BER Molecular Science Challenges

Workshop Report

Atmosphere
Oceans
Continents

Aquatic Basins
Regions

Aquatic Biomes
Terrestrial Biomes

Ecosystem
Ecology, Watersheds,
and Field Plots

Communities,
Organisms, and Cells

Biobased Factories

Pathways and
Networks

Proteins and
Protein Machines

Genomes

Atmosphere-Land Surface Interactions

Near-Surface and Below-Surface Interactions

Synthetic and Genomic Bioscience

U.S. DEPARTMENT OF ENERGY | Office of Science

Office of Biological and Environmental Research

Office of Biological and Environmental Research
Molecular Science Challenges

Report from the May 27–29, 2014, Workshop
Germantown, Maryland

Co-Chairs

James C. Liao
University of California, Los Angeles

Judy D. Wall
University of Missouri, Columbia

Breakout Group Chairs

Vicki Grassian
University of Iowa

Michael Thomashow
Michigan State University

Norman Dovichi
University of Notre Dame

Breakout Group Rapporteurs

Scott Bridgham
University of Oregon

John Bargar
SLAC National Accelerator Laboratory

Michael Crowley
National Renewable Energy Laboratory

BER Liaisons

Paul Bayer
Climate and Environmental Sciences Division
Paul.Bayer@science.doe.gov, 301.903.5324

Roland Hirsch
Biological Systems Science Division
Roland.Hirsch@science.doe.gov, 301.903.9009

Workshop convened by the

U.S. Department of Energy
Office of Science
Office of Biological and Environmental Research

Mission

The Office of Biological and Environmental Research (BER) advances world-class fundamental research programs and scientific user facilities to support the Department of Energy's energy, environment, and basic research missions. Addressing diverse and critical global challenges, the BER program seeks to understand how genomic information is translated to functional capabilities, enabling more confident redesign of microbes and plants for sustainable biofuel production, improved carbon storage, or contaminant bioremediation. BER research advances understanding of the roles of Earth's biogeochemical systems (the atmosphere, land, oceans, sea ice, and subsurface) in determining climate so that it can be predicted decades or centuries into the future, information needed to plan for energy and resource needs. Solutions to these challenges are driven by a foundation of scientific knowledge and inquiry in atmospheric chemistry and physics, ecology, biology, and biogeochemistry.

Office of Biological and Environmental Research
Molecular Science Challenges

Workshop Report

Published: April 2015

U.S. DEPARTMENT OF **ENERGY** | Office of Science

Office of Biological and Environmental Research

Acknowledgements

Table of Contents

List of Figures

Executive Summary

In its *Global Risks 2014* report, the World Economic Forum (WEF; WEF 2014) lists 10 current global risks of highest concern. Of these, four are science-related and all are central to the mission scope of the U.S. Department of Energy's (DOE) Office of Biological and Environmental Research (BER): (1) water crises, (2) failure of climate change mitigation and adaptation, (3) greater incidence of extreme weather events, and (4) food crises tightly linked to the use of biomass for sustainable fuel production. The WEF report also notes the interdependencies between environmental and societal risks that make solutions to these stresses ever more important for global stability.

These solutions will hinge on understanding the underlying processes of entire systems—from molecular to global scales—and will range from redesigning microbes and plants to understanding the roles of Earth's biogeochemical systems in determining climate and contaminant transport and transformation. Clearly, however, the success of these endeavors critically depends on understanding the reliable coupling between molecular and mesoscale phenomena.

Advancing efforts to determine sustainable solutions was the impetus for the BER Advisory Committee's (BERAC) report *Grand Challenges for BER: A Long-Term Vision* (BERAC 2010). A subsequent report, *BER Virtual Laboratory: Innovative Framework for Biological and Environmental Grand Challenges* (BERAC 2013), focused and expanded on the "technology and tools most needed to support the biological and environmental research necessary" to meet the challenges and opportunities addressed in the 2010 report. BER is faced with the monumental challenge of building a foundation for a continuum of understanding across a range of spatiotemporal scales—from the molecular to the global scale and from picoseconds to millennia—that is perhaps unique in the sciences.

For this effort to succeed, integration of multiscale observations, experiments, computation, theory, and knowledge is essential, but much of today's research is still discipline-limited, focusing on single-scale system components with little consideration of any impact on interacting disciplines. Fragmentation hinders potential advances in constructing predictive models by limiting the synergies that come from integrating research and data across disciplines and scales.

Operating within DOE's Office of Science, BER convened the Molecular Science Challenges workshop (see Fig. 1. Molecular Science Challenges in Biological and Environmental Research, p. vi), which brought together scientists representing the full spectrum of BER program elements that depend on molecular science. BER research—spanning biological, environmental, and climate sciences—has evolved over recent years to require a much more robust understanding of the molecular systems and processes that underpin program goals. The workshop was held May 27–29, 2014, in Germantown, Maryland. Participants were tasked with assembling 10-year projections of scientific and technological challenges and opportunities in molecular science relevant to BER's mission and, once these were identified, with developing high-level progressions of scientific objectives to address these challenges and opportunities. Workshop participants were assigned to one of three breakout groups: (1) atmosphere–land surface interactions, (2) near- and below-surface interactions, and (3) synthetic and genomic bioscience. Although the disparate disciplines and expertise of the participants led to far-ranging discussions and revealed the enormity of the scope of BER activities, the workshop did result in identifying gaps in
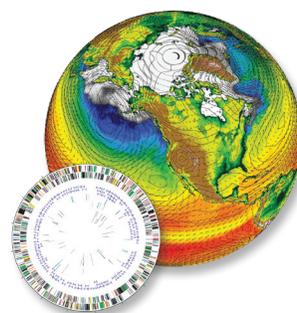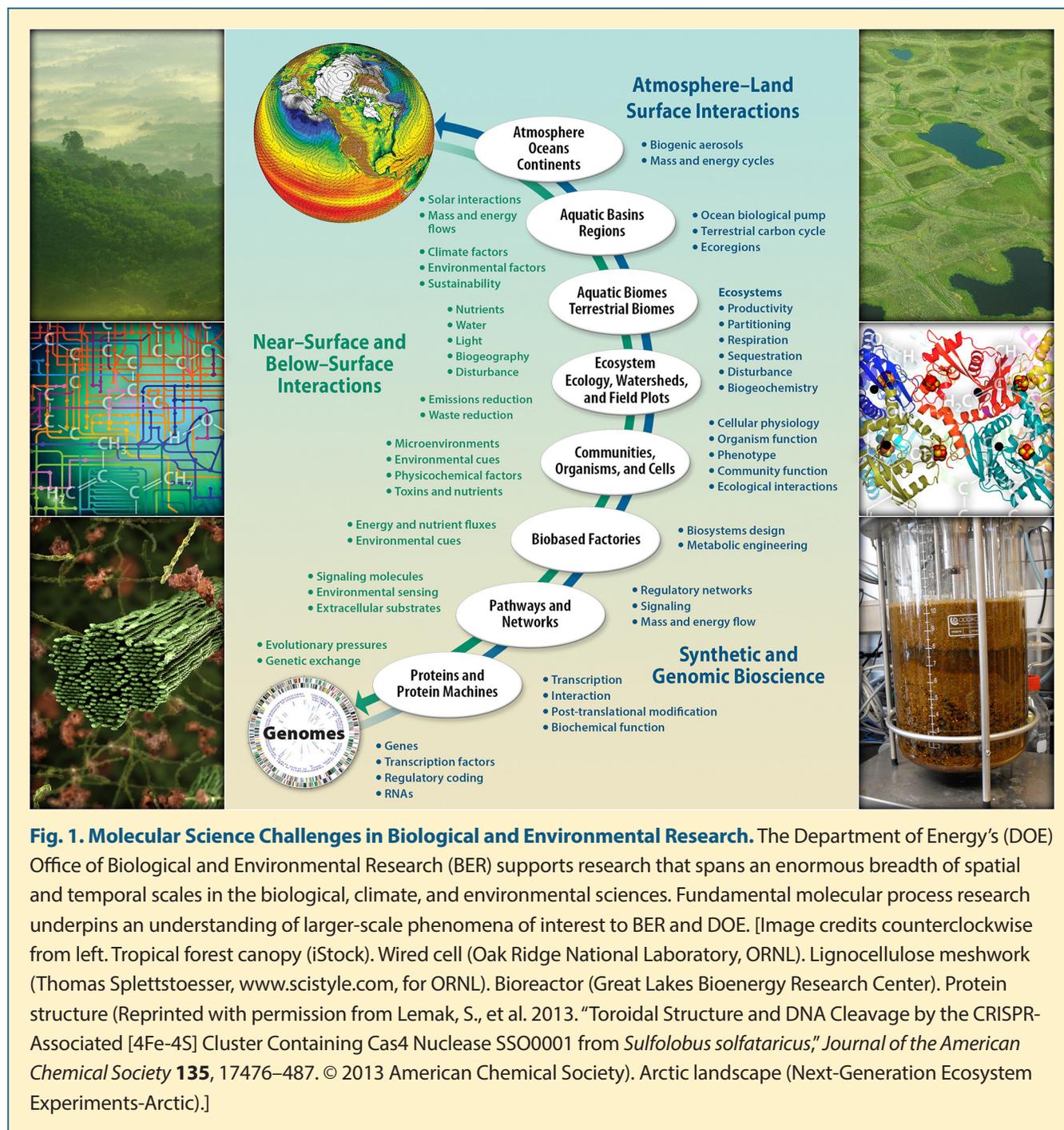
Fig. 1. Molecular Science Challenges in Biological and Environmental Research. The Department of Energy's (DOE) Office of Biological and Environmental Research (BER) supports research that spans an enormous breadth of spatial and temporal scales in the biological, climate, and environmental sciences. Fundamental molecular process research underpins an understanding of larger-scale phenomena of interest to BER and DOE. [Image credits counterclockwise from left. Tropical forest canopy (iStock). Wired cell (Oak Ridge National Laboratory, ORNL). Lignocellulose meshwork (Thomas Splettstoesser, www.scistyle.com, for ORNL). Bioreactor (Great Lakes Bioenergy Research Center). Protein structure (Reprinted with permission from Lemak, S., et al. 2013. "Toroidal Structure and DNA Cleavage by the CRISPR-Associated [4Fe-4S] Cluster Containing Cas4 Nuclease SSO0001 from *Sulfolobus solfataricus*," *Journal of the American Chemical Society* **135**, 17476–487. © 2013 American Chemical Society). Arctic landscape (Next-Generation Ecosystem Experiments-Arctic).]

knowledge and tools needed for molecular science research and development. The following workshop charges guided the breakout group discussions:

- Understand the molecular systems and processes that underpin BER program goals.

