
Region of an antibody molecule that is most important in antigen binding

Figure 3.

Here, it can be seen how two different FAPs bind to a protein (Aga2p) that occurs in yeast bud scars. Both types of FAPS are displayed simultaneously on the yeast cell as they bind to both fluorescein (green) and modified malachite green (red).



Bud:
Asexual reproduction in yeast characterized by asymmetrical division that results in a daughter cell that is much smaller than the original cell.

Visualizing drug targets

MBIC researchers adapted the FAP technology to create a new biosensor that could aid in the development of drugs that target proteins called G protein-coupled receptors (GPCRs), which are involved in regulating many functions critical to health, including heart and lung function, mood, cognition and memory, digestion and the inflammatory response. Drugs that target GPCRs make up approximately 30% of pharmaceuticals on the market.

GPCRs are located in the cell membrane, where they interact with molecules responsible for cellular communication such as neurotransmitters and hormones. When one of the receptors encounters such a molecule, it relays a signal across the cell membrane that, in turn, initiates a response. After the response is triggered, the receptor retreats from the membrane into the cell's interior.

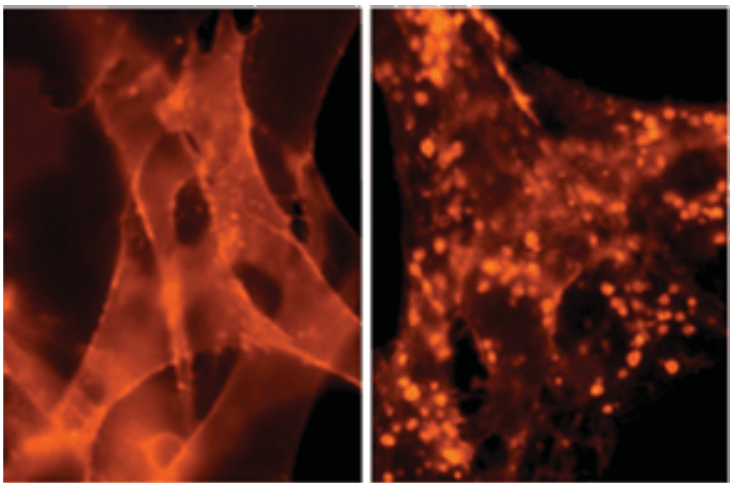
The MBIC team engineered GPCR fluoromodules consisting of a FAP that fuses to the beta-2 andrenergic receptor (β2AR), a GPCR that is present in brain, heart, lung, and other tissues. When the researchers introduce the fluorogen, it binds to the FAP-tagged β2AR on the cell surface, emitting a bright fluorescent glow (see figure). When the receptor is activated and retreats from the membrane into the interior of the cell, the fluorescence dims. In the future, this promising technology could be scaled up to screen large numbers of molecules for GPCR activity.

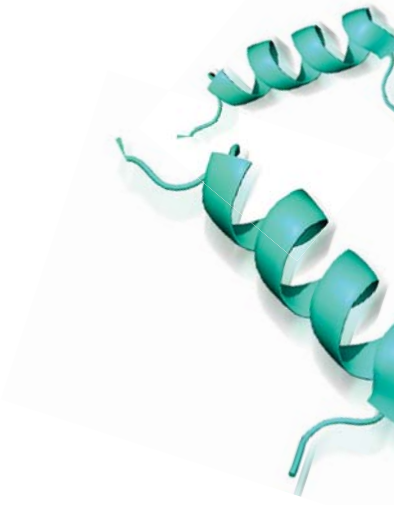
Holleran J, Brown D, Fuhrman MH, et al. Fluorogen-activating proteins as biosensors of cell-surface proteins in living cells. Cytometry A. 2010;77:776–82.

Fisher GW, Adler SA, Fuhrman MH, et al. Detection and quantification of beta2AR internalization in living cells using FAP-based biosensor technology. J Biomol Screen. 2010; 15(6):703–9.

Figure 3.

Pictured in the left panel, a group of FAP-labeled cells glow orange in the presence of a fluorogen dye. After the cells are treated with a drug that activates GPCRs, the receptors move inside the cell, leaving many fewer on the surface, as shown in the image on the right.

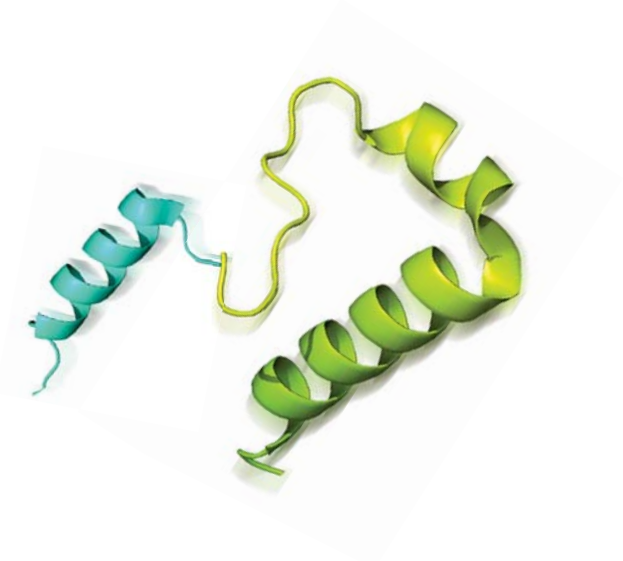




Future Directions

In the next few years, the Molecular Biosensor and Imaging Center hopes to have commercial entities manufacturing and distributing the receptor-tracking probes they have already developed, their protease biosensors, and their first targeted physiology probes. They are expecting to complete development of a new class of protein-protein interaction probes and create targeted indicator probes for calcium, potassium, membrane potential, and redox potential. Some of these new probes will take MBIC research into the important new area of autophagy for which they have already established several new collaborations. They especially want to continue to see these new probes become part of high-throughput screening efforts in drug discovery programs especially as manufacturers extend their generic biosensor technologies to a wide range of regulatory protein activities.

Autophagy:
A normal process in which a cell destroys proteins and other substances in its cytoplasm (the fliud inside the cell membrane but outside the nucleus), which may lead to cell death (apoptosis).



Available Resources

Workshops & Courses

A comprehensive list of MBIC resources can be accessed through the center's Resources webpage.

MBIC maintains an online booking system for its shared equipment and resources. Instructions for accessing the booking system can be found on the center's website. In addition, MBIC users have access to the imaging resources at the Center for Biological Imaging at the University of Pittsburgh. To access the CBI scheduling system for using the fluorescence, confocal and electron microscopy resources, go to their website.

Materials and Protocols

Among the resources available to the scientific community through MBIC are datasheets for the various fluorogenic dyes; vector map datasheets; primary scFv sequences; annotated scFv sequences; and protocols for plasmid production and archiving, preparation of fluorogens and dye stocks, and protein production and purification.

Software

Also available to MBIC personnel is the Imaris scientific software module, which allows visualization of original and derived data objects in a real-time interactive manner. Researchers can quickly make visual assignments of their experiments in 3D and 4D to discsover relationships that are otherwise hidden. Its rendering quality, speed, precision, and interactivity are unrivalled.

Image Processing & Analysis

A variety of technologies are available for visualizing and measuring experimental results, including a fluorescence-activated cell sorting flow cytometer, a Carl Zeiss Axio Imaging 2 ApoTome fluorescence microscope, an Olympus IX70 Laser TIRF system with EM-CCD camera, and a Tecan Safire2 fluorescence plate reader -- a fully modular monochromator-based microplate detection system that offers a range of high-speed fluorescence techniques. The plate reader is capable of measuring top and bottom fluorescence intensity, absorbance, and luminescence.

Available Resources Continued

Databases

MBIC has set up and maintains several databases that are useful to researchers using fluorescent microscopy tools to interrogate cellular networks and pathways:

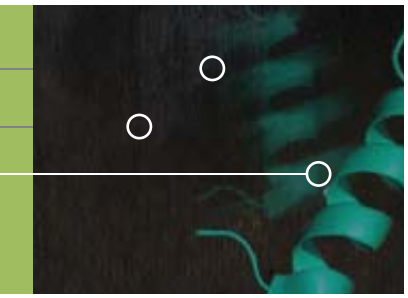
- MBIC Biosensor Database: The newly constructed MBIC Biosensor Database is now in the beta-testing phase. Access is by authorized user login only.
- Subcellular Location Image Finder (SLIF): SLIF locates fluorescence microscope images in online journal articles and indexes them according to cell line, proteins visualized, and resolution.
- Protein Subcellular Location Image Database (PSLID): PSLID provides an open-access, structured, well-annotated databse of fluorescence microscope images depicting subcellular location proteins.

Training Resources

MBIC facilities and equipment are available to center personnel and collaborators. Training is provided on the proper care and operation of the following systems:

- Confocal microscope systems (i.e., Carl Zeiss LSM 510Meta MP Confocor3 inverted spectral microscope, Carl Zeiss LSM 510Meta DuoScan inverted confocal microscope; Andor Revolution XD System with spinning disk confocal/TIRF/FRAPPA).
- Carl Zeiss Axio Imaging 2 ApoTome fluorescence microscope.
- Imaris scientific software module.
- FACS sorting flow cytometer with BD FACSDiva workstation.
- Olympus IX70 Laser TIRF system with EM-CCD camera.
- Tecan Safire2 fluorescence plate reader.
- Biosafety level 1 transfer cabinets.

The names and contact information for staff who provide the training are available via MBIC's website.



Networks & Pathways of Lysine Modification

Johns Hopkins University
Principal Investigator: Jef Boeke
www.hopkinsmedicine.org/ibbs/research/TCNP/index.html

Introduction

The amino acid lysine is emerging as important in a wide range of biological processes, from control of gene expression to recycling of proteins. The goal of the Networks & Pathways of Lysine Modification Center, a National Technology Centers for Networks and Pathways project, is to unravel the signaling networks and pathways in which lysine is involved by developing and applying genetic and computational approaches, proteomics and mass spectrometry technologies, single-cell profiling, and other novel methods to detect, quantify, and monitor lysine modifications.

