NEW CHALLENGES IN THE LIFE SCIENCES

PRIORITIZING EUROPEAN RESEARCH IN MOLECULAR SYSTEMS BIOLOGY

A DOCUMENT OF THE COORDINATION ACTION NMR-LIFE

AND THE FORUM FOR EUROPEAN STRUCTURAL PROTEOMICS (FESP)

Motivation

Research in the post-genomic era is moving toward new horizons, which are largely embraced by the broad definition of *Molecular Systems Biology*. To encourage discovery in this epoch, we need to prioritize and focus on the future of such research. Where do we want to go, and how will we get there?

Systems biology addresses the properties of entire biological systems and subsystems as opposed to the isolated study of their individual components. Fundamental properties of biological systems rely on the spatial and temporal interactions of the macromolecules that compose the system and can only be understood by looking at the system as a whole. An understanding of *molecular* systems biology, i.e. the ability to *model* systems to predict biological outcome at the molecular level, is unthinkable without an appreciation of the dynamic structure of proteins, the specificity of protein-protein interactions and the resulting properties of molecular machines, pathways and entire networks. We believe that this molecular "protein" perspective is critical for systems biology to have a much-needed impact on medicine and pharmacology.

A symposium entitled, "New Challenges in the Life Sciences: Prioritizing European Research in Molecular Systems Biology", was held on October 18-19, 2007 in Florence, Italy, to provide a forum for the scientific community involved in on-going major European post-genomic projects to discuss the importance of a molecular research in advancing systems biology. The meeting was sponsored by the European Commission (EC) through the *Coordination Action NMR-Life*, in conjunction with the *Forum for European Structural Proteomics (FESP)*. It is hoped that the newly established European Infrastructures in the Biomedical Sciences, as outlined by the *European Strategy Forum for Research Infrastructures (ESFRI)* and implemented by the European Commission, will find this discussion helpful in fine-tuning their funding objectives. The symposium provided an opportunity for the participants to share their work and make their voices heard as we look toward the future of molecular systems biology.

The meeting involved about 30 experts, representing the multidisciplinary nature of the field, and included scientists and policy-makers from the US, China, Japan, India and Europe. The meeting consisted of four sessions that included presentations from invited speakers followed by group discussions. The first session addressed the new challenges in protein chemistry associated with molecular systems biology, the second focused on the role of structural biology from a systems biology perspective, and the third on the impact of molecular systems biology on molecular medicine. The fourth session reviewed funding opportunities and strategies.

This document represents a synthesis of the contributions of the participating scientists and strives to define the importance of a molecular foundation for a systems-level understanding of biology, to identify the opportunities that will be afforded by a '3D' view of biological systems, and to recommend funding priorities for advancing molecular systems biology."

A vision for molecular systems biology

Individual protein and RNA structures, as well as protein-protein and protein nucleic acid complexes provide a knowledge base that is robust and detailed. This structural foundation, in turn, constitutes a basis for extrapolation from a given biological system to produce testable hypotheses about its response to perturbations. The information derived from one system can then be extrapolated to homologous processes. Structural data are essential to generate predictions of how genetic variation affects protein activity and, ultimately, impacts phenotype. Likewise, the resulting framework offers the ability to hypothesize how networks in humans may operate and respond by analogy with studies of the corresponding networks in model organisms.

The most critical opportunities:

- To understand a biological system it is critical to move beyond a focus on its individual components to the next level of complexity, i.e. an understanding of how the components assemble into functional units.

- Biomolecules for which a three-dimensional structure can be obtained represent particularly robust building blocks for starting systems-level integration.

- Experimentally-derived multi-scale structural information will enable the objective and reliable visualization of cellular organization from the molecular to the systems level.

- A structure-based scaffold will provide an ideal platform for integrating available biological data and will be synergistic with other systems biology efforts.

- Using this structure-based platform will be a particularly effective means for validating and interpreting genetic variation as it relates to disease and will guide more informed and precise therapeutic interventions.

State of the art

Many problems of key importance to human quality of life, health, economic development, agriculture and the environment depend critically on being able to predict the behavior of complex biological networks in the healthy organism as well as how they are perturbed in disease. Being able to predict the behavior of these networks is key both to tailoring available therapies to individual patients and to developing new therapies. The essence of systems biology lies in its ability to formulate these predictions. Proteins constitute a fundamental element of biological systems. The comprehensive analysis of their structure, location, function and mechanism of action, as well as of their interactions with each other

and with other components of the cell, provides information crucial to the development of systems biology.