- Integrate across the breadth of spatial and temporal scales of BER research areas.

- Take advantage of DOE national laboratory and facility resources.

- Identify molecular science challenges and opportunities.

- Describe research pathways to overcome barriers in BER-relevant molecular science over a 10-year timeframe.

The workshop's goal was to identify knowledge gaps that must be filled and to imagine and suggest tools—either not yet readily accessible or not yet in existence—that could provide data for knowledge growth and development. Workshop organizers guided the participants through structured working sessions to identify broad areas of opportunity for future research and capability development. These opportunities include understanding, describing, and modeling molecular- to global-scale processes based on synergistic, integrated, multidisciplinary approaches that will enable better informed policy decisions to address future needs.

This report provides individual summaries of the molecular science research needs developed from the (1) atmosphere–land surface interactions, (2) near- and below-surface interactions, and (3) synthetic and genomic bioscience breakout discussions and concludes with a summary of the workshop's cross-cutting themes.

Atmosphere–Land Surface Interactions

From an overall decadal standpoint energy, climate, and the environment are intimately woven together. Therefore, balancing the increasing need for energy with avoiding harm to Earth's climate and environment remains a challenge. To date, there is only superficial understanding of the molecular-level chemical, physical, and biological processes that underpin large-scale climate and environment systems.

1. Exchange Processes Between Land and Atmosphere

Complex land–atmosphere exchanges and interactions affect emissions of gases and particulate matter. In turn, these gases and particles undergo atmospheric reactions that ultimately determine the abundance and efficacy of cloud condensation nuclei and ice nuclei, essential precursors to cloud formation and the hydrologic cycle. Because precipitation from clouds enables the growth of plants and microbial communities, understanding these detailed processes and their interactions is necessary to predict how future energy scenarios will impact the climate and environment. To determine how energy policies could affect global cycles, validated models are needed for identifying the key feedbacks governing Earth's radiation balance, cloud formation, and precipitation, and the dependence of these atmospheric events on plant and microbial community development and evolution.

Fundamental knowledge of molecular-scale processes that results from laboratory studies is essential for understanding the complex interactions and dynamics of chemical and biological processes in the different environmental media, but information from larger scales is needed to describe environmental responses to changes in emissions and the predicted climate change that will result. Therefore, while deciphering molecular-scale processes is essential, mesoscale and macroscale decryption of the interplay between different processes is equally important for forecasting environmental responses to climate changes. However,

the full complexity of the ambient environment cannot be reproduced in the laboratory, and ambient measurements are not likely to elucidate and quantify the links between different molecular-scale factors. Comprehensive macro- and global-scale models are needed to elucidate atmospheric impacts. Other needs are to identify and quantify the chemical and particulate emissions and depositions between land and atmosphere and to determine whether unidentified phenomena, both biogenic and anthropogenic, are operating across interfaces.

2. *Atmospheric Aerosol Links to Radiative Balance, Cloud Formation, and Precipitation*

The capability to predict climate change on a global level is acknowledged to be insufficiently robust. Developing useful models will require knowledge of both aerosol and gas-phase anthropogenic and natural sources, as well as their interplay. Further, coupling the most fundamental molecular, physical, and optical properties of individual aerosols with macroscopic measurements of cloud cover and atmospheric radiation measurements will be essential to improve models. Incorporating these details into models will require correlating molecular data with field measurements and parameterization. Several areas of particular importance are clouds and cloud formation; radiation, with a focus on light-absorbing aerosols; secondary organic aerosols; and particle heterogeneity. Other needs include understanding water and aerosol interactions and the mechanisms and release rates of biogenic emissions.

3. *Terrestrial Ecosystem Impacts from Transfer of Energy, Water, Gases, Organics, and Particles to and from the Atmosphere*

Terrestrial ecosystems are fundamentally affected by transfers of energy, water, gases, and particles to and from the atmosphere. These transfers are regulated by mechanisms that function at multiple levels of biological organization. Therefore, to understand these exchanges, controlling mechanisms need to be coupled across multiple scales, ranging from molecules to landscapes and, in turn, linking regional to global consequences in terms of climate and air quality. Changes in land surfaces are likely to become more intense in the future, and types of land use may become one of the most important terrestrial–atmospheric feedbacks to climate. Although understanding of molecular-scale processes is essential, mesoscale and macroscale understanding of the interplay among different processes and sources—both anthropogenic and natural—is equally important to understanding the environment. Comprehensive models will need to consider exchanges among the atmosphere, soil, plants, surface water, groundwater, and oceans, including transport of mineral-rich particles and organics.

4. *Technology and Capabilities Needed*

To address the molecular science questions posed in the domain of atmosphere–land surface interactions, the essential tools include environmental sensors with extended temporal range, methodologies for analysis of ultralow concentrations and volumes, and high-speed computationally powered access to data archives. To meet workforce requirements, it is necessary to establish, promote, and train a scientific community equipped with the skills to address multidisciplinary science in complex systems. A number of capacities must be generated (or retained) that require a motivated workforce, such as high-performance computing, synchrotron light and neutron source applications, instrument development, knowledge dissemination, and data sharing.

Near- and Below-Surface Interactions

A research goal for BER is to obtain a predictive knowledge of terrestrial ecosystems extending from bedrock to treetops and from global to molecular scales. To do so requires a fundamental understanding of coupled hydrobiogeochemical processes and interactions in complex near- and below-surface ecosystems.

1. Molecular Dependence of Ecosystem Biogeochemistry

Developing new knowledge about linkages between molecular- and ecosystem-scale processes is profoundly important to achieving this BER goal. Ecosystem biogeochemistry is driven by reactions occurring at the molecular scale—at enzymes, cell membranes, organic substrates, and surfaces of minerals and detrital organic material. Moreover, biogeochemical processes are initiated and controlled by the transcription of microbial and plant genes. This molecular machinery and the associated processes reside within the interstitial spaces between sediment (or soil) grains and at interfaces among water, minerals, and living organisms. These physical spaces, their arrangements within ecosystems, and their biotic and mineralogical/chemical makeup are highly heterogeneous in space and time. This complex multiscale heterogeneity significantly influences molecular processes and ecosystem responses to environmental perturbations. To make major advances in our knowledge of biogeochemical and ecological processes, it will be necessary to address this heterogeneity across spatial and temporal scales beyond snapshot studies. The ability to predict ecosystem biogeochemistry rests upon foundational knowledge of these molecular processes; their impact on electron, nutrient, and contaminant fluxes; their locations within the terrestrial environment; and their impacts on systems of increasing scales through the complete ecosystem.

2. Spatial and Temporal Factors Controlling Ecosystem Behavior

Ongoing research reveals important biogeochemical drivers of ecosystem behavior across a range of scales, such as mineral surface-hosted reactions, microbial processes, plant physiology, and "hot spots" and "hot moments" in ecosystems. Transformational discoveries are expected from a robust scientific understanding of interfaces that incorporates the complexity existing among biogeochemical system components,

such as the dynamic interactions between plants and microbes. Linking these ecosystem components in space and time is a major research horizon. Consequently, the decadal goal for this area is to understand the interdependencies of biogeochemical processes and how these processes quantitatively scale from the molecular to ecosystem levels. This understanding can then be used to predict ecosystem response to environmental perturbations such as changing climate and land-use patterns.

Achieving this vision will require studies of genes, plants, microorganisms, enzymes, pore networks within sediments and soils, minerals, solutes, and the interfaces among these ecosystem elements. Length scales ranging from Ångstroms (10^{-10} m) to meters to regional scales, along with temporal behavior from nanoseconds to decades, will need to be coupled. Essential areas to address include warming, elevated atmospheric carbon dioxide concentration, alterations in nutrient cycles, the responses of key plant functional types to perturbations, altered precipitation timing and amounts, and contaminant transport and transformation. Major advances will require systems approaches that incorporate iterations between model-driven laboratory and field experimentation, keen observations, and skill in constructing computational models and simulations from large datasets. Contaminant fates, carbon budgets, biological responses, and nutrient cycles must be determined in diverse biogeochemical settings to provide data that enable predictions of the mobility, reactivity, and stability of these components in complex near- and below-surface environments and in surface water, pore water, and groundwater. This knowledge is needed to improve models that can project both the ecological effects of climate change and the feedbacks between the terrestrial ecosystem and the rest of the Earth system.

3. Objectives to Achieve a Predictive Understanding of Ecosystem Behavior

Specific objectives identified at the workshop are the following:

1. Link organism genotype and system biogeochemistry to predict phenotype and ecosystem processes at cellular, organismal, and community levels.

2. Improve the mechanistic understanding of plant-microbe interactions, with emphasis on the linkage of metabolomes to ecosystem models.

3. Incorporate mechanistic and quantitative knowledge of molecular processes occurring in pores into biogeochemical and ecosystem models through ongoing process discovery and investigation.

4. Determine the physical locations and times in an ecosystem across which biogeochemical reactivity is high and link these hot spots and hot moments to predictively understand ecosystem behavior.