Why lysine?

Lysine is unique among the amino acids because it can be modified in many ways, with profound biological effects. One critical question in biology is how different types of cells, for example muscle cells and brain cells, do such different things, even though all cells in a person or other organism have exactly the same set of genes. The general answer is that each different kind of cell manufactures different proteins, despite all having the DNA template to manufacture every protein the body can make. That statement only pushes the question further back, however. Scientists want to know how this happens, not simply that it does.

Lysine modification turns out to be an important part of the answer. For example, protein modification of lysines found on DNA-coiling histones is critical for epigenetic control of gene expression, the process that dictates which proteins are produced in which cells.

Other lysine modifications include:

- Acetylation—adding an acetyl group. Lysine acetylation preferentially targets large macromolecular complexes involved in diverse cellular processes, including chromatin remodeling and the cell cycle.
- Methylation—adding a methyl group.

Epigenetic Control: Through this mechanism, the underlying DNA is not changed, yet the control system is inherited, at least over the short term. That's how the lineage remains the same as cells reproduce—skin cells don't become heart cells.

- Ubiquitylation—adding ubiquitin, a small regulatory protein found in all eukaryotic cells. This versatile control mechanism is involved in many aspects of cell life. As one example, attaching a molecule of ubiquitin is the first step in a protein's degradation in the proteasome.
- SUMOylation—attachment of a SUMO protein (Small Ubiquitin-like Modifier). SUMO modification of proteins has many functions, including protein stability, transport into and out of the cell nucleus, and regulation of transcription.

Among these types of lysine modification, ubiquitylation is best understood; it controls proteins' life, death, and trafficking, among other roles. In general, however, these modifications and the pathways regulating their interactions are poorly understood. Studying them is no small task—many lysine-modifying proteins have multiple substrates and are encoded by a number of gene families. The Center is working to develop and use a number of specific technologies to understand the pathways in which lysine modification is involved.

How Will the Goal Be Met?

The Lysine Modification Center integrates "driving biological projects" centered on lysine acetylation, methylation, and ubiquitylation with technology development, training, and technology dissemination efforts. The key questions involved are related to defining the pathways and understanding their structures, i.e., how protein a connects to protein b, to protein c, and so on. To get at these questions, the Center has developed three key analytical technologies:

- SLAM—"Synthetic Lethality Analyzed by Microarray." Instead of merely determining if protein a interacts with or binds to protein b, SLAM allows the knockout of two activities at once to test the hypothesis that they represent parallel or redundant pathways. If so, knocking them both out should be lethal to the organism (usually yeasts in the experiments).
- Protein microarrays. Once lethality has been established (i.e., the pathways are important), the next step is to determine the proteins in the pathways. Protein microarrays step in as a complementary technology for that purpose.
- Innovative mass spectrometry technologies. In this setting, the purpose of mass spectrometry is to detect and quantify lysine modifications in the identified pathway. For example, an experiment might involve purifying a particular histone protein and looking at the lysine modifications that occur under different biological conditions.

Such currently experimental approaches may permit an understanding of the dynamic processes in which lysine modifications are involved. This could lead not only to solving some of biology's most puzzling riddles, but also to better understanding of disease and better treatments. For example, the epigenetic changes that occur in individual cells appear to be vitally important in understanding how cancer gets started and progresses.

Because lysine modification is involved with human health, aging, and disease at many levels, what is learned will have far-reaching implications.

Highlights

Weight Control by Inhibiting Ghrelin Activity?

Ghrelin is a peptide hormone that promotes weight gain in vertebrates. To be functional, however, it must be activated by adding an octanoic acid to a specific serine residue. This is done with the help of an enzyme called ghrelin O-acyltransferase (GOAT). Lysine Center investigators designed and synthesized a molecule that inhibits the GOAT enzyme, as a way to alter ghrelin's activity. They then tested this inhibitor, called GO-CoA-Tat, in both cell cultures and in vivo, injecting it into both wild-type mice, fully capable of producing the GOAT enzyme, and mutant mice unable to produce

Investigators

Jef D. Boeke, Ph.D., works to develop improved methods to identify genetic interactions (synthetic lethality) using microarrays (SLAM) and to build and interpret cellular "wiring diagrams," probe networks and pathways of lysine modifications, and to synthesize an engineered version of the yeast genome called Sc 2.0.

Philip A. Cole, M.D., Ph.D., focuses on chemical activity probes for identification of novel acetyltransferases, protein kinases, histone-modifying enzymes, and analyzing hormone biosynthesis. This technology involves the use of chemical affinity-labeling reagents to discover and characterize novel histone acetyltransferase (HAT) enzymes.

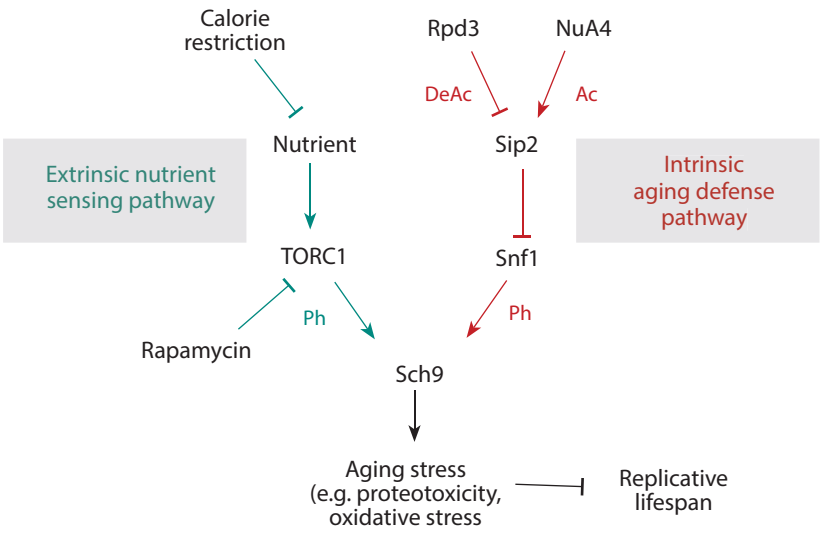
Robert Cotter, Ph.D., seeks to develop new mass spectrometric instrumentation and analytical methodology to improve capabilities for elucidating lysine modifications, particularly acetylation and ubiquitylation. He is particularly interested in developing miniaturized mass spectrometers for use as portable detectors of biological agents, environmental contaminants, and clinical diagnostics.

Yingming Zhao, Ph.D., develops mass spectrometry–based proteomics technologies to dissect post-translational modification pathways. He uses proteomics to characterize the functions of lysine acetylation, propionylation, butyrylation, crotonylation, succinylation, and malongylation pathways; the last five of these modifications were recently discovered in his lab.

Heng Zhu, Ph.D., applies protein chip technology to profile the dynamics of various post-translational modifications in humans. His lab is currently focused on (1) identification of SUMO E3 ligase–dependent substrates and their binding partners; (2) characterization of ubiquitin E2/3 relationships; (3) profiling phosphorylation changes using lysates from cultured cells or tissues; (4) trying to discover novel enzymatic activity; and (5) cancer biomarker identification and validation.

Figure 1.

Model of the Effects of Sip2 Acetylation on Control of Intrinsic Aging



Red lines indicate the intrinsic aging defense pathway identified in this study. Blue lines indicate the extrinsic nutrient-sensing pathway. Ac, lysine acetylation; DeAc, lysine deacetylation; Ph, phosphorylation of serine and threonine.

Lu J-Y, Lin Y-Y, Sheu J-C, et al. Acetylation of yeast AMPK controls intrinsic aging independently of calorie restriction. Cell, 2011;146:969–979.

GOAT. They found that GO-CoA-Tat improved glucose tolerance and reduced weight gain in the wild-type mice but not in the mutant mice. This is important evidence that the beneficial metabolic effects were due specifically to inhibiting GOAT. This work suggests that GOAT might be a new clinical target for the treatment of human metabolic disease.

Barnett BP, Hwang Y, Taylor MS, et al. Glucose and weight control in mice with a designed ghrelin O-acyltransferase inhibitor. Science, 2010;330:1689–1692. PubMed Link