A long term goal of the life sciences is the complete comprehension, at the molecular level, of the processes at the basis of Life, which are essential for living organisms to survive and account for the function or malfunction of cells and tissues. Molecular systems biology is an integrative discipline that has evolved out of the need to describe the behavior of complex biological systems in terms of their molecular components and interactions; it provides a platform for data collection, data analysis, integration of data from various sources, and modeling of molecular and cellular phenomena.

Historically, molecular biology has progressed through the identification of individual genes and proteins and the study of their individual functions - components of biochemical pathways have been analyzed as though they were independent pieces of a larger puzzle. This approach imposes limitations, however; an organism is more than the sum of its individual functional processes, and each process is affected by all of the others. In the case of the human, we are left with an unclear picture of how the human body functions, and how we can best approach attempts to predict, prevent, and treat health problems. Efforts to cure complex diseases have met with limited success because only individual aspects of the organism have been studied at one time. Molecular systems biology approaches the study of an organism as an integrated and interacting network of genes, proteins, and biochemical reactions – and it is now well recognized that it is this *system* that gives rise to and maintains Life. Individual functional processes must therefore be studied in the context of an entire cell or organism, and not in isolation. It is the *interactions* that are ultimately responsible for the form and function of any organism.

Structural genomics efforts have provided a large number of structures that can be used for the next level of integration. In parallel, it has been possible to obtain a first coarse map of the cellular machinery of the model organism yeast by purification of complexes and mass spectrometry. Several additional technological platforms are available to systems biology, such proteomics, metabolomics, etc. Taken together, these provide the basis for 3D systems biology.

The Frontiers of Molecular Systems Biology

In order to achieve the goals outlined above, we must obtain a detailed description of the molecular components, their interactions, and abundances under a wide range of cellular and *in vitro* conditions. To this end, we require a set of new methodologies and technologies, and the acquisition of appropriate large-scale datasets. Specifically:

- Systematic analysis of structures and interactions of proteins, protein-protein and protein-nucleic acid complexes

- Innovative computational and experimental approaches to studying the effect of variations in the sequence and of post-translational modification on the structure and function of proteins and their complexes, e.g. somatic mutations in cancer, by experimental and computational approaches

- Development of tools for measuring the often transient interactions between components of the complex systems involved (including dynamics and structural changes affecting function)

- Development of novel knowledge management systems capable of describing the components of such complex systems and the assurance of their reliability, for the purpose of being incorporated into models

- Systematic and parallel *in vivo* and *in vitro* studies of protein interactions and biochemistry at intermediate levels of system complexity

- Elucidation and modeling of the principles underlying signaling circuits

- Analysis of host-pathogen interactions

- Understanding microbial communities and interpreting metagenomic data

- Development of models incorporating dynamic, structural and mutational information able to predict interactions in biological systems

- Identification of therapeutic opportunities for treating common human diseases

The Role of Structural Biology from a Systems Biology Perspective

A cell may be compared to a car, for which, in order to understand its workings, we need to examine and understand the component subsystems, e.g. the gear train, transmission, fuel system, etc. Structural biology can make a crucial contribution to analyzing these subsystems, ultimately working towards understanding how proteins function in an *in vivo* context. A goal would be to obtain a mathematical description of networks and systems with the ultimate goal of achieving predictive understanding of a system or of a subsystem. This encompasses dynamic, often transient, interactions in a cellular context, and examining the nature of response to structural disorder in such interactions, moving away from solely looking at ordered domains. The goal is an understanding of the impact of these features on cellular systems and their role in disease mechanisms, e.g. in Alzheimer's, prion and other neurodegenerative diseases, in diabetes, cardiovascular disease, infectious diseases, both bacterial and viral, and in cancer.

We need to emphasize steps forward in key structural biology techniques to facilitate this. This requires continued development of individual technologies, but, importantly, also their key synergistic implementation. To combine and integrate the different techniques, we need advances in methods but also improved and new computational tools. A key objective will be to integrate the whole system as a function of time. Quantitative information is crucial for systems biology.