4. Technology Capabilities Needed

Decadal needs include novel and multidisciplinary technical advances allowing two-dimensional and three-dimensional (3D) imaging of dynamic processes from the molecular scale to the nanoscale. It will be necessary to improve the sensitivity of spectroscopic techniques while increasing throughput. Noninvasive geophysical probing techniques and real-time sensors will be needed for subsurface deployment at scales from microns to meters. Computational science will be needed for molecular modeling and simulation and for deriving methods to link processes across scales accurately. Experimental and computational tools will be needed to determine thermodynamic/kinetic constants for specific species and for "omics" analysis and phenotype screening. Advances in supercomputing, imaging, X-ray light sources, and neutron source capabilities will play a major role in solving these problems and are crucial to moving the science forward.

Synthetic and Genomic Bioscience

1. Toward a Multiscale 3D View of Cells

Biology originates from the molecular scale, and rational engineering of biological systems in the energy and environmental biosciences must start at this level. The revolution in the biological sciences that led to the derivation of the "laws of life" has resulted in a mechanistic understanding that forms the basis of modern life sciences, for example, the double helix structure—the "central dogma" of DNA-RNA-protein information transfer and redox-driven energy conversion.

Fueled by molecular-scale information, systems biology goes beyond the consideration of single macromolecules to obtain holistic information about interacting biological systems by exploring metabolic networks, genomics, and proteomics. Enabled by this information, biosystems are being re-engineered for various purposes. Ongoing efforts to collect and interpret omics-based information defining molecular systems for metabolism, regulation, and signaling are supported by BER's investment in a computational system for the integration of disparate types of data: the DOE Systems Biology Knowledgebase or KBase (kbase.us).

The next challenging step will be to incorporate those data, together with high-resolution imaging, into 4D spatiotemporal models that simulate microbial and plant cells. Computational and experimental science will combine—with computational science "integrating" detailed and disparate information from experiments into working models capable of providing causative explanations of the phenotypical behavior of cellular systems. This combination will lead to the discovery of new rules that govern systems-level interaction and the evolution of biomolecules and processes. Although these principles will be more general than individual molecular mechanisms, they nonetheless will be guided by molecular science and the laws of physics and chemistry and thus will be inherently multiscale. New process principles will drastically reduce the complexity of data

analysis and will be used to formulate guidelines for the redesign and construction of biological systems.

2. Multi-Resolution Experiments

Physical descriptions of microbial and plant cells critically depend on experimental capabilities to visualize, conceptualize, and test on relevant length and time scales. Although the use of systems models to discover new guiding principles may not require a complete "parts list," focused, high-resolution 3D imaging of biological events will be critical for identifying the players in key processes. Identifying critical components will require detailed knowledge of the structures and dynamics of interacting macromolecules; imaging of subcellular structures; determination of the positions of macromolecules with respect to these expressed structures; and knowledge of the distributions of smaller species such as solvents, metabolites, and ions. Another requirement is the ability to observe and measure the impacts of these molecular systems across interfaces and at successively linked system scales. Spectroscopic, chemical mapping, and imaging techniques will need to be developed to enable high-resolution, time-resolved investigation of complex processes, such as energy conversion and electron flow, in four dimensions. BER is ideally equipped for coupling molecular-scale research with mesoscale phenomena by virtue of its large-scale facilities for molecular research, including next-generation synchrotron radiation, neutron scattering, and supercomputing.

3. Computational Science

Physical simulation models need to be constructed at the molecular, mesoscale, and cellular levels. At each level, techniques are needed to improve sampling and simulation efficiency. Computational tools are needed to (1) understand biological and biogeochemical reactions, (2) simulate hundreds of interacting biological macromolecules on a systems level, and (3) elucidate processes at cellular interfaces. New theoretical and computational methodologies are required

to couple scales while retaining essential driving information. These methodologies will allow multiphysics descriptions of biological phenomena, leading to cellular-level simulations on timescales of up to one second, such as macromolecule and metabolite diffusion across the cell. Finally, whole-cell and colony simulations are needed to handle spatial heterogeneity and efficiently simulate timescales on the order of the cell cycle. These will require techniques for identifying individual cellular macromolecular and small-molecule species.

4. Predictive Redesign

The grand challenge in the molecular energy and environmental biosciences is to integrate the plethora of data being generated to enable a predictive understanding of complex systems at multiple scales—from the cell to organisms and communities and eventually to a global level. The coupling of high-resolution imaging, omics-driven data, and advanced computational technology will lead to the derivation of new principles of cell function that will enable the rational design of microbes and plants for producing renewable energy and ensuring the sustainability of the environment.

Cross-Cutting Themes

BER research includes disciplines that span many spatial and temporal scales—biology/physics interfaces at the subatomic level, subsurface biogeochemistry affecting contaminant behavior and nutrient cycling, sustainable biofuels development through genomics and systems biology of plants and microbes, and climate science research. Integration of these research areas will require maintaining a well-trained and highly motivated workforce and supporting the development of, and access to, advanced computational tools. Experimental research provides both targeted data for developing powerful predictive models of energy use and climate change and also large datasets that may provide drivers for molecular and cellular

research for plant and microbial systems. However, a major impediment to passing information from one discipline to its interface with another is the heterogeneity existing at every level (e.g., particle size and composition in aerosols, the variety of structures in a cell, the composition and functions of microbial communities, the geochemistry and physical properties of Earth materials, genome expression dynamics, evolutionary influences, and process dynamics). This heterogeneity complicates rational inquiry, hindering the predictability of the macroscopic consequences of microscopic behavior. Developing reliable tools to make these predictions is a major challenge. Much of this research may need to incorporate ideas from complex systems, such as principles guiding nonlinear

and nonequilibrium phenomena. In integrating across scales, there also is the classical challenge of extrapolating from laboratory results to field studies and the reverse (e.g., simultaneously accounting for the complexity and heterogeneity in the real world). Hence, cross-cutting themes point to both opportunities and obstacles for integration across scales, including spatial and temporal factors, discipline-unique assumptions, and data acquisition hurdles. Linking disciplines and scales creates new opportunities for data exchange and expanded communication via computational simulation and modeling at all scales, leading to improved understanding of processes and providing definitive information for engineering and policy decisions.

Introduction

In its *Global Risks 2014* report, the World Economic Forum (WEF; WEF 2014) lists 10 current global risks of highest concern. Of those, four are science-related and all are central to the mission scope of the Office of Biological and Environmental Research (BER) within the U.S. Department of Energy's (DOE) Office of Science: (1) water crises, (2) failure of climate change mitigation and adaptation, (3) greater incidence of extreme weather events, and (4) food crises tightly linked to the use of biomass for sustainable fuel production. According to this report, "the risks considered *high impact* and *high likelihood* are mostly environmental and economic in nature." The WEF report also notes the interdependencies between environmental and societal risks that make solutions to these stresses ever more important for global stability.

These solutions will hinge on understanding the underlying processes of entire systems—from molecular to global scales, and they will range from redesigning microbes and plants to understanding the roles of Earth's biogeochemical systems in determining climate and contaminant transport and transformation. Clearly, however, the success of these endeavors critically depends on understanding the reliable coupling between molecular and mesoscale phenomena.

Advancing efforts to determine sustainable solutions to these challenges was the impetus for the BER Advisory Committee's (BERAC) report *Grand Challenges for BER: A Long-Term Vision* (BERAC 2010). A subsequent report, *BER Virtual Laboratory: Innovative Framework for Biological and Environmental Grand Challenges* (BERAC 2013), focused and expanded on the "technology and tools most needed to support the biological and environmental research necessary" to meet the challenges and opportunities addressed in the 2010 report. BER is faced with the monumental challenge of building a foundation for a continuum of understanding across a range of spatiotemporal scales—from the molecular to the global scale and from picoseconds to millennia—that is perhaps unique in the sciences. For this work to succeed, integration of multiscale observations, experiments, computation, theory, and knowledge is essential, but much of today's research is still discipline-limited, focusing on single-scale system components with little consideration of any impact on interacting disciplines. Fragmentation hinders potential advances in constructing predictive models by limiting the synergies that come from integrating research and data across disciplines and scales.

BER convened the Molecular Science Challenges workshop to bring together scientists representing the full spectrum of BER program elements that depend on molecular science. BER research, which spans the biological, environmental, and climate sciences, has evolved over recent years to require a much more robust understanding of the molecular systems and processes that underpin program goals. The workshop was held May 27–29, 2014, in Germantown, Maryland, with 22 scientists from across the United States, representing the wide range of BER's research portfolio. These experts were given the task of assembling 10-year projections of scientific and technological challenges and opportunities in molecular science relevant to BER's mission and, once these were identified, to develop high-level progressions of scientific objectives to address these challenges and opportunities. In addition, there was a need to identify knowledge gaps that must be filled to expand the ability to understand, describe, and model molecular- to global-scale processes based on synergistic and multidisciplinary approaches. Participants also were asked to suggest what tools—not yet in existence or not yet readily accessible—could be imagined to provide data for knowledge development.

Workshop participants were assigned to one of three breakout groups: (1) atmosphere–land surface interactions, (2) near- and below-surface interactions, and (3) synthetic and genomic bioscience. Because the assembled scientists were much more disparate in their research expertise than these groups imply, extensive communication across disciplines was required for consensus formation. Early exchanges in the breakout sessions required forming concepts of the range of expertise available and of the commonly held assumptions of the decadal trajectory of the scientific fields and global events. Although the disparate disciplines and expertise of the participants led to far-ranging discussions and revealed the enormity of the scope of BER activities, the workshop did result in identifying gaps in knowledge and tools needed for molecular science research and development. The following workshop charges guided the breakout group discussions:

- Understand the molecular systems and processes that underpin BER program goals.

- Integrate across the breadth of spatial and temporal scales of BER research areas.

- Take advantage of DOE national laboratory and facility resources.

- Identify molecular science challenges and opportunities.

- Describe research pathways to overcome barriers in BER-relevant molecular science over a 10-year timeframe.

The workshop organizers guided the participants through structured working sessions to identify broad areas of opportunity for future research and capability development. These opportunities included understanding, describing, and modeling molecular- to global-scale processes based on synergistic, integrated, multidisciplinary approaches that will enable more informed policy decisions to address future needs.