An Aging Clock in Yeast

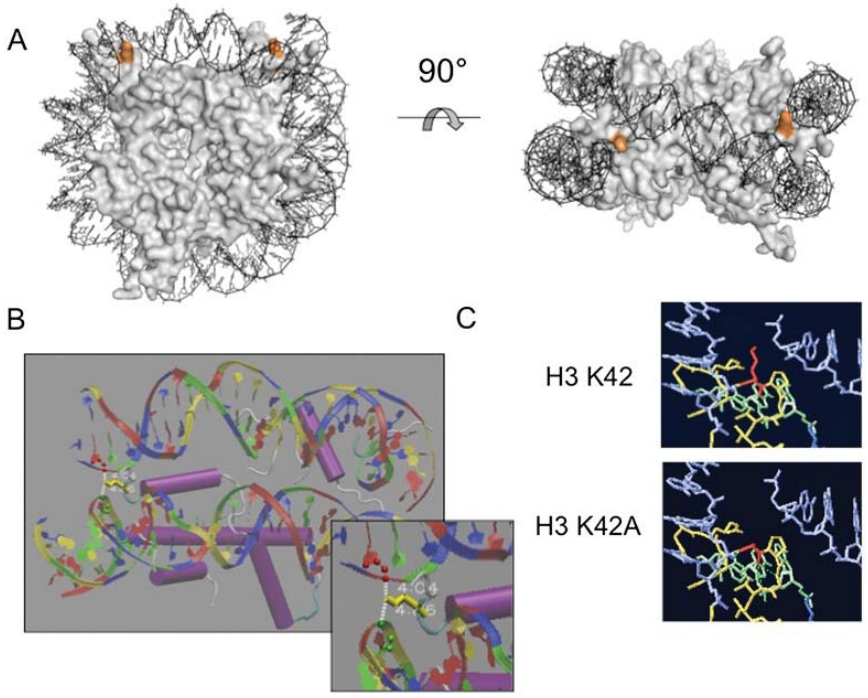
In humans and other higher organisms, a protein called AMPK (5′ AMP-activated kinase) serves as a sort of metabolic master regulating switch. It increases cellular energy when more is needed by activating catabolic processes (breaking down such molecules as proteins and lipids), while shutting down energy-using anabolic processes (synthesis of new proteins, lipids, etc.). The important questions are how does this happen and what are the results. Lysine Center investigators have recently demonstrated that the counterpart of AMPK in the yeast Saccharomyces cerevisiae, called Snf1, controls intrinsic aging and extends replicative life span through progressive deacetylation of the Snf1 subunit Sip2. The team identified 4 lysines that progressively showed less acetylation as cell lineages aged. To test whether acetylation of these lysine residues was the controlling mechanism, they created a Sip2 lysine-to-arginine mutant that could not be acetylated. These mutants were very short lived. Another mutant, lysine-to-glutamine, which mimics the normal acetylated Sip2, reversed the shortened life span. Careful sorting of yeast cells by age, then testing for acetylation levels, confirmed the progressive deacetylation with age. Because AMPK is highly conserved among eukaryotes, these results suggest that a similar intrinsic aging control system occurs in higher organisms, raising the interesting possibility of pharmacologic anti-aging tools.

A Lysine Transcriptional Gatekeeper in Yeast

Lysine Center investigators believe they have found a transcriptional ‘‘gatekeeper’’in the αN-helix region of histone H3, where Lys 42

Figure 2.

Histone H3-K42A is a mutation that disrupts the nucleosome surface

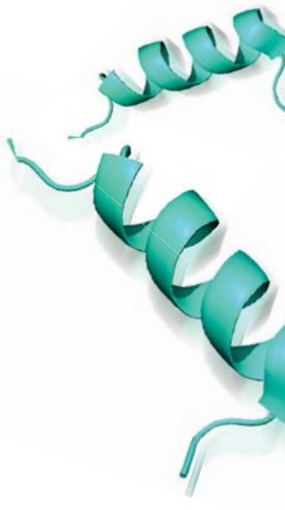


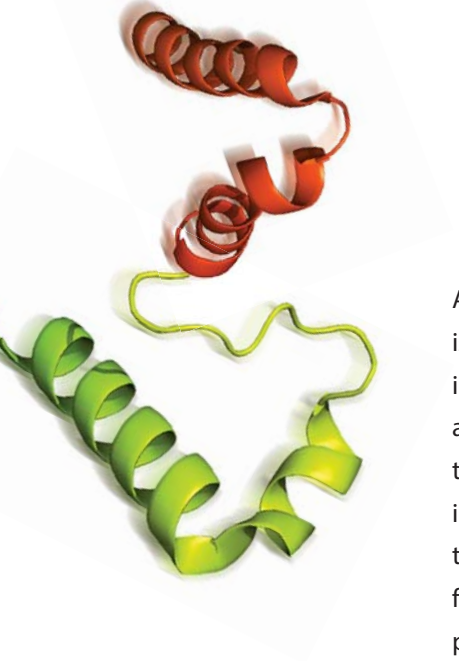
(A) The crystal structure of the S. cerevisiae nucleosome, highlighting histone H3 Lys 42 (orange) positioned at the DNA entry and exit points. (B) Nucleosome crystal structure surrounding H3-K42 (yellow), indicating hydrogen bonds between the lysine side chain and DNA. (C) Predicted structures of the H3-K42A mutation in its nucleosomal context. (Red) Residue 42; (blue) DNA; (green/yellow) histone H3.

Hyland EM, Molina H, Poorey K, et al. An evolutionarily 'young' lysine residue in histone H3 attenuates transcriptional output in Saccharomyces cerevisiae. Genes and Development, 2011;25:1306–1319.

interacts directly with the DNA (Lys is the 3-letter code for lysine, and Lys 42 is amino acid number 42 from the N-terminal end of the protein). Furthermore, methylation of this lysine is the likely mechanism by which the cell regulates this gatekeeping function.

Histones are proteins found in the nucleus of eukaryotic cells. They serve to organize DNA physically and are involved in gene regulation. This is a story of H3, one of the "core" histones, so called because they combine with 3 other types to form the nucleosome core, a kind of spindle around which the DNA wraps. Each nucleosome holds 147 base pairs (bp) of DNA; these nucleosomes are separated by up to about 80 bp of DNA, giving the appearance of beads on a string.



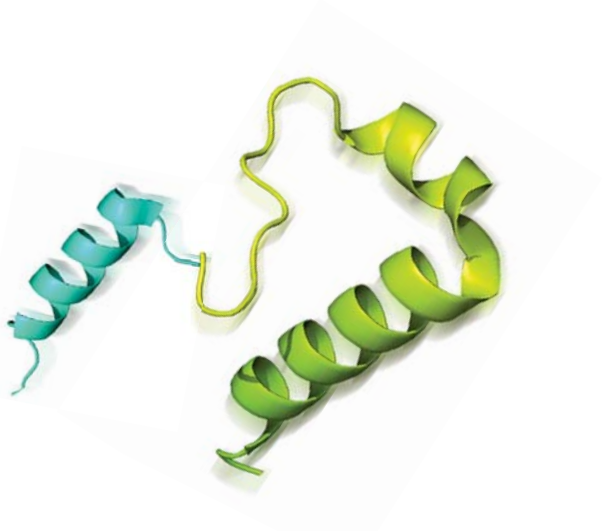


An aggregation of these proteins with DNA is called chromatin. When carefully arranged in a 3-dimensional construct that preserves access to the DNA for transcription and replication, chromatin forms chromosomes. Histones interact directly with DNA and histone mutations that disrupt these interactions are often fatal, underlining their importance to biological processes.

This study focused on H3 Lys 42 because it interacts directly with the DNA and is a site for methylation. A non-lethal H3 mutation exists (K42A) in Saccharomyces cerevisiae, in which the lysine (K) is converted to alanine (A). Cells with this mutation grow at about half the normal rate and are overly sensitive to a number of chemicals, indicating that this specific lysine is critical to numerous cellular processes or that its effect on nucleosome structure indirectly influences multiple pathways, possibly by altering gene expression.

To examine this question, the Lysine Center investigators looked at the gene expression profile in this mutant and found that the expression of about 42% of the genes was altered, most (90%) increased, on average about 2-fold, but up to 12-fold over normal wild-type yeast cells. Other amino acid substitutions in the same region of H3 also increased gene expression, but not uniformly, and K42A was the most severe.

These results indicate that Lys 42 is involved in controlling gene expression and altering it, as in the K42A mutation (which means the protein cannot be methylated at this position) disrupts gene control.



Available Resources

Workshops & Courses

See a list of current and past courses, workshops, and meetings on the Networks & Pathways of Lysine Modification website.

Protocols

The Lysine Modification Center currently has more than a dozen protocols available on their website.

Databases

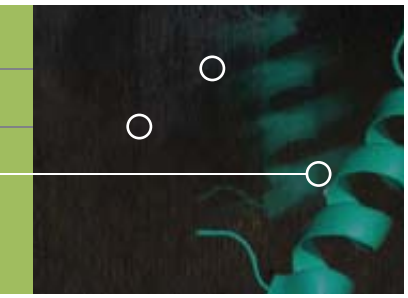
Currently available databases include:

- Histone Systematic Mutation Database, available at HistoneHits.
- Human Protein Reference Database, available at HPRD.
- Human Proteinpedia, a community portal for sharing & integration of human protein data, available at Human Proteinpedia.

Other Resources

- Synthetic Lethality Analyzed by Microarray (SLAM). Inquire via Jef Boeke or Pamela Meluh.
- A frozen strain and plasmid distribution system called IcePick.
- A web-based program called GeneDesign for the design of synthetic genes.
- Middle Atlantic Mass Spectrometry Laboratory.
- Stable Isotope labeling with amino acids in cell culture, available at SILAC.
- The Yeast Histone Identification Factory, available at YeHIF.
- A full calendar of Center activities, seminars, etc., is available on Networks & Pathways of Lysine Modification website.

http://www.hopkinsmedicine.org/institute_basic_biomedical_sciences/research_centers/high_throughput_biology_hit/technology_center_networks_pathways/#



National Center for Dynamic Interactome Research

Rockefeller University
Principal Investigator: Michael Rout
www.ncdir.org

Introduction

Genomes of entire organisms can now be sequenced in a matter of days. The National Center for Biotechnology Information database has millions of gene entries, and genes are just the start. They encode a comparable number of proteins, with all their homologs and variants. If "Life" is a Broadway play, the genome is its book and proteins the actors. And just like a play, life is never static. Proteins are interacting like actors on a stage—or perhaps more like an old-time Hollywood spectacle, with "a cast of thousands" and an almost unimaginably intricate plot.

What living cells do cannot be captured in a static picture or blueprint. Instead, we must understand the many dynamic interactions of the players, as defined by that blueprint. Proteins are highly variable in both structure and function. They are made up of molecular building blocks called amino acids, and a given protein might contain tens or hundreds of these, strung together in a specific order. Then they twist and fold and curl back on themselves; often proteins combine with others to become functional. Sometimes they act alone, sometimes they team up, but always they are built to carry out the cell's work.