Important techniques where further improvements will be necessary span many areas, from structural to computational biology and from mass spectrometry to knowledge management.

Examples include: 1) X-ray crystallography (Remote access; automatic analysis of maps for ligand recognition; improvement in collection of data on microcrystals); 2) NMR (*in vivo*

methods; transient interactions; disordered systems; solid state methods; improved computational methods); 3) Electron Microscopy (Single particle cryo EM for large complexes; cryoelectron tomography of cellular structures); 4) SAXS; 5) Protein Production

Impact of molecular systems biology on molecular medicine

The following opportunities will be attainable through a molecular systems biology approach:

A. Understanding Disease

- Systems level understanding of genetic perturbation reflecting disease. Efforts from medical genetics and mouse models of disease will become more easily interpretable.
- The nature of many disease mutations will become interpretable by mapping effects of the molecular defects on 3D protein complex structures, including inter-complex protein-protein interactions
- Understanding the mechanisms of complex diseases will profit from integration of data at the level of the molecular machine, of the entire pathway and at the inter-pathway level
- Extrapolation of mutant-protein-effect relationships from a validated system to other diseases becomes possible

B. Treating Disease

- Entire disease pathways in three-dimensions will instruct the identification and use of research compounds, interfering peptides and drug leads
- “Poly-pharmacology” (off-target effects) may be turned into “Systems Pharmacology” (where multiple targets are hit on purpose)
- The 3D systems biology understanding of disease may allow the informed treatment of multifactorial diseases and the use of combination therapy
- More efficient target and scaffold hopping (the same drug class targeting a particular protein class can be usedto hit an analogous target and the same target may be hit by unrelated chemical scaffolds)

The recommended research path to achieve the goals above should comprise the following steps:

Identification and structural characterization of individual elements → correlation of structure-function relations of these elements→ experimental super-structural studies → 3D modeling → experimental perturbation of the system by mapping disease mutations → modeling and simulation to predict the molecular disease mechanism(s) and to propose drug targets / diagnostic markers

These steps would have as their foundation and strength:

1. Build on the continuation of the structural genomics effort to deliver the shapes and structures of components to enable 3D systems biology
2. Map human cellular machinery
3. Obtain hybrid structural information from patchwork analysis of suitable technologies
4. 3D computational modeling
5. Integrate experimental studies, disease genetics and pharmacological perturbation of the system of interest
6. Modeling/Simulation of diseased states. Prediction of diseased network output and small molecule effect

Appendix. Available funding instruments in the EU

Within the 7th Framework Programme:

» Cooperation «: Health 3rd call 2008, 4th call 2009

ICT programme: (bioinformatics, databases)

» Capacities «: Research infrastructures: resources, facilities and related services aimed at top-level research

- Support to existing research infrastructures provided for Integrated Activities (bottom-up and targeted approaches) and e-Infrastructures
- Support to new research infrastructures provided for Design Studies and Construction of new infrastructures
- Support for policy development and program implementation

ERA-NET schemes permit coordination between relevant national research programmes of Member States.

International dimension

<u>EU-USA interaction</u>

Project participation level

FP7: US partners can participate in proposals, but, in most cases, cannot receive EU funding

NIH: EU and Canadian partners could be funded (like EBI) based on decision of the appropriate governing board, or could be subcontractors of US partners

Programme cooperation and coordination level

Project cooperation and coordination: example of PSI: international cooperation with Wellcome Trust (UK) established: 2000 Cambridge meeting (tackled: international coordination; policies on data-release, international coordination of target selection, facilities)

IKMC: international knock-out mouse consortium – running since spring 2007: EC-NIH-Genome Canada

ICGC: International Cancer genome consortium – in preparation

EU -Third countries

Bilateral cooperation agreements: (EU-China, EU-India, EU-Russia): areas identified; implementation through » SICA « calls in FP7 (3rd call: Health, topics with China and Russia foreseen)

Mechanisms for identifying joint areas of cooperation at the international level

Agreements (top-down)

EC-US Task Force on biotechnology (bottom-up); workshops on yearly basis in selected areas, produce policy documents for joint actions (Workshop on infrastructures in Systems biology, May 2007)

» Ad-hoc initiatives « Systems biology of cancer (EC-US) May 2008

EU: Identification of inputs to the creation and » fine tuning « of FP7

1) Legal procedure for adoption of the FP7 (codecision between Council and European Parliament)

2) Inputs to annual work programme (scientific topics and instruments: how are the broad topics defined for the FP7?)