The report provides individual summaries of the molecular science research needs identified by the breakout discussions for (1) atmosphere–land surface interactions, (2) near- and below-surface interactions, and (3) synthetic and genomic bioscience and concludes with a summary of the workshop's cross-cutting themes.

I. Atmosphere–Land Surface Interactions

A. Decadal Overall Vision

Energy, climate, and the environment are intimately woven together. Therefore, balancing the increasing need for energy with avoiding harm to Earth's climate and environment remains a challenge. To date, understanding of the chemical, physical, and biological factors that control molecular-level processes, which is inherently needed to understand climate and the environment, remains poor.

For atmosphere–land surface interactions, there are four areas of emphasis:

- Exchange processes: Emissions and deposition between land and atmosphere.

- Atmospheric aerosols: Links to radiative balance, cloud formation, and precipitation.

- Terrestrial ecosystem impacts: Fundamental effects on terrestrial ecosystems of the transfer of energy, water, gases, organics, and particles to and from the atmosphere.

- Technology and capabilities needed.

B. Exchange Processes: Emissions and Deposition Between Land and Atmosphere

Vision

Land-atmosphere exchanges and interactions are complex. Feedbacks between land and atmosphere affect emissions of gases and particulate matter. In turn, these gases and particles undergo atmospheric reactions that ultimately determine the abundance and efficacy of cloud condensation nuclei and ice nuclei, essential precursors to cloud formation and the hydrologic cycle. Precipitation from those clouds enables the growth of plants and microbial communities. Without understanding of these detailed processes and their

interactions, it will be exceedingly difficult to predict how future energy scenarios will impact the climate and environment. Thus, there is a critical need for validated models of the key feedbacks that govern Earth's radiation balance, cloud formation and precipitation, and plant and microbial community development and evolution to determine how energy policies affect global cycles.

Molecular-scale knowledge of the underlying biological and chemical processes provides the boundary condition to understand this interplay at the smallest scale. However, models are required that extend through the mesoscale and macroscale and that, ultimately, describe the regional- through global-scale interactions. Currently, fundamental, molecular-scale descriptions of the many processes involved cannot be integrated directly into models that describe the largest scales, or even the much more modest scales of most field studies, both because of a lack of knowledge of these processes and because of their inherently detailed complexity (see Fig 2. Solid-Fluid Interface Character Varying Over All Scales, p. 4). Models must be parameterized in ways that account for the intrinsic heterogeneity in the environment and yet retain the essential features of the processes being modeled. These models need to be validated through a combination of controlled laboratory experiments and field studies.

Laboratory studies of environmental, chemical, and biological processes reveal the dependence of the observed rates, steady states, and equilibria on local chemical, biological, and physical variables. However, conditions in the natural environment are inherently heterogeneous. Additionally, a much more diverse range of chemical and biological species are present in the natural environment than can be explored in controlled laboratory experiments, especially as the scale increases beyond that at which the underlying molecular processes control system responses. Moreover, feedbacks from processes at all scales can profoundly

influence molecular and biological processes. *Thus, fundamental knowledge of the molecular-scale processes that result from laboratory studies is essential for understanding the dynamics of chemical and biological processes in the different environmental media, but information from larger scales is needed to describe environmental responses to changes in emissions and the climate change that will result.*

One of the grand challenges for integrating molecular science in biological and environmental research is linking fundamental science into environmental models spanning from the microscale through the global scale, from scales that control underlying chemical and biological processes to those that govern environmental dynamics. Biological, chemical, environmental, and modeling research communities address different aspects of these complex systems. The ultimate development of models that span the complete range of scales will require contributions from, and collaboration among, researchers in all these different communities. Models will necessarily involve parameterizations that simplify descriptions of the complex, molecular-scale interactions.

Such approximations already are being employed to describe atmospheric chemistry, aerosol and cloud dynamics, and the hydrologic cycle on regional to global scales (see sidebar, Soil Microbial Community Models: Spanning the Scales, p. 5). Detailed kinetic mechanisms have been developed to describe the photochemical reactions of individual hydrocarbon precursors, but hundreds of molecular species are produced, many of which cannot be measured directly. Laboratory studies of selected compounds from a number of classes of reactive hydrocarbon vapors (e.g., olefins, aromatics, and terpenes) guide the development of general models of atmospheric photochemistry, but the models must address the much larger numbers and greater
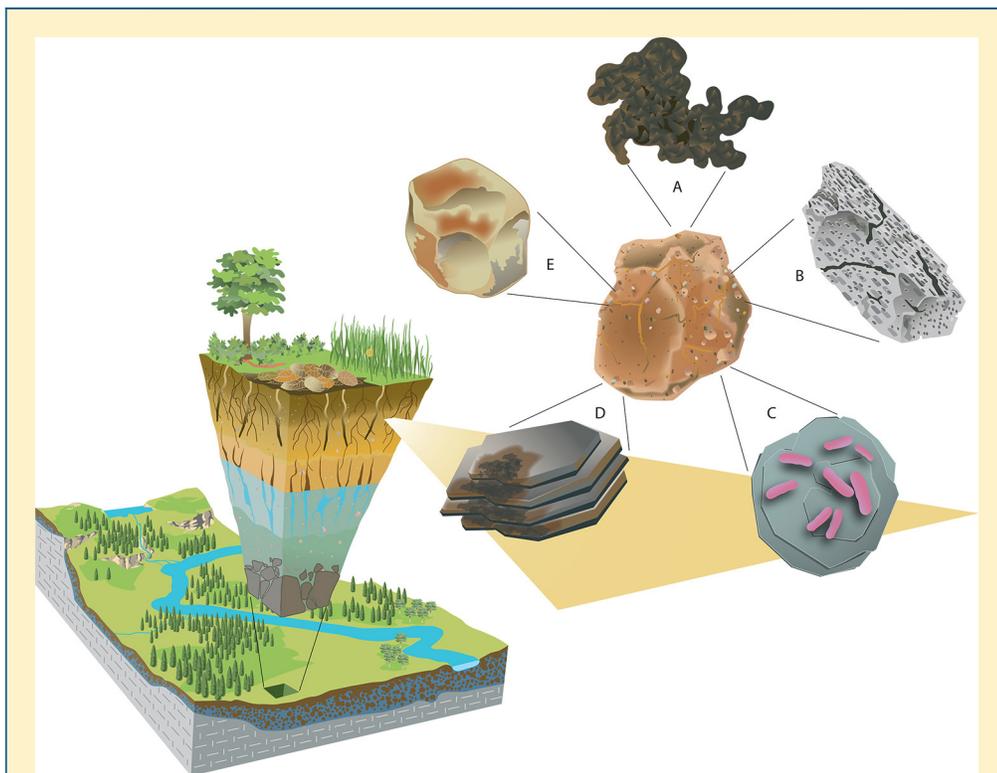
Fig. 2. Solid-Fluid Interface Character Varying Over All Scales. For a given lithology, climate, and landscape position, interface composition depends on biogeochemical conditions at the pore scale. Within a typical soil aggregate, reactive surfaces include (A) natural organic matter, (B) nanoporous silicate minerals, (C) mineral-microbe complexes, (D) secondary aluminosilicate clays and their surface organic coatings, and (E) oxide and/or carbonate coatings. [Image courtesy Chorover, J., et al. 2007. "Soil Biogeochemical Processes Within the Critical Zone," *Elements* **3**, 321–26. DOI: 10.2113/gselements.3.5.321]

Soil Microbial Community Models: Spanning the Scales

The evolution and productivity of a soil microbial community depend on water and nutrient availability that, in turn, is affected by atmospheric dynamics. Precipitation drives fluctuations in the water table that enhance carbon fixation. Cloud dynamics and precipitation intensity depend on atmospheric aerosols that may contribute a range of nutrients and toxins to the soil. Terrain and soil structure introduce local spatial variability and may enhance temporal variations.
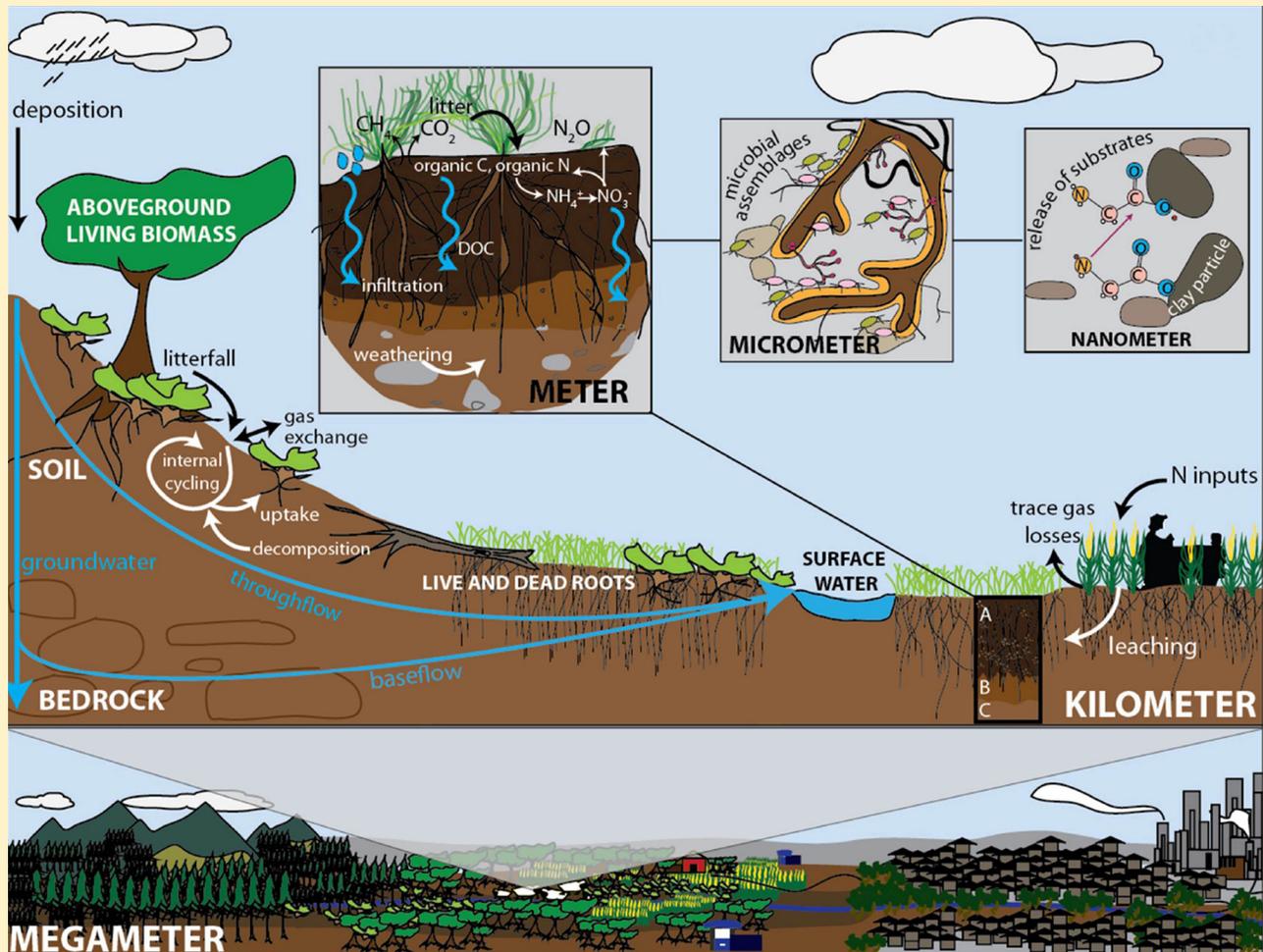
Fig. 3. Carbon (De)Stabilization in the Soil Fabric—Spanning the Scales. Incongruent geochemical weathering reactions (e.g., primary to secondary or to tertiary) occur within a milieu enriched in an active subsurface biota including roots, microbial cells, and biomolecules. These same processes affect contaminant attenuation and mobilization in the subsurface. [Image courtesy Hinckley, E.-L., et al. 2014. "Digging into the World Beneath Our Feet: Bridging Across Scales in the Age of Global Change," *Eos, Transactions American Geophysical Union* **95**(11) 96–97. DOI: 10.1002/2014EO110004]