Proteome:
The total of all proteins that can be produced by a given organism, tissue, or cell

Interactome:
All interactions between proteins and between proteins and other types of organic and inorganic molecules within a given organism, tissue, or cell, in a specified process or set of conditions

Tools for Dissecting the Macromolecular Machinery of Dynamic Cellular Processes

Even though scientists have today identified thousands of protein-protein interactions, using current technologies at today's pace it would still take hundreds of years to know and understand all the functions and dynamic interactions carried out by all the proteins in the human proteome—the sum total of all the proteins humans are capable of producing. This total picture of dynamic interactions is called the interactome. With so many interactions remaining to be uncovered, new technologies are desperately needed that can reveal the interactome, illuminating the dynamic interplay of the macromolecular characters in cells.

The National Center for Dynamic Interactome Research (NCDIR) is creating techniques that can follow this interplay at the timescale and spatial resolution needed to "see" these dynamic cellular processes. With these new technologies, structures can be determined and interactions uncovered hundreds or even thousands of times faster than using the present methods.

NCDIR's first goal is thus to create robust, innovative new tools for isolating, analyzing, and visualizing macromolecular complexes to reveal their dynamic behaviors, and to test these tools on biological systems that typify many of the major problems encountered in proteomic investigations of cellular processes. Its second goal is to place these proven tools in the hands of biomedical researchers, through publications and seminars, through Web-based and hands-on training, and through new reagents and strains, so investigators can realize the full potential of the genomic revolution.

Technologies for Solving Today's Scientific Questions

Organelle:
A functionally specialized part of a cell, such as a ribosome or a mitochondrion, that is analogous to an organ in an organism

Stoichiometric Quantification:
The number of molecules of chemical a needed to completely react with chemical b in a specified reaction c

NCDIR's technology development is driven by problems in biology and involves the synergy of several groups working together to leverage each other's strengths. For example, organelle biogenesis: how does a cell build a nuclear pore complex? This question is far from simple. When unpacked, it's really many questions. What are the steps? Which proteins interact in each of the steps? What do the intermediate complexes look like? To answer such questions, it must be possible to observe the process with great sensitivity.

Technologies that make this possible include:

- Fluorescent and chemical tags on the proteins in a complex, and affinity molecules that bind to the tags, allowing the complex to be visualized and purified.
- A protocol for rapidly freezing cells and keeping them at liquid nitrogen temperatures while breaking them open, so all interactions ongoing at the time of freezing are preserved.
- Tools for high-resolution spatial analysis to "see" the proximity and spatial relationships of individual proteins in the emerging complex.
- Mass spectrometry tools that enable the identification of the components in a complex; the stoichiometric quantification of interacting proteins (the number of molecules of protein x that bind to protein y, where x and y are in the same complex); the chemical modification of these proteins (such as phosphorylation),

serving as signals to initiate other events; and quantitative analysis of the changes in the type, number, and modifications of proteins in a complex over time.

- The Integrative Modeling Platform (IMP). For details, see Highlights.

The next step is to test these new technologies using model systems that not only represent important biological questions, but also epitomize the sorts of problems the technology is intended to solve. Imagine being able to tag and isolate all proteins interacting in a given process at very fine-grained time intervals, freeze the action at each time so the molecules can be identified, with no non-specific interactions to muddy the picture, then combine these high-quality data with spatial analyses and stoichiometric relationships at each step to generate a series of "pictures" of the developing complex. The result is not just a static model of a complex, but something more akin to a movie of the complex being built.

NCDIR Director Michael Rout likens the Center's research groups to the age-old analogy of blind men groping an elephant. The Center takes all the various data and integrates them into one coherent picture that satisfies the constraints those data impose. Instead of the "rope" found by one group or the "pillar" found by another, the Center's integrative multi-scale approach allows development of a complete picture of cellular systems in all their dynamic complexity.

Investigators

John Aitchison, Ph.D., studies functional genomics, proteomics, and genetics to understand the regulation of nuclear function and peroxisome development. Understanding of these processes provides insight into the fundamental biology nuclear function, and in particular peroxisome biogenesis and disease processes associated with peroxisomes.

Brian Chait, D. Phil., is the Camille and Henry Dreyfus Professor at Rockefeller University in New York and head of the Laboratory for Mass Spectrometry and Gaseous Ion Chemistry. He specializes in the use of mass spectrometry as a tool for investigating a variety of biological and biochemical phenomena.

Michael Rout, Ph.D., is NCDIR director. He uses biochemical, biophysical, and structural approaches to characterize macromolecular assemblies, with an emphasis on the nuclear pore complex, a key part of the pathway that relays information between the nucleus and cytoplasm. His goal is to further develop proteomic technologies that will enable the community to assemble detailed, dynamic representations of the interactions in the cell.

Andrej Sali, Ph.D., maximizes the accuracy, precision, completeness, and efficiency of the structural coverage of proteins and their assemblies through use of the Integrative Modeling Platform, which tightly integrates different types of experimental data, physical theories, and statistical inference, spanning all relevant sizes and time scales, to determine the structures of macromolecular complexes.

Highlights

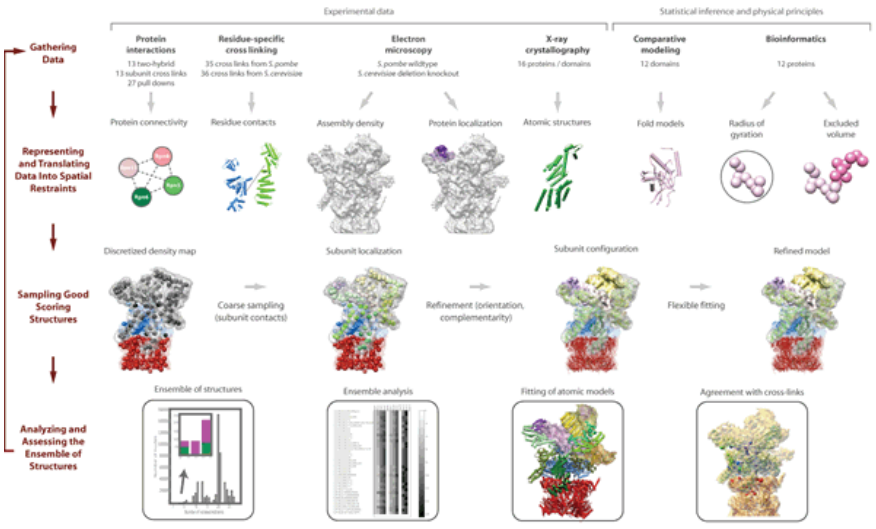
Putting the pieces together: Integrative Modeling Platform Software for structure determination of macromolecular assemblies

Ribonucleoprotein Complex: An assembly formed of ribonucleic acid (RNA) and proteins

Building explicit models of a biological system that are consistent with the myriad available data is one of the key challenges in biology. Modeling the structure and dynamics of macromolecular assemblies, for example, can give insights into how biological systems work, how they have evolved, and how they might be controlled. Modeling can also suggest future experiments. Figure 1 shows how the process of integrative structure modeling, which aims to use information from any source, improves the accuracy, precision, completeness, and efficiency of structure determination. Integrative modeling may be thought of as a computational optimization problem: information about the assembly is encoded into a scoring function that evaluates candidate models. In fact, integrative modeling replicates the standard scientific cycle of gathering data, proposing hypotheses, and then gathering more data to test and refine those hypotheses. It repeats the stages of gathering information, choosing how to represent and evaluate models, finding models that score well, and analyzing both the models and the information. The cycle ends when the models converge to fit all current information and are judged to be satisfactory. When new information is gathered, the cycle resumes. This approach enables assessment of both the output models and the input data, and offers the opportunity to plan the most useful future experiments. To facilitate widespread adoption, the Center is developing, applying, and distributing the open-source software package for this integrative modeling approach, called the Integrative Modeling Platform (IMP).

Figure 1.

Application of IMP to determination of the molecular architecture of 26S proteasome



Red lines indicate the intrinsic aging defense pathway identified in this study. Blue lines indicate the extrinsic nutrient-sensing pathway. Ac, lysine acetylation; DeAc, lysine deacetylation; Ph, phosphorylation of serine and threonine.

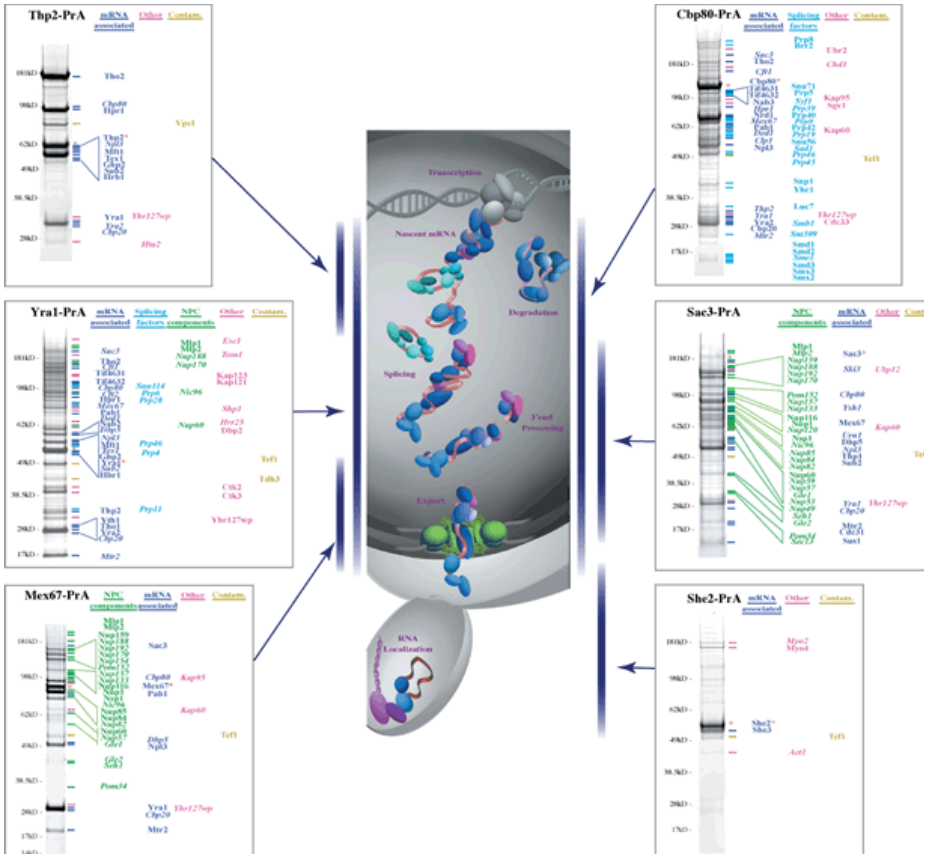
Lasker K, Förster F, Bohn S, et al. Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proceedings of the National Academy of Sciences, in press.