1. Advisory groups for health research and research infrastructures
2. Input from on-going projects: SSA, CA (FESP)
3. Programme committees (Member States)
4. International dimension: implementation of political initiatives from agreements between EC and other countries
5. Addressing the needs of emerging international consortia (Mouse genome)

Top-down (FP7) vs. bottom-up:

- Bottom-up: expression of interest; 2-stage procedure;

Choice of instruments: depends on the scale and ambition of the issue addressed

The situation in USA (NIH)

The PSI is ongoing: large-scale initiatives started in late 90's. At present, the 2nd stage of the PSI is focused on selected targets in order to achieve coverage of large protein families. In the future a possible target will be the human gut microbiome, but this may

depend on a decision with respect to funding of the 3rd stage of the PSI, which should be taken in early 2008.

Systems biology: 6 National centers for systems biology active: interdisciplinary, based on collaborative research

PSI-systems biology: possibilities of interaction will be explored in the future (e.g. interactions with Functional Glycomics Center, the National Center for Research Resources-NCRR, Synchrotrons, the National Cancer Institute-NCI)

US-Third countries: - policy initiatives

NIH -Wellcome trust; Japan (RIKEN); China – ISGO conferences

Official agreements with China, Japan, India…

Research Infrastructure financing to participants from third countries is possible, depending on evaluators.

FLORENCE, OCTOBER 18-19, 2007

Participants in the Symposium:

Lucia Banci	*CERM, University of Florence and FESP*
Wolfgang Baumeister	*Max Planck Institute for Biochemistry, Martinsried*
Ivano Bertini	*Coordinator, Coordination Action NMR-Life Director, Magnetic Resonance Center (CERM), University of Florence*
Roslyn M. Bill	*Aston University, Birmingham, Coordinator of E-MEP*
Steven E. Brenner	*University of California, Berkeley*
Bruno Canard	*CNRS, Université de Provence Université de la Méditerranée*
John Christodoulou	*University of Cambridge*
Bruno Domon	*Institute of Molecular Systems Biology, ETH Zurich*
Peter Henderson	*University of Leeds, Coordinator EU EDICT project*
Sasa Jenko	*European Commission, Directorate General for Research Genomics and Systems Biology*
Patrik Kolar	*European Commission, Directorate General for Research Genomics and Systems Biology*
Joanne Kotz	*Nature Chemical Biology, Senior Editor*
Andreas Kremer	*SienaBiotech SpA*
Peter Kuhn	*The Scripps Research Institute, La Jolla*
Hans Lehrach	*Max Planck Institute for Molecular Genetics, Berlin*
Zhi-Jie Liu	*Chinese Academy of Science, National Laboratory of Biomacromolecules*
Claudio Luchinat	*Magnetic Resonance Center (CERM), University of Florence*
John Moult	*Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute*
Pär Nordlund	*Karolinska Institutet, Stockholm, SGC Consortium*
John Norvell	*NIH, Director of the Protein Structure Initiative*
Carlo Rizzuto	*Chair European Strategy Forum on Research Infrastructures (ESFRI)*

Antonio Rosato	*Magnetic Resonance Center (CERM), University of Florence*
Harald Schwalbe	*Biological NMR Group, Johann Wolfgang Goethe University, Frankfurt*
Israel Silman	*Weizmann Institute of Science, Rehovot*
Lalji Singh	*Director of the Centre for Cellular & Molecular Biology, India*
David Stuart	*Oxford University, Coordinator of SPINE2*
Giulio Superti-Furga	*Austrian Academy of Sciences, Director of the Center for Molecular Medicine*
Joel Sussman	*Coordinator Forum for European Structural Proteomics (FESP)*
Purificacion Tejedor Del Real	*European Commission, Directorate General for Research, Research Infrastructures (Unit B3)*
Georg Terstappen	*Siena Biotech SpA*
Shigeyuki Yokoyama	*RIKEN - Yokohama, Director of the Protein Research Group*

Welcome Address by: Carlo Rizzuto Chair European Strategy Forum on Research Infrastructures (ESFRI)