diversity of molecules that contribute to atmospheric chemistry. Even if all the detailed kinetics were known, the computational demands of simulating such complex kinetic networks would require compromises that capture the essence of molecular-level processes while sufficiently reducing the level of detail to enable integration into global climate models. Important steps have been taken in this direction, but much remains to be done.

A case in point is the modeling of secondary organic aerosol (SOA) formation. Early models of SOA formation divided the products of reaction into two volatility classes to describe the dynamics of particle formation and growth (Odum et al. 1996, 1997; Hoffmann et al. 1997), first in chamber studies and later in regional- to global-scale atmospheric models. A recent extension of this approach, the volatility-basis-set (Donahue et al. 2006), accounts for the wide molecular weight range of atmospheric reaction products by segregating vapors into broad classes that span orders of magnitude in vapor pressure. As the range of molecular weights found in atmospheric aerosols increases, the influence of increasing viscosity on the phase, structure, and homogeneity of atmospheric particles has received increasing attention, stimulating efforts to model the influence of composition on the diffusion within particles (Koop et al. 2011). However, the phase heterogeneity of atmospheric particles remains incompletely understood.

Integration of molecular-scale models of chemical and biological processes in other environmental media will need similar parameterization efforts, although the numbers of variables required to describe these processes may dwarf those used in current atmospheric models. This parameterization effort requires fundamental data generated in laboratory and theoretical studies by chemists and biologists to guide model development. However, experiments also will be needed that span the range of molecular-scale conditions encountered in the natural environment. Measurements at environmental mean conditions are not adequate. Assuming a normal distribution of conditions, the tails

of the distribution may well dominate the dynamics of nonlinear processes or the complex dynamics of networks. Both the distribution of environmental parameters individually, with their many correlations, and the temporal variations of those distributions in the natural environment need to be understood.

Thus, although deciphering of molecular-scale processes is essential, mesoscale and macroscale decryption of the interplays among different processes is equally important to understanding the environment. The full complexity of the ambient environment cannot be reproduced in the laboratory, and ambient measurements are not likely to elucidate and quantify the links among different molecular-scale factors. Comprehensive models will need to consider exchanges among the atmosphere, soil, microbial communities, plants, surface water and groundwater, and oceans. These include emissions of gases and particles to the atmosphere from both anthropogenic and natural sources (see sidebar, Biogenic Emissions Model, p. 7).

Decadal Thrusts

- Integrate molecular-scale understanding into macroscale to global-scale modeling to elucidate impacts of energy usage and policy.

- Determine whether atmospheric concentrations of primary biological particles and biogenic vapors are predictors of an ecosystem's state.

- Elucidate the links among the atmosphere/climate and successful phenotypes/species distributions/biodiversity.

- Develop a theoretical understanding of the interconnections among atmosphere, clouds, carbon dioxide (CO_2) levels, soil, microbial communities, plants, surface water and groundwater, oceans, and the hydrologic cycle.

- Determine the mechanisms and release rates of biogenic vapors and particles to the atmosphere from soil, natural waters, plant surfaces, and other

Biogenic Emissions Model

Biogenic hydrocarbons from plants are integrated into atmospheric models through the Model of Emissions of Gases and Aerosols from Nature (MEGAN; Guenther et al. 1995, 2006, 2012; see Fig. 4. Schematic of MEGAN2.1 Model, below). MEGAN describes groundcover, with 1-km resolution globally and 30-m resolution in the contiguous United States, and recently has been extended to describe one important class of primary biological particles, pollen. Many exchanges are lacking in present models, however, including vapor emissions from soil and microbial communities. Atmospheric influences on adaptation or evolution of microbial and plant communities, as well as the resulting shifts in species distributions and biodiversity, need to be understood and modeled.
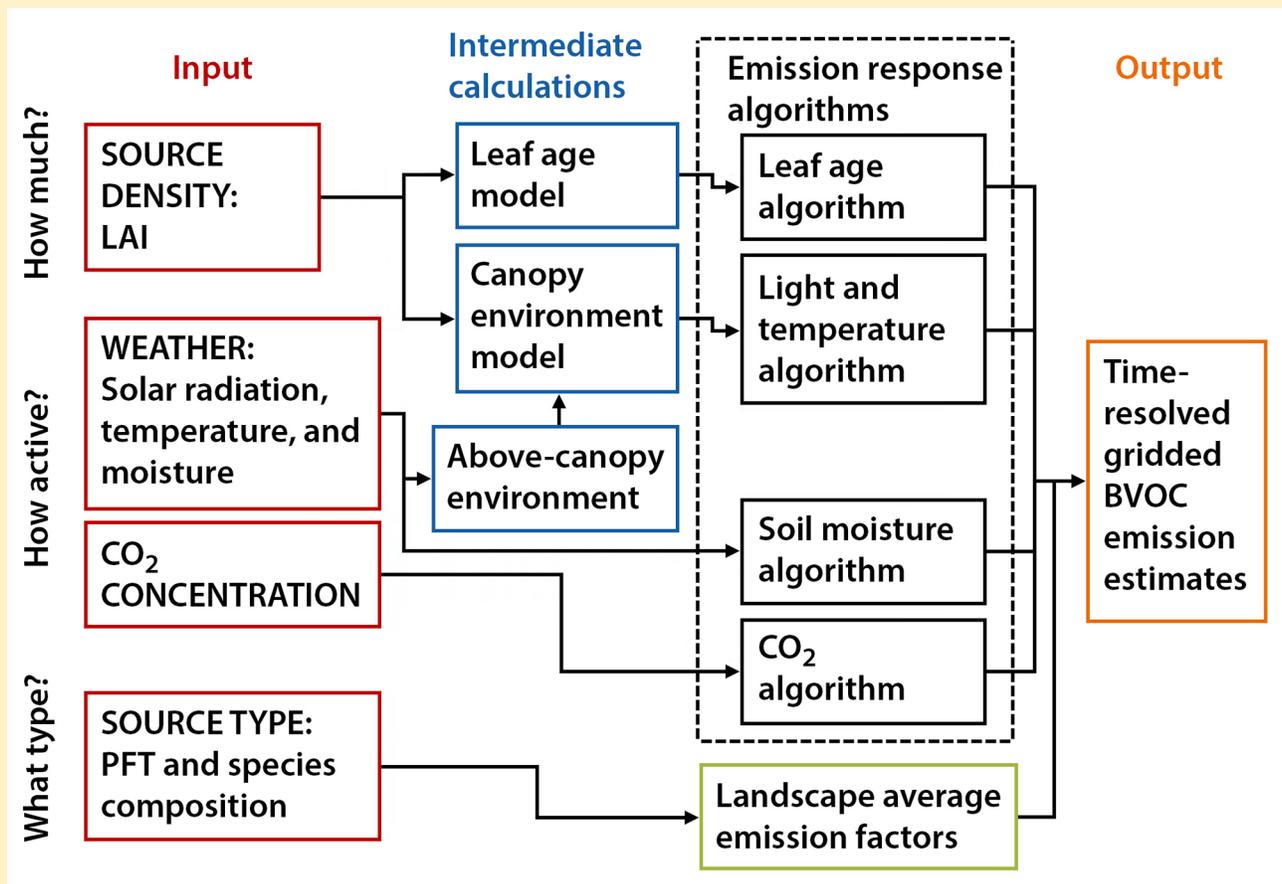
Fig. 4. Schematic of MEGAN2.1 Model. The two major model components are the algorithms describing emission response to variations in environmental conditions and the calculation of landscape average emission factors. LAI: leaf area index; CO_2: carbon dioxide; PFT: plant functional type; BVOC: biogenic volatile organic compounds. [From Guenther et al. 2012. "The Model of Emissions of Gases and Aerosols from Nature Version 2.1 (MEGAN2.1): An Extended and Updated Framework for Modeling Biogenic Emissions," *Geoscience Model Development* **5**, 1471–92. DOI: 10.5194/gmd-5-1471-2012. © Author(s) 2012. This work is distributed under the Creative Commons Attribution 3.0 License]

media through the actions of bacteria, fungi, and other communities to understand and quantify their impacts on atmospheric processes.