Proteomic and RNomic Analyses of Diverse RNP Complexes

Biology's central dogma has long been that genetic information in DNA is transcribed into RNA to produce proteins and cellular machinery. However, before protein production can begin, a variety of protein factors associate with the RNA, forming large ribonucleoprotein complexes (RNPs), to modify, splice, fold, and transport the RNA out of the nucleus. RNPs come in two forms:

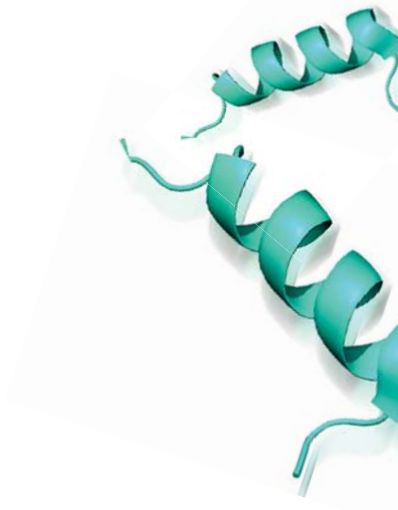
Figure 2.

Steps in mRNP development



Oeffinger M, Wei KE, Rogers R, et al. Comprehensive analysis of diverse ribonucleoprotein complexes. Nature Methods, 2007;4:951–956.

mRNPs contain messenger RNA (mRNA) and rRNPs contain ribosomal RNA (rRNA). Efforts to study the dynamic assembly and disassembly of RNPs have been hampered by RNA's extreme sensitivity and the short life of the complexes. Weakly interacting components can be lost and nonspecific material can associate. NCDIR investigators have developed a rapid, single-step affinity purification method to overcome these limitations and capture the dynamic interactome of maturing RNPs. Figure 2 illustrates how the method isolates each stage of RNP maturation and processing, from transcription at the top to export out of the nucleus at the bottom, so the material in each can be analyzed for its RNA and protein content. Detailed study of the mRNP maturation pathway is shining light on whether all mRNAs have a common maturation pathway or if different maturation processes are needed for different classes. The results will help to determine whether particular mRNAs associate more strongly with particular classes of RNP and provide a dynamic, high-resolution map of RNP assembly.

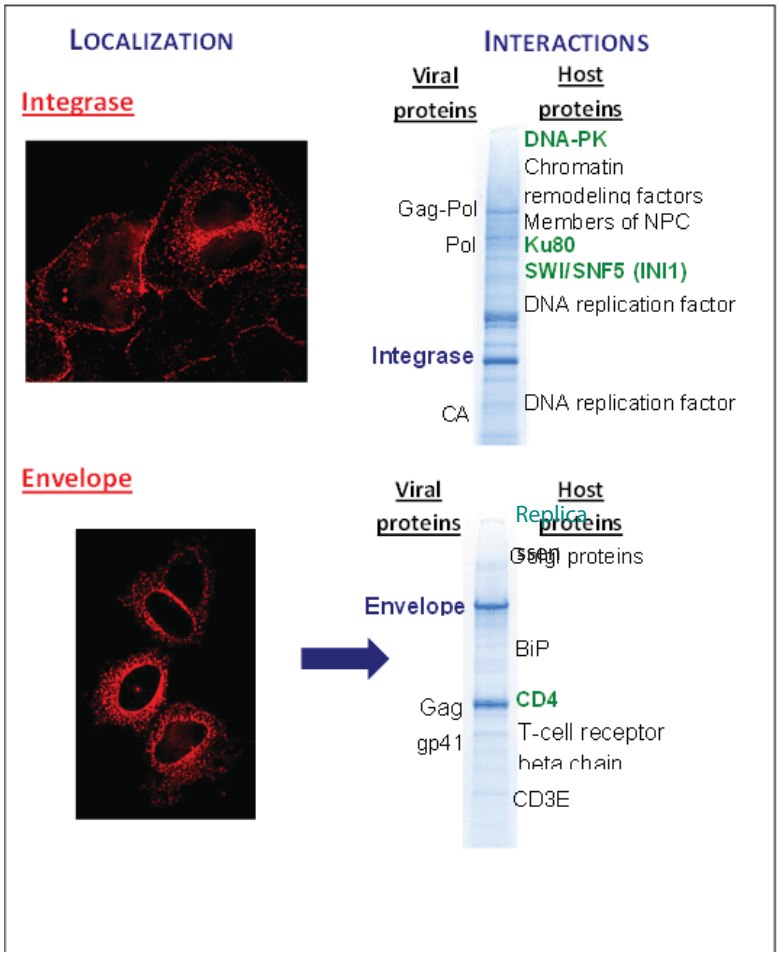


Replication Fork:
The essential structure that forms DNA is replicated. This fork starts at a "point of origin" and moves along the DNA strand, opening it in a Y-shape so replication can proceed behind it.

Mapping the Virus-Host Cell Interactome

When a virus—such as Dengue, influenza, HIV, or cytomegalovirus—invades a host cell, it must rely on many host cell factors to complete the infection because of limited viral genetic capacity. An NCDIR focus is on identifying these host factors, and determining why their interaction with viral proteins leads to the viral disease progressing. One example involves the devastating pandemic HIV-1, the causative agent of AIDS. NCDIR has developed a random insertional tagging and selection system in which viruses are engineered to incorporate an immunological or biochemical tag. These viruses can replicate, and the modifications are stably maintained in culture. Figure 3 illustrates the host protein–HIV-1 integrase and envelope protein interactions revealed by this work as these tagged viruseds move through their natural life cycle. A number of these were previously unknown. These findings may provide insight into the strategies HIV-1 uses to maneuver the cellular machinery. Ongoing experiments will investigate further how viruses use host factors. Synchronized infections allow recovery of protein complexes at different stages of infection. These methods provide a powerful tool to identify cellular proteins that associate with any viral protein during the viral life cycle. Some of these might be absolutely required by the virus but be at least partially nonessential to a host cell, thus offering potential new targets for therapeutic intervention.

Figure 3.
Localization of host protein interactions with viral proteins.



Chromosomal Interactomes and Dynamics

The control of DNA in our cells depends on particular protein-DNA associations. The resulting complexes and their chemical modifications determine which genes a cell should transcribe at what time and are also responsible for DNA replication. Determining what these proteins are and where on the DNA they bind is the key to understanding both transcription and replication and ultimately, to preventing and treating many diseases. NCDIR investigators studied the dynamics of DNA replication in yeast, looking at progression of the essential template for replication, called the replication fork.

Previous studies suggested that DNA replication

proceeds at variable rates over different parts of the budding yeast genome, strongly affected by local chromatin structure and by interaction with machineries controlling transcription, repair, and epigenetic maintenance. NCDIR investigators used microarray analysis to study an integral member of the replication fork called the GINS complex and found that it moves uniformly regardless of genomic location, leading to a fundamentally new understanding of DNA replication. Such dynamic studies may ultimately provide insights into disease processes caused by defects in replication.

Sekedat MD, Fenyö D, Rogers RS, Tackett AJ, Aitchison JD, Chait BT. GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. Molecular Systems Biology, 2010; 6:353.

GINS Complex:
A recently discovered, essential part of the replication fork. The GINS complex includes 4 different proteins and appears to be key to not only establishing the replication fork, but also in its progression.

Future Directions

Impact on Science

One major goal of the center is to be able to freeze any tagged macromolecular complex in place so that its position in the cell can be visualized and the intact complex isolated with all its components and specific neighbors. This will allow its macromolecular composition to be analyzed, its morphology visualized, and by comparisons of the complex at defined time points, its dynamic changes tracked and quantified. These data can then be integrated into a detailed representation of the changing interactions of the macromolecular players in almost any dynamic subcellular assembly. Like any good mechanic, we need a detailed understanding of how all the parts dynamically interact in these "molecular machines" if we are to understand how they function in a living cell—and thus how they can cause disease when they go wrong.

Impact on Health Care

Humans have around 20,500 protein-encoding genes, each potentially involved in one or more disease processes; for example, an estimated 200 proteins contribute to cancer. The number of the potential interaction states of so many proteins during the healthy or diseased life of a cell is enormous. This means we need robust, easily transferable methodologies to characterize systems so that groups targeting a particular disease can readily incorporate them into their strategies. NCDIR's goals are the creation, application, and dissemination of such methodologies.

Available Resources

NCDIR Director Michael Rout notes that communicating research results is just as important as doing the research. When the point of your research is to develop tools, other investigators who would find those tools useful must not only hear about them but must also learn to use them. "As has been pointed out, science is really a blue collar enterprise—it's very hands on. Interested investigators have to get their hands dirty learning how to use these tools if they are to do any good," Rout says.

Workshops & Courses

NCDIR currently has a number of workshops and course offerings in systems biology, including an introduction to systems biology and a proteomics workshop. For current offerings and registration, go to www.ncdir.org/workshop.php.

Protocols

Current offerings number about two dozen, including several that are available online in video form. For details visit www.ncdir.org/protocols.php.