- Determine the total burden of primary biological particles (e.g., bacteria, virions, fungi, and spores) in the atmosphere and the extent to which they act as cloud condensation nuclei and ice nuclei.

- Develop networks of sensors to probe the different environmental media over the range of length scales and timescales that govern transport and reactions among air, soil, and water, and elucidate their effects on the development and growth of plants and microbes.

- Define how atmospheric deposition of vapors and particles affects the nutrient balance and reactivity in soils and water. Elucidate how these feedbacks influence biological activity, mineral mobilization, acidity, and so on.

- Generate comprehensive models of the interactions among the multiple environmental media and actors, intensive molecular to mesoscale measurements, and data from distributed sensor networks probing the influence of the statistical distribution of state variables (and their correlations and cross-correlations) on integrated environmental systems over the full range of scales.

C. Atmospheric Aerosols: Links to Radiative Balance, Cloud Formation, and Precipitation

Vision

Improved understanding and predictive capabilities on a global level are needed with respect to (1) cloud formation and lifetime and the resulting effects on Earth's radiation balance; (2) the effects of anthropogenic and biogenic emissions (gas and particle phases) and their interplay in the global radiation budget; (3) how energy policies affect global cycles; and (4) how changes in global climate will affect human migration,

quality of life, and whether such effects can be mitigated solely by energy policies.

Developing the capability to predict climate change will require knowledge of both aerosol- and gas-phase anthropogenic and natural sources, as well as their interplay. Further, coupling the most fundamental molecular, physical, and optical properties of individual aerosols with macroscopic measurements of cloud cover and atmospheric radiation will be essential to improve models. Incorporating these details into models will require correlating molecular-scale information with field measurements and parameterization. Several areas of particular importance that require further study are highlighted in the following sections.

C.1 Clouds and Cloud Formation

The role, reactions, rates, nucleation sites, and dynamics of water in its entirety are some of the most fundamental and critical issues in understanding the Earth system. A major challenge is to understand the formation and evolution of aerosol particles and their interactions with clouds. Under what conditions do clouds form and do aerosols affect cloud size, lifetime, and coverage? The molecular composition of aerosol particle surfaces must be known to predict the propensity for ice nucleation and the uptake of organics.

Data for the interaction of water with atmospheric aerosols and particulate matter and the effect of water on particle size, properties, and reactivity are crucial for predicting the effects of aerosols on clouds. Many parameters govern aerosol formation and aerosol physical processes, including their water and ice cloud formation potential. These parameters include vapor pressure, activity, solubility, surface tension, viscosity, density, chemical reactivity, and surface structure. Ultimately these properties are governed by molecular-scale phenomena such as interactions between solute moieties and water. Nucleation of ice from super-cooled liquid droplets in the atmosphere relies on the presence of so-called ice nuclei. Measurements indicate that these ice

nuclei exist at very small concentrations, and various types of particles have been classified according to their efficiency as ice nuclei. Predictive capabilities currently are limited by the lack of a fundamental understanding of what molecular properties create "good" liquid and ice cloud nuclei. Can molecular-level insight be obtained to improve our predictive capabilities?

Decadal Thrusts

- Develop a theoretical understanding of water interactions with different types of atmospheric aerosols, modeling the fundamental sites for water adsorption, and determine how these sites influence or control ice nucleation and cloud condensation nucleation.

- Conduct detailed surface analysis of single particles under ambient conditions of pressure, temperature, and relative humidity to examine roles of sites versus functionality.

- Determine the chemical and physical properties of aerosol particles, including their water solubility and viscosity, components, and ice nucleating sites and functional groups.

- Elucidate the effect of microscopic particle "roughness" versus particle size on ice or cloud nucleation ability.

- Determine the roles of processing and composition of particles to elucidate whether processing leads to favorable nucleation sites or to inhibition of nucleation.

- Establish the effects of organic surfactants on the ability of particles to act as cloud condensation nuclei and any other potential impacts of surfactants.

- Obtain a molecular-level understanding of the time-scales and partitioning of different aerosol components in micron-size (or smaller) droplets to provide insight into the importance of surfactant effects.

- Develop a molecular-level understanding of immersion freezing, contact ice nucleation, and deposition ice nucleation processes.

- Determine whether the drivers for ice nucleation are outliers (i.e., lie outside the normal distribution describing the ensemble of each particular aerosol type). Establish whether nucleation is stochastic in whole or in part.

- Determine the particle composition and distribution of the nuclei present at the onset of sudden nucleation events. Ascertain the role of particle types, if any, in sudden nucleation events.

- Examine the effects of molecular-level chemistry on new particle formation.

- Decipher the role and relative importance of biogenic particles in nucleation.

- Because of the profound effect of aqueous processes on the molecular composition of cycled atmospheric particulates, establish whether the high rates of these processes during repeated cycles—water condensation, aqueous photochemistry, and evaporation—affect aerosol cloud interactions.

C.2 Radiation: Particularly Light-Absorbing Aerosols

The capability to predict the optical properties of strongly light-absorbing aerosols (black and brown carbons) will require detailed knowledge of their sources (e.g., biomass versus fossil fuel combustion), composition, and morphology, as well as the effects of atmospheric cycling. Molecular-scale details govern how particles mix with other components, determine particle morphology, and control particle evolution with atmospheric processing and age. Measurements of these changes over time are needed to develop useful models of this evolving mixture. A predictive understanding of light absorption by organic aerosols (brown carbon)—minor quantities of very strong chromophores—may determine the overall absorption of organic aerosols. Hence, there is a need to know the specific molecular structures of chromophores and their formation, evolution, and disappearance in atmospheric processes. What are typical chromophores of atmospheric relevance?

Are they emitted directly to the atmosphere (primary particles) or formed in the atmosphere (secondary aerosols)? What are the precursors and mechanisms of secondary chromophore formation? Light absorption could be enhanced as a result of metal-organic complexes in mixed organic-inorganic particles, but this mechanism has not yet been studied. A framework to predict the optical properties of an ensemble mixture needs to be developed and related to field measurements to parameterize these effects to a level such that they can practically be incorporated into models. The fundamental chemical and physical natures of the different species that absorb solar radiation remain a critical uncertainty. The role of gas-phase emissions, their products, and their impacts (particle nucleation, growth, and influence on mixing state) remains an important issue.

Decadal Thrusts

- Develop methods of measuring black carbon that eliminate dependence on processing recipes.

- Elucidate a molecular-level understanding of the surface chemistry of black carbon. Determine the impact of this chemistry on the mixing (internal distribution) of "coating" materials, particularly water. What role do porous structures play? Identify chemical and photolytic aging activities.

- Characterize the contributing components for brown carbons at the molecular level to elucidate the explicit chromophores of ambient (or laboratory) brown carbon.

- Identify any trace macromolecular compounds responsible for absorption of near-ultraviolet visible solar radiation by light-absorbing aerosols (e.g., brown carbon).

- Improve the understanding and quantification of optical properties as a function of chemical composition, size distribution, and mixing state. Correlate the molecular-scale information with macroscale observables, parameterize it, and incorporate the resulting data into models.

- Perform high–time-resolution (~10 min.) field characterization of chromophores responsible for brown carbon aerosols.

- Quantify the impacts of the use of different fuels, including gasoline, diesel, and ethanol, on air quality (ground-level ozone and particulate matter) and climate through direct and indirect radiative forcing.

- Clarify which organic aerosol types increase and which decrease radiative forcing.

C.3 Secondary Organic Aerosols

Biogenic emissions are important both to SOA formation and to the interplay among biogenic and anthropogenic emissions in gas-phase and particle chemistry. A predictive understanding of the physical states of organic particles requires more in-depth information on their molecular composition. Particle composition determines viscosity, in-particle diffusion, phase separations, water uptake, and multiphase reactivity. Currently, these processes are considered to be measurable in bulk solutions or solids with minimal input from a molecular description. Ultimately, molecular speciation and functional group analysis will be required to predict these processes with an approach based on fundamental, molecular-level characteristics of organics, followed by prediction of the physical properties of particles. Fundamental information is lacking regarding the rate-limiting steps in SOA production and the effects of relative humidity and temperature on phase and particle chemistry. The integration of molecular-scale processes will require parameterization and validation with field measurements to incorporate the processes into atmospheric models. Developing climatology with the temporal and spatial variability of the chemical composition of atmospheric particles will require increased observations, particularly for the organic components that remain poorly characterized. Gas-phase field measurements of unidentified anthropogenic and biogenic volatile organic compounds and their oxidation products need in-depth study.

Decadal Thrusts

- Determine the reactions and processes responsible for SOA formation, phase, viscosity, and morphology.

- Obtain size-resolved measurement of particle-phase oligomers, carboxylic acids, organic nitrates, and other multifunctional organic molecules (aldehydes, ketones, and alcohols) to identify which particle-phase reaction mechanisms govern organic mass concentrations that are observed in field measurements.

- Identify, quantify, and parameterize the rate-limiting steps in
  - the production of SOAs (especially sulfate and nitrate)
  - thermal and photochemical reactions in the gas phase
  - transport through the gas-particle interfaces and through viscous liquid phases
  - condensed-phase chemical reactions.

- To model quantitatively and understand fundamentally the impacts of oligomer formation and decomposition on atmospheric aerosol properties, determine the formation pathways and thermochemical properties associated with organic oligomers.

- Develop new measurement techniques and strategies to characterize interfacial chemistry.

- Determine how temperature and relative humidity affect the uptake of organic compounds, adsorption, and desorption.

C.4 Particle Heterogeneity

Atmospheric aerosols are frequently heterogeneous. Aerosols may be formed under a wide range of conditions and take a variety of forms: a heterogeneous mixture of black carbons and non-combusted fuel, dust containing mineral and carbonaceous components, nonvolatile gas-phase products, mixtures of sea salt and organics, and industrial emissions. During aging, aerosols can interact via photolysis or gas-phase chemistry so that SOAs condense onto them. Chemical diversity within and among individual particles remains poorly understood (Prather, Hatch, and Grassian 2008), yet it has substantial implications for understanding of the atmospheric impacts of particles. Particle heterogeneity and its effects must be explored to understand nascent optical properties of aerosols and how aerosols evolve as a function of atmospheric processing, as well as to understand the effects of heterogeneity on water vapor uptake and subsequent cloud or ice nucleation.

Decadal Thrusts

- Determine the types of chemical reactions and interactions (influence on phase) possible between mixed inorganic and organic aerosols. Although they are treated independently in model simulations, are there important interactions that currently are neglected?

- Explore the role of particle heterogeneity in nucleation processes.

- Further analyze multiphase chemical kinetics, such as mass accommodation and reactive uptake coefficients, competitive adsorption processes (water vapor and other trace gases), and the effects of temperature and relative humidity.

- Determine the effects of changes in phase states resulting from photolytic and chemical reactions and the effects of changes in water vapor uptake on competition between surface- and volume-limited chemistry within a particle.

- Establish whether aerosol aging involves phase separation of mixed inorganic and organic aerosol components and/or separation of only organic components in aerosols.

- Measure the influence of anthropogenic and biogenic gas-phase emissions on particle heterogeneity.

- Determine the effects of non-metal components on bonding of aerosol metals and changes in oxidation states within heterogeneous particles.

- Develop rapid characterization of individual particles in the field to the molecular scale.

- Develop a method to parameterize the complexities of particle heterogeneity for incorporation into models.

- Expand field observations temporally and spatially to contribute to an understanding of the variability in the chemical composition of atmospheric particles. This information is needed to fill the gap between detailed measurements and translate the results into the generality of models.

- Increase the skilled workforce dedicated to integrating and translating data from these complex aerosol systems to generate useful predictive models.

D. Terrestrial Ecosystem Impacts

Vision

Terrestrial ecosystems are fundamentally affected by multiple exchanges with the atmosphere. The mechanisms controlling each exchange need to be coupled across multiple scales.

Terrestrial ecosystems are fundamentally connected to transfers of energy, water, and gaseous and particulate materials to and from the atmosphere. These interactions occur on many different scales, from molecules to landscapes, and have regional to global consequences in terms of climate and air quality. There are a number of well documented examples (e.g., in the tropics and the Arctic) showing that these land–atmosphere exchanges are dramatically altered by land-surface changes due to land use (see Fig. 3, p. 5), vegetation community structure and canopy phenology feedbacks with climate, and changing snow and ice cover. These land-surface changes likely will become even more intense in the future, and they may become one of the most important terrestrial–atmospheric feedbacks into climate. Although understanding of molecular-scale processes is essential, mesoscale and macroscale understanding of the interplay among different processes and sources is equally important

to understanding the environment. These include emissions of gases and particles to the atmosphere from both anthropogenic and natural sources (see Fig. 4, p. 7). Comprehensive models will need to consider exchanges among the atmosphere, soil, plants, surface water and groundwater, and oceans. Transport of mineral-rich particles and organics impacts nutrient availability in oceans and on land.

Decadal Thrusts

- Elucidate the molecular mechanisms of chemical and biological cycling of minerals and transformations of organic compounds that contain nitrogen, sulfur, phosphorus, arsenic, and other environmentally relevant elements.

- Improve the quantification of particulate matter containing redox-active compounds and metals (i.e., mercury and other metal emissions) that can be responsible for harmful effects to animals and plants.

- Determine whether atmospheric processing plays a role in making molecular forms of elements bioavailable (e.g., when aerosols deposit in nitrogen- or iron-limited ecosystems).

- Quantify more accurately the removal processes affecting aerosols. The quantification should not be limited to wet deposition but also consider heterogeneous (gas-particle) and photolytic processes, evaporative adjustments to changes in gas-phase composition, and consequences for the spatial and temporal aerosol distributions.

- Carry out direct research to elucidate the effects of changing land surfaces on energy and water transfers with the atmosphere, and how transfers are propagated to affect local, regional, and global climate systems. This research may take the form of using eddy covariance technology with flux towers to derive the landscape effects of land-use change, modeling of regional energy and water budgets, and finer-scale measurements at terrestrial surfaces.

- Explore changes in energy and water balances to determine their effects on fluxes of CO_2 and trace gases.

- Determine the impacts of changing land surfaces on the generation of dust, other types of atmospheric aerosols, and volatile organic compounds.

Figure 5. Global Models, this page, illustrates the data types and sources needed for model building.

E. Technology and Capabilities Needed

Vision

To address the molecular science questions posed in the domain of atmosphere–land surface interactions, it will be essential to establish, promote, and train a scientific community equipped with the skills to address multidisciplinary science in complex systems. To enable the establishment of this community, a number of capacities must be generated (or retained) and coordinated actions undertaken.

E.1 High-Performance Computing

Understanding complex processes at the atmosphere-land interface will require the development of sophisticated computational simulation and modeling software capable of describing complex interfacial properties. Increasing amounts of experimental and observational data are needed to provide meaningful predictions for climate change. In turn, the ever expanding data to be processed drive new computational approaches to data integration, modeling, and knowledge generation across multiple spatial and temporal dimensions. For maximal impact, the software developed must make use of high-performance computing hardware currently existing and planned within the Department of Energy (DOE) system. Moreover, making these

advances possible will require the flexible adaptation of new and existing codes to the various heterogeneous architectures planned for subexascale and exascale supercomputers.

E.2 Access to Synchrotron Light Sources

The research community looking at the surface chemistry of environmental interfaces, aerosol production, and climate impacts has broad needs in the spectral range to be addressed—from 100s of eV to >20 keV—that will require coordinated access to synchrotron facilities. The use of synchrotron light sources will be an essential research component for scientists working on minerals, soils, aerosols, and microbes. Indeed, in light of the possibility of studying ultra-dilute samples (pico- to

Fig. 5. Global Models. Contributions from all levels are needed for the generation of accurate predictive *global models* in an iterative process. *Laboratory measurements* are needed for detailed studies of aerosol composition, structure, and chemical and physical activities. *Field measurements* must be expanded in both areas and in time to provide cause-and-effect predictions. Interactions at the interfaces of *air, terrestrial surfaces, and oceans* are critical to determine exchanges and their effects. Phenomena must obey *quantum mechanical principles* and eventually be explained in those terms.

femto-molar concentrations) and small sample volumes (femtoliter and below), with elemental- and molecular-specific sample environments that will enable the investigation of many more real-world samples, this research community is expected to grow. The diversity of the research community also will place demands on light sources. Understanding the key processes described in this chapter will require even greater access to synchrotron light sources than in the past. Meeting this need will necessitate reducing the barriers to synchrotron access for these researchers via training and appropriate access models. Although specialized instrumentation may be needed, there will be cases in which the science drivers do not require leading-edge synchrotron technology, but rather increased capacity at relevant beam lines—including micro-XANES (microprobe X-ray absorption near edge structure), microdiffraction, and X-ray microscopes. To establish good physical models, standards must be established for characterizing complex samples, particularly through the analysis of bulk samples. Leaders in the field will need to work more closely with the synchrotron facilities to actively engage their colleagues. This engagement also will involve work with free-electron laser instruments at the Linac Coherent Light Source.

E.3 Stimulate Instrumentation Development

Understanding atmosphere–land surface interactions presents particular needs for instrumentation development. The scale differences in the data to be collected require the design and deployment of experiments spanning great distances. Consequently, low-cost, versatile, long-duration sensors that can operate across those distances must be developed. Associated with these studies will be the need for high-speed distributed data archives with high-speed analysis in place.

E.4 Establish Multidisciplinary Training Schemes

Meeting the challenges of the research topics in this field will require the formation, training, and maintenance of a research community. Multidisciplinary research

schools based on the European and Asian models (e.g., the European Research Course on Atmospheres sponsored by the European Commission in Grenoble, France) should be established. Students should be trained to develop hypotheses across temporal and spatial scales and to integrate and translate molecular-level information into models. DOE is in a position to play a leadership role by developing funding mechanisms that address the need for more effective student training at DOE national laboratories and user facilities.

Decadal Thrusts

- Create distributed networks that can provide large datasets over long time periods. These might take the form of advanced sensor networks that can be dynamically changed to different resolutions to detect different physical and chemical phenomena.

- Increase the trained workforce with the right types of education and personal skills to integrate and translate data among different domains and to aid with the interfaces between different sciences (including mathematics and computer sciences).

- Design a mechanism to spark important research questions. For example, a "collaboratory" could bring all the sciences together much like the national laboratories do. One can imagine a sensor network detecting an unusual signal, triggering a focusing event that picks up a new chemical component that was not in the chemical models. A sensitivity analysis is then performed, suggesting additional fundamental experiments and computations that need to be completed. Once these are completed, the chemical models are modified and fed into the larger-scale models to understand the unusual event. Multiple research areas could be of interest, including aerosol formation (where there are some social mechanisms in place for existing teams, but new researchers have difficulty integrating) and SOAs over cities, forests, and prairies.