Databases & Software

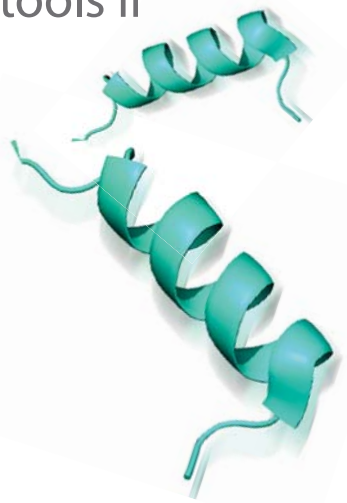
NCDIR has a number of databases that unregistered users may browse and open source software that any one can download. See http://mylabdatabase.com/cgi-bin/rout_labdatabase/main.fpl and http://salilab.org/imp for more information.

Summary Statistics

The center has produced 155 publications as of 2011. A total of 195 investigators use the facilities, from 26 states and 21 countries. Ten institutes at NIH fund the NCDIR's research.

"As has been pointed out, science is really a blue collar enterprise -- it's very hands on. Interested investigators have to get their hands dirty learning how to use these tools if they are to do any good."

-Michael Rout



R.D. Berlin Center for Cell Analysis and Modeling

University of Connecticut Health Center
Principal Investigator: Les Loew
www.ccam.uchc.edu/index.html

Introduction

Networks and pathways consist of series of individual reactions or processes, each of which is inherently reversible. But, under conditions in living cells, a pathway often proceeds in one preferred direction because downstream or upstream molecular interactions among pathway components confer polarity, or directionality, on the overall pathway. The interactions that determine the route by which a signal traverses a branched network can also be biased by how molecules are organized within the cellular space.

The R.D. Berlin Center for Cell Analysis and Modeling (CCAM) at the University of Connecticut Health Center is tackling tough but important questions about how molecules in cells move and how their arrangements in relation to other cell components influence cell function. Established in 1994, CCAM brings together faculty, instrumentation, and research dedicated to a quantitative multidisciplinary approach to the study of cellular systems.

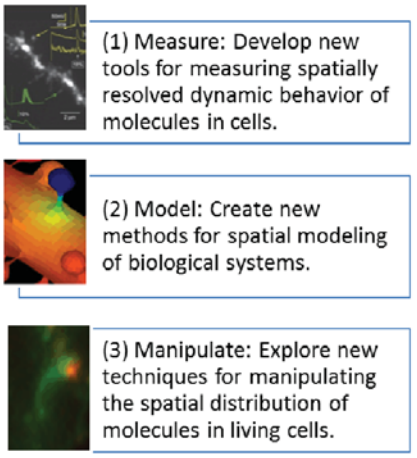
New Technologies to Measure, Model, Manipulate

CCAM hosts a confluence of expertise in physics, chemistry, software engineering, and experimental cell biology immersed in a biomedical research setting that values interdisciplinary collaborations. Center researchers have integrated new microscope technologies capable of making measurements in living cells with new physical formulations and computational tools. The result is spatially realistic quantitative models and experimental manipulations that are providing new insights into the changing distribution, behavior, interactions, and functions of specific molecules within the cell. The information gleaned by applying these powerful new tools are pulling back the curtain to reveal how signaling networks regulate and are regulated by the spatial organization of molecules in cells.

Investigating the influence of spatial organization of molecules on networks and pathways in living cells, the researchers use a three-way approach as shown in Figure 1.

Figure 1.

Three-way approach to investigating the influence of spatial organization of molecules on networks and pathways in living cells



Technologies for Measuring

- Fluorescence correlation spectroscopy (FCS): This technology can provide quantitative data for analyzing molecular interaction networks underlying complex cell biological processes by measuring concentrations, diffusion coefficients, kinetic constants, and conditional operators in living cells. This high-resolution method of imaging the movement of fluorescent molecules in and out of a volume has been used to understand the complex regulation of RNA nuclear–cytoplasmic transport and RNA granule trafficking in neurons. In addition carrying out basic research, CCAM investigators are creating protocols that will make FCS easier for the scientific community to use.
- Optical probes: Fluorescent dyes are key tools in the study of cellular processes. At CCAM, novel chromophores from environmentally sensitive dyes are synthesized and characterized. Membrane-associated styryl dyes are a particular interest in the study of transmembrane potentials and signaling events that occur due to voltage changes at the plasma membrane. Combining these dyes with additionally innovative methods of optics creates new quantitative ways of looking at cellular behaviors.
- Second Harmonic Generation: This physical phenomenon occurs when an intense laser beam interacts with a polarizable

substance that is without a centered symmetry. This nonlinear optical process has been optimized to examine biological tissues and is being adapted to advance the resolution and characterization of tissue structure.

- Single-molecule imaging: This technology allows cell biologists to analyze cellular behaviors at a new scale of resolution. Faculty at CCAM have used photo-switchable fluorescent proteins for single-molecule tracking in live cells to analyze the dynamics of membrane-cytoskeletal interactions as well as properties of gene transcription and translation. Single-molecule tracking involves the exploitation of chromophores for in vivo imaging, millisecond and nanometer resolution from two-photon laser microscopy, deconvolution, and tracking algorithms. Single-molecule imaging, similar to FCS, is providing exceptionally fine resolution to cellular behaviors such that stochastic behaviors can be quantitatively measured.

Technologies for Modeling

- BioNetGen: This software tool automatically generates a biochemical reaction network from user-specified rules for biomolecular interactions on the level of protein domains. Rule-based models allow the biologist and modeler to generate complex networks that emerge from the logical consequences of knowledge and assumptions about the mechanistic details of cellular signaling (http://vcell.org/bionetgen/features.html).

Investigators

Leslie M. Loew, Ph.D., the director of CCAM, served as the TCNP principal investigator. He focused particularly on collaborative research, training cores, and the administrative and managerial aspects of the center. Dr. Loew has expertise in developing and characterizing fluorescent probes of membrane potential. He was also a major force behind the VIrtual Cell project for which he created a framework for using computer simulation to explore cell biological mechanisms.

As deputy director of CCAM, Ann Cowan, Ph.D., oversees the light microscopy user facility, which provides access to state-of-the-art instrumentation in high-resolution fluorescence microscopy to the general research community. For NRCAM, she administered the service and dissemination cores.

Other Investigators

Measurement Tools - John Carson, Ph.D., University of Connecticut

Analysis Tools - James Schaff, Ph.D., University of Connecticut

Computer Science Tools - Michael Blinov, Ph.D., University of Connecticut

Theoretical Tools - Charles Wolgemuth, Ph.D., University of Connecticut

- Gfit: This open-access software tool can perform global analysis of large, heterogeneous data sets. Global analysis performs simultaneous regression analysis for multiple parameters in the biological system. This tool and approach has been used successfully with FCS data and surface plasmon resonance to study nuclear transport, transcription, and DNA-protein interactions. http://gfit.sourceforge.net/

Technologies for Manipulating

- Cellular tissues and development: Cellular adhesion and fusion are critical mechanisms to the formation of tissues during development. Genetics and new optical methods are being used to study cell fusion and tissue formation in Caenorhabditis elegans (a soil nematode often used as a model organism) and tissue properties of mammalian ovaries. Live imaging combined with genetic and nano-fabrication manipulations are powerful tools for delineating the mechanical and cellular behaviors associated with normal tissue development and cancers.

- Cytoskeletal dynamics: Actin and microtubule dynamics play an important role in vesicular transport, e.g., melanophores in Xenopus melanocytes, and signaling transduction, nck signaling and macroscopic cellular behaviors such as cell motility, crawling, and shape. Cellular, molecular, and biochemical approaches are used to understand the mechanisms of self-organization of cytoplasmic microtubules in living cells, regulation of microtubule dynamics, and cellular control over the activity of microtubule-based molecular motors. Biophysical modeling, such as solid and fluid mechanics, coupled with cellular imaging are used to probe the physical and biochemical mechanisms by which cells move, grow, and maintain their shape.

Highlights

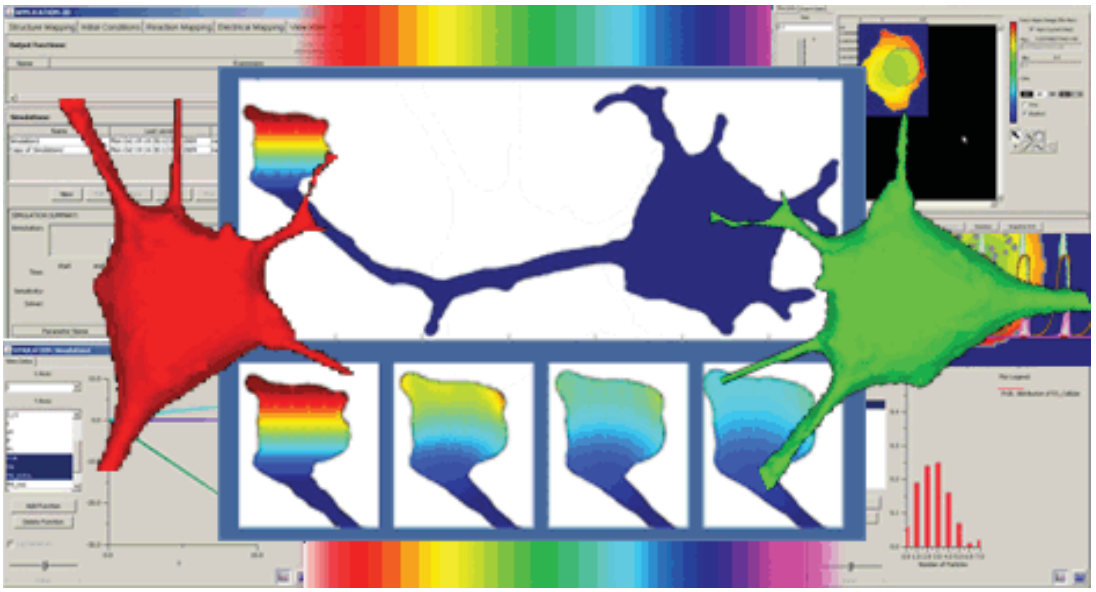
Virtual Cell

CCAM is the home of the Virtual Cell, a computational environment for cell biological modeling developed as an NIH-designated National Resource. By harnessing the power of computational science and systems biology, Virtual Cell has revolutionized cellular and molecular biology by creating realistic simulations of intracellular processes. The modeling software is useful for scientists ranging from cell biologists to theoretical biophysicists who can create simple or complex multi-layered models with its Java web-based interface. It is freely accessible to all members of the scientific community.

Virtual Cell automatically converts biological descriptions into a corresponding mathematical system of ordinary and/or partial differential equations. The models are built naturally from experimental images of cell and subcellular structures combined with biochemical and electrophysiological data. Results can be displayed

Figure 2.

Screenshot from Virtual Cell (http://vcell.org/images/vcell.png)



Moraru II, Schaff JC, Slepchenko BM, Blinov ML, Morgan F, Lakshminarayana A, Gao F, Li Y, Loew LM. Virtual Cell modelling and simulation software environment. IET Syst Biol. 2008 Sep;2(5):352-62. Available at: www.ncbi.nlm.nih.gov/pmc/articles/PMC2711391/.

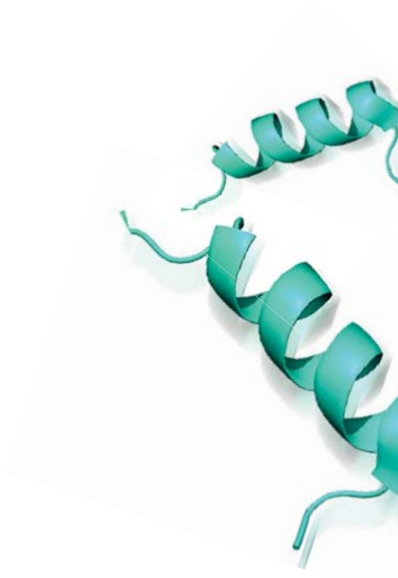
and analyzed online or downloaded to the user's computer in a variety of formats.

Because Virtual Cell is web-based, it has become a huge hit within the cell biology community. Biologists were able to create powerful simulations without having their own sophisticated computing hardware or high-end software expertise. It also allowed users to store data in the centralized database at the Health Center. It enables them to easily share models, and collaborate. Moreover, the Virtual Cell team can continuously update and improve the software without worrying about backwards compatibility and maintenance of many different versions of software.

RNA Trafficking

Research at CCAM has been at the forefront of utilizing advanced fluorescence imaging techniques to study RNA trafficking in cells. Research focuses on intracellular RNA trafficking in neural cells investigated through a combination of quantitative live cell imaging, fluorescence correlation spectroscopy, and mathematical modeling. Such studies have identified RNA granules—an RNA trafficking intermediate—and characterized spatial transcriptional and translational regulation of RNAs localized to granules.

Carson JH, Barbarese E. Systems analysis of RNA trafficking in neural cells. Biol Cell. 2005 Jan;97(1):51-62.



Available Resources

Microscopy Resources

Microscopy instrumentation includes four confocal laser scanning microscopes, including ultraviolet excitation and real-time imaging, nonlinear optical microscopy utilizing a titanium sapphire pulsed laser, confocal-based fluorescence correlation spectroscopy, wide-field imaging workstation with cooled charge-coupled device and rapid excitation filter wheel, and dual-wavelength spectrofluorometry.

Virtual FRAP Tool

The Virtual FRAP tool, a downloadable executable program, is designed to analyze data from FRAP (fluorescence recovery after photo-bleaching) experiments. It can be used to analyze experiments that collect all of the fluorescence associated with the cell.

Microfilament Detector

This tool is used to detect and classify cellular structures in microscopic images. The software is modular in design to support easy modifications and integration with other software applications.

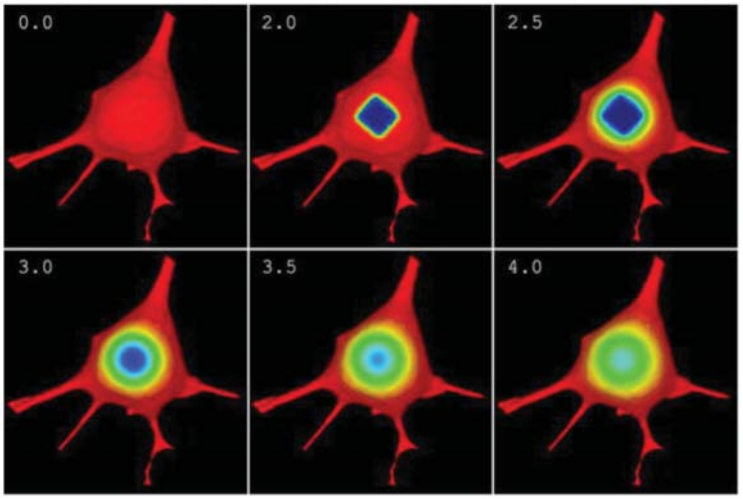
SyBiL/SBPax

Sybil is short for Systems Biology Linker. It is a type of software that can load, analyze, and display data in the Biological Pathway eXchange format (BioPAX). Biological pathways, reaction and signaling networks are made publicly available in formats standardized to the BioPAX format. Simulation tools for biological pathways are standardized to other markup languages such as Systems Biology Markup Language (SBML) and Virtual Cell Markup Language (VCML).

5D Visualization MicroApp

This ImageJ plugin overlays 4D movies into a 5D display. This allows comparison of separate 4D

Figure 3.
Virtual FRAP Tool



A time sequence illustrating diffusion of a molecular species across a 3D cellular membrane after photobleaching in a FRAP experiment. Time 0.0 shows an evenly distributed concentration of a diffusible molecule (in red, high concentration) followed by a bleaching event at time 2.0, which decreases the molecule's concentration (shown in blue). As the molecules diffuse from the unbleached membrane to the bleached area, a concentration gradient (green, yellow) can be seen.

Moraru II, Schaff JC, Slepchenko BM, Blinov ML, Morgan F, Lakshminarayana A, Gao F, Li Y, Loew LM. Virtual Cell modelling and simulation software environment. IET Syst Biol. 2008 Sep;2(5):352-62. Available at: www.ncbi.nlm.nih.gov/pmc/articles/PMC2711391/.

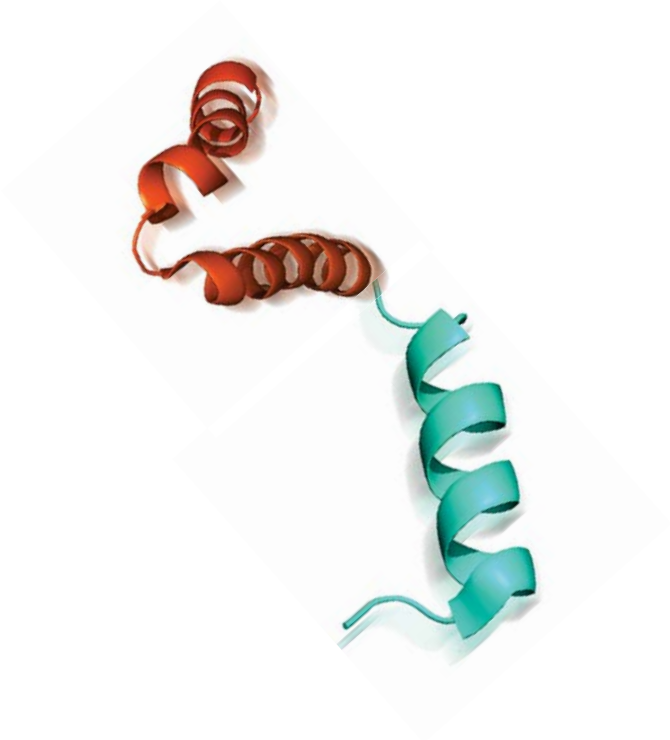
microscopy recordings, combination of different microscopy channels, or overlay of microscopy and model data in 4 dimensions.

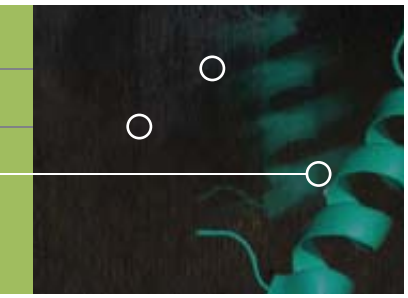
Computational Resources

Computational resources include two dedicated environmentally controlled server rooms housing a variety of hardware resources, currently with >600 gigaflops computer power, >20 terabyte storage, and >500 gigabits per second switching bandwidth, and undergoing continuous expansion.

Training Opportunities

The Virtual Cell short course is held each June. A live-cell microscopy workshop is held each fall. Tutorials on Virtual Cell and BioNetGen are available online through CCAM's website.





Center on Proteolytic Pathways

Transforming the way the world looks at proteases
Burnham Institute for Medical Research
Principal Investigator: Jeff W. Smith, Ph.D.
http://cpp.burnham.org/

Introduction

Why a Center on Proteolytic Pathways? The study of proteases, their specific targets and products, and the biological pathways in which they are involved promises to shed light on hundreds of disease processes. More than 60 human diseases result directly from mutations in proteases or their natural inhibitors, and the Human Genome Project revealed at least 553 human proteases. These are now known to be involved in virtually every type of human pathology—from infection to inflammation, from cancer to Alzheimer's, yet only about 100 of them have been characterized in any detail.

Diseases and pathologies are only part of the story, however. Proteases are essential in many necessary "normal" physiological processes. For example, they make it possible to digest nutrient proteins, they are involved in clipping off parts to activate other proteins, and they facilitate the necessary degradation of old and no-longer-needed proteins so their constituent amino acids may be recycled.

What is a protease?

A protease is an enzyme, the technical name for a protein that catalyzes or speeds up chemical reactions in a living organism. Proteases specifically break or "hydrolyze" peptide bonds. This reaction is also referred to as "proteolysis," the breaking up of proteins. The peptide bond is extremely stable at neutral pH (half-life on the order of 10 to 100 years), yet in the presence of the proper protease, the reaction can occur within milliseconds, as is frequently necessary in biological systems.

Proteases come in five major categories, serine, cysteine, threonine, aspartic, and metalloproteases. The names tell us something about the basic mechanism of each type. For example, serine proteases use serine as the active agent (the nucleophile) in the reaction.

Peptide Bonds:
The chemical link between two amino acids, formed when the carboxyl end of one combines with the amino end of another. This bond is essential to all proteins and peptides.

Nucleophile:
The reactive portion of a protease, active because it is attracted to a positively charged atom of another protein and can give up an electron pair to balance the charge. The term literally means "nucleus loving."

The Center's mission

The Center of Proteolytic Pathways (CPP) was created to develop technologies designed to transform the current reductionist and static understanding of biological systems, particularly proteolytic networks, into a comprehensive and dynamic view.

Bringing Proteolysis to Life

Because proteolysis is so fundamental to biology, the CPP has focused on six basic questions that need to be answered about every protease:

1. Which proteases are involved in a given biological event?
2. What is the magnitude of the proteolytic event?
3. Where does the event take place?
4. What molecule(s) does the protease act on?
5. What products are produced and are they biologically active?
6. What pathways precede and follow the proteolytic event?

In a way, these questions can be boiled down to four: What protease is acting in a given biological process? What is the protease acting on? What is the product of the proteolysis? Where does all this occur?

To answer them, the center has focused on developing four technologies.

1. Protease activity–based profiling is a major step in answering the first question. It uses probes that selectively attach to the active site of the protease in question, based on the specific activity of the protease or of the category of proteases to which it belongs. So far, such probes are available for serine and cysteine proteases, and work is ongoing to develop similar probes for metalloproteases.
2. Protease Activity Imaging Technology (PAIT), using phage display technology, moves toward answering the second question. PAIT is a high-throughput screening technique used to identify the peptide sequence on which a given protease operates. Ultimately, this technique will allow prediction of both the normal and pathophysiological substrates of any protease.
3. Product Terminal Isotope Coding (PROTIC) is a tool to answer the third question. This isotopic tagging system attaches to the N-terminal ends of all proteins and peptide products in a system so that the products of proteolysis can be identified and quantified.
4. Using data input from these three technologies, and answers to the first three questions—Who is acting? Who is being acted upon? What is the result?—the Proteolysis Map (PMAP) is a novel set of computational tools and databases designed to model the proteolytic event—a step toward understanding precisely where and when it occurs. The PMAP Web site already includes a database containing the sequences of every protease coded

Demyelination:
The destructive removal of myelin, the protective fatty protein that surrounds nerve cells. This tissue is essential to the nerve cell's ability to transmit nerve signals.

by the human genome, a database of documented proteolytic events, and tools for examining protease structure and proteolytic pathways (see Available Resources, below).

Highlights

MMP-25 Is a Possible Drug Target in Multiple Sclerosis

Researchers at the Center on Proteolytic Pathways have identified a metalloprotease as a possible drug target for the treatment of multiple sclerosis (MS). This progressive autoimmune disease involves demyelination, which inhibits the ability of brain and other central nervous cells to communicate with each other effectively. Progression of this disease is highly variable; it can be episodic, intermittent, or steady. Likewise, the degree of disability caused by MS varies greatly and may advance over decades. Two-thirds of deaths among those with MS are directly related to its consequences and MS patients live on average 5 to 10 years less than unaffected people. No cure is currently known.

Center investigators used an animal model of MS, called experimental autoimmune encephalomyelitis (EAE), to study the role of matrix metalloproteases (MMPs) in fragmenting myelin basic protein (MBP) and demyelination. The investigators found that bone marrow–derived MMP-25 in immune cells called lymphocytes cleaves MBP in the myelin better than all other MMPs and readily inactivates crystallin αB (CRYAB), which normally acts as a

Investigators

Jeffrey W. Smith, Ph.D., is the principal investigator of the Center on Proteolytic Pathways. He has a longstanding interest in developing new activity-based proteomics methods, in which small molecules are used to tag active sites within enzymes. Probes that react broadly with many members of an enzyme class allow quantification and identification of active enzymes within any biological sample. This puts the power of chemistry to work as a "separation tool" to parse the proteome into easily understandable classes. Chemical probes that are used for activity-based proteomics bind at the active site of a class of enzymes, so they can be used to screen for drugs that inhibit a particular enzyme.

Matthew Bogyo, Ph.D.
TRP 1 Protease Activity-Based Profiling

Adam Godzik, Ph.D.
TRP 4 Proteolysis Map (PMAP)

Andrei Osterman, Ph.D.
TRP 4 Proteolysis Map (PMAP)

Guy Salvesen, Ph.D.
TRP 3 Product Terminal Isotope Coding

Bonnie Sloane, Ph.D.
TRP 3 Product Terminal Isotope Coding

brake on several inflammatory pathways. CRYAB inactivation makes the EAE worse in affected mice.

To test the hypothesis that persistent inflammation caused by viral and bacterial infections is an important step toward autoimmune MS, the investigators showed that such inflammation activated MMP-25 in immune cells. MMP-25 action created MBP fragments, which caused production of reactive T cells that cross the blood-brain barrier and produce MS. The evidence gained in these experiments strongly suggests that MMP-25 could be a useful drug target in MS.

Shiryaev SA, Remacle AG, Savinov AY, et al. Inflammatory proprotein convertase-matrix metallanoprotease proteolytic pathway in antigen-presenting cells as a step to autoimmune multiple sclerosis. Journal of Biological Chemistry, 2009;284:30615–30626.

Studying the Roles of Proteases in Cancer

Investigators at the Center on Proteolytic Pathways have developed live cell assays and a model to study the dynamic roles played by proteases in cancer. Proteolysis in cancer is complex because the source of the proteases can be cells inside or outside of the tumor, and proteolytic activity varies from one type of cancer to another. A single protease can play different parts within the same cancer from one stage to another. Tumors show increased protease expression, but whether proteases cause malignant progression is still unknown in most cases. A final complication is how little is known about what the proteases are acting on and what the products of the action are.

New techniques developed by the Center on Proteolytic Pathways for live cell imaging of protease action show potential for answering some of these questions. These new methodologies are making it possible to measure and localize protease activity, both essential for determining protease activity in a tumor and whether the activity is involved in the cancer's progression.

The investigators call their model MAME (mammary architecture and microenvironment engineering) because it was designed to mimic human breast architecture. Used with their live cell–imaging assay, they can assess the contribution of various breast cells and cell-to-cell interactions to proteolysis. This model system makes it possible to discriminate between the activity of one class of protease and another, using selective probes or protease inhibitors. (See Figure 1 for a schematic representation of the model. Note the presence of DQ, or dye-quenched collagens, which reveal their proteolytic digestion by giving off fluorescent particles.)

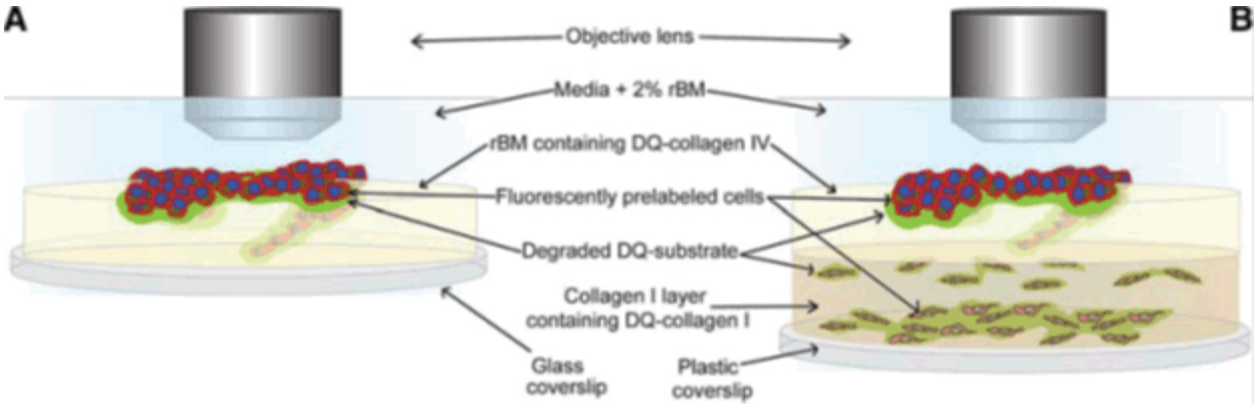
Because the development of new blood vessels (called angiogenesis) is critical to tumor growth and metastasis, and proteases are involved in degrading extracellular tissues so angiogenesis can proceed, the investigators looked at which cells and proteases are involved. They found that several proteases participate, suggesting the need to target multiple classes of proteases in therapeutic interventions.

These new assays and models can be used to (1) identify the proteases and pathways that are potential therapeutic targets, (2) test reagents such as inhibitors to stop protease activity, and (3) develop imaging probes for in vivo use. Models

comparable to MAME can be developed for the study of proteolysis in other tissues of interest.

Figure 1.

A. 3-dimensional fluorescently labeled "overlay cell culture" grown on reconstituted basement membrane (rBM). B. The tripartite MAME model. (Used by permission)



Sameni M, Cavallo-Medved D, Dosescu J, et al. Imaging and quantifying the dynamics of tumor-associated proteolysis. Clinical and Experimental Metastasis, 2009;26:299–309.

Available Resources

The Proteolysis Map (PMAP) contains a variety of databases, including proteolytic events, protease sequences, proteolytic substrates, proteolytic networks and signaling pathways, substrate prediction, and a toolkit, all available at www.proteolysis.org/proteases.