II. Near- and Below-Surface Biogeochemical Interactions

A primary research objective of the U.S. Department of Energy's (DOE) Office of Biological and Environmental Research (BER) is to obtain predictive knowledge of terrestrial ecosystems extending from bedrock to treetops (i.e., the "critical zone") and from global to molecular scales. Climate warming, increasing atmospheric carbon dioxide (CO_2) concentrations, the resulting perturbations of precipitation timing and amounts, and the impacts of all these changes on nutrient cycling are driving ecosystem changes, which all have important ramifications for biological productivity, contaminant behavior, and land use. A robust understanding of ecosystem responses to environmental change is essential to improve projections of climatically induced effects and feedbacks between the terrestrial ecosystem and Earth's atmosphere and hydrosphere. To advance our global understanding of ecosystem function, biological and geochemical (biogeochemical) processes must be studied in space and time in diverse terrestrial settings. BER has long had an interest in biogeochemical processes in natural environments, from the near surface occupied by plants and other photosynthetic organisms down to subsurface aquifers. In the future, understanding of processes that determine ecosystem responses to changing environmental conditions; insight into their impacts on plants, microbes, biogeochemical cycles, and contaminant mobility; and an ability to detect and understand these changes will be critical to enabling a predictive capability.

Knowledge of molecular-scale processes is profoundly important to achieving this BER vision. Near- and below-surface biogeochemistry is driven by reactions occurring at the molecular scale on enzymes, cell membranes, organic substrates, and mineral surfaces.

Moreover, biogeochemical processes are initiated and controlled at the molecular scale by the expression of microbial and plant genes. These biotic processes are complemented and influenced by molecular chemistry occurring within the interstitial spaces of soil or sediment particles, including interfaces between water and minerals, water and bacteria, bacteria and minerals, and plant cells and microorganisms. These physical spaces; their arrangements within ecosystems; and their microbial, mineralogical, and chemical compositions are heterogeneous at length scales of nanometers to meters. This heterogeneity profoundly influences molecular-scale processes and ecosystem responses to environmental perturbations. The ability to predict biogeochemical processes at the ecosystem scale rests upon foundational knowledge of these molecular-scale processes; their impacts on electron, nutrient, and contaminant fluxes; their locations within the terrestrial environment; and their projection to the ecosystem scale in space and time.

Because molecular-scale biogeochemical processes drive ecosystem behavior across scales, they must be considered in this larger context. Consequently, the subjects discussed here are diverse and cut across ecosystem scales, link subenvironments, and require new scientific methodologies and approaches. "Subenvironments" and "subsystems" are used interchangeably in this chapter to refer to discrete portions of ecosystems that are identifiable by location, saturation state, redox status, and other soil or sediment characteristics (e.g., grain size, mineralogy, texture, permeability, organic content, and depositional environment). Subsystems may be microns or hundreds of meters in dimension and include pore spaces, soil (micro) aggregates, the root-soil interface, soil horizons, and hyporheic zones.

Obtaining molecular-scale information directly within ecosystems is difficult because of their chemical, physical, and temporal complexity. For example, subsystems exhibiting intense biogeochemical reactivity ("hot spots") may have small dimensions (microns to centimeters) and may be active only intermittently (in "hot moments") during climatic events such as rainfall or meltwater runoff. This intermittency challenges their detection and characterization. Even larger (meterscale) eco-subsystems, such as thick organic-rich sediment lenses, are difficult to observe if they are buried within the subsurface. Often, the distribution of such structures is not known. This complexity presents major challenges to predicting the impacts of important biogeochemical hot spot and hot moment activity on an ecosystem trajectory under changing climate or land-use conditions.

Macroscale interactions between ecosystem components (e.g., between plants and microbes or plants and minerals) and between ecosystem subenvironments, (e.g., along groundwater–surface water interfaces) are critical to understanding the complexity of natural ecosystems and predicting their responses to environmental change. Interfaces between saturated and unsaturated sediments, or organic-poor and organic-rich zones, create biogeochemical gradients in which intense redox cycling of carbon, nutrients, and contaminants occurs. Although these interfacial zones are relatively thin (often microns to centimeters), a large fraction of the molecular-scale biogeochemical activity may reside there, and overall reaction rates may be controlled significantly by the physical and biogeochemical heterogeneity of gradients.

The spatial arrangement and connectivity of ecosystem components, critical to ecosystem behavior, are generally site-specific. Consequently, field-based studies are critical. However, resolving fundamental biogeochemical drivers operating in the field often is not possible without controlled laboratory research on well-constructed model systems. Studies that link mechanistic breakthroughs in the laboratory to *in situ*

field processes are critical. Complementary statistical and modeling studies that shed light on field and laboratory experiments are essential to successfully scale from molecules to ecosystems.

Biogeochemical studies to date have highlighted the importance of molecular-scale and microscale ecosystem components—such as mineral surfaces, enzymes, and bacterial cells—as drivers of chemical reactions in near-surface and subsurface environments (Borch et al. 2010; Williams et al. 2013; DiChristina, Fredrickson, and Zachara 2005; Paerl and Pinckney 1996; Gilbert and Banfield 2005; Brown and Calas 2011, 2012, and 2013). Studies also have identified ecosystem subenvironments as nexuses of biogeochemical activity that produce and consume nutrients and greenhouse gases and modify the chemical forms of contaminants (Vidon et al. 2010; McClain et al. 2003; Laanbroek 2010). However, important details necessary to construct predictive models of these active regions are not known, including their spatial distributions and temporal variability. Field-relevant reaction rates frequently cannot be determined; in some cases, key reactions may be missing from reaction networks. In the case of metal contaminants, their molecular-scale chemical and physical forms, along with their properties in the field, are poorly understood. Knowledge of nutrient and contaminant fluxes among ecosystem subenvironments, critical for constructing ecosystem models, is often not available. The environmental roles of many abundant microbes and proteins that serve as catalysts for biogeochemical reactions are unknown. Integrating subsystem temporal and spatial behaviors into ecosystem models is necessary for accurate predictions of contaminant mobility, system tipping points, and biogeochemical transitions. The consequences of changes in subsystem behaviors might include ecosystem succession resulting in potentially novel functional states. Collecting this information at the molecular scale while generating knowledge about processes impacted at the ecosystem scale is a grand challenge.

A. Decadal Overall Vision

Our goal is to understand the interdependencies of biogeochemical processes and how these processes quantitatively scale from the molecular to ecosystem level. This understanding can then be used to predict ecosystem response to changing climate, tipping points, and land-use patterns. Achieving this vision will require studying genes, plants, microorganisms, enzymes, pore networks within sediments and soils, minerals, groundwater, and the interfaces between these components (see sidebar, Molecular-Scale Biological and Biogeochemical Processes, p. 18). Length scales ranging from Ångstroms (10^{-10} m) to meters must be considered. Moreover, the *temporal* behavior of processes and its impacts on larger-scale components must be considered, which requires crossing scales from nanoseconds to decades.

Investments in the following scientific areas have the potential to advance BER toward this vision:

- Link genome composition of microorganisms and plants with phenotypic traits at cellular, organismal, and community levels.

- Determine plant-microbe interaction mechanisms.

- Elucidate biogeochemical cycling of carbon, nutrients, and contaminants.

- Link subsystems and processes across scales to describe ecosystem behavior.

Addressing these challenges will require comprehensive and creative approaches that leverage existing knowledge and technologies in new ways, such as combining approaches involving nano-SIMS (nanometer-scale secondary ion mass spectrometry), X-ray nanoprobe, fluorescence *in situ* hybridization (FISH), and Chip-SIP (stable isotope probing performed on a phylogenetic microarray). Existing analytical capabilities will need to be significantly improved or adapted from other fields. For example, detection limits and spatial resolution for synchrotron-based

X-ray techniques and mass spectroscopy need to be advanced by orders of magnitude. "Omics" approaches must be applied to plants and scaled for high-throughput application; the resulting datasets must then be reduced to tractable algorithms that can be visualized or incorporated into models. Molecular and geochemical modeling approaches will need to be significantly advanced. Beyond this, entirely new techniques and approaches need to be developed to couple, for example, plant or microbial genotype directly with phenotype or function in the environment and detect and quantify metabolites in living plant or microbial tissues in real time without pretreatment or amendment.

B. Linking Plant and Microbial Genotypes to Phenotypes Across Scales

Vision

Through advances in DNA sequencing technologies, we can determine, at nucleotide resolution, the complete genotype of any organism. These advances are ushering in a new era of genomic science in which the goal turns from defining an organism's genotype (i.e., its genetic code) to developing the knowledgebase required to predict phenotype (i.e., its physical characteristics) from genotype (Jelier et al. 2011; Lehner 2013; Gottlieb, Beitel, and Trifiro 2014; Yvert 2014). A realization of these relationships ultimately will enable accurate predictions ranging from knowing which genotypes of a given energy crop species are likely to perform best in a water- or nitrogen-limited environment to predicting ecosystem flora and fauna responses to climate and land-use change.

Predicting phenotype from genotype will require measuring a wide range of biological parameters— molecular, physiological, and structural—in populations of a given organism that differ in genotype and developing computational approaches enabling such predictions. These efforts will require applying the full array of current and developing omics

Molecular-Scale Biological and Biogeochemical Processes

Molecular-scale biological and biogeochemical processes control nutrient cycling, contaminant mobility, and gas fluxes at the ecosystem scale in near- and below-surface environments. Understanding how these processes interact across space and time scales will aid prediction of ecosystem response to a changing climate. Illustrated and described here are key components of these interdependent processes that require better understanding to enable such predictions (see Fig. 6. Molecular-Scale Processes and Biogeochemical Hot Spots, p. 19).

Organic-rich sediment layers (upper left quadrant) often contain abundant fine-grained and redox-active minerals and may host a large fraction of the biogeochemical activity in a subsurface environment. These biogeochemical hot spots export organic carbon as carbon dioxide (CO_2), methane (CH_4), and dissolved organic matter to other portions of the surrounding ecosystem. Other important chemical exports include dissolved reduced inorganic species such as hydrogen sulfide (HS^-), Fe^{2+}, and nitrous oxide (N_2O). Organic-rich sediments release and acquire metal contaminants such as uranium and are believed to be important controls over contaminant mobility.

Reactive nanominerals, microbes, organic molecules, and water reside in the pore spaces between sediment grains (middle left). Mineral grain surfaces often are coated with these reactive components. Consequently, the molecular-scale reactions that drive ecosystem biogeochemistry—and the processes by which they couple in space and time—occur at the pore scale.

Plant-subsurface interfaces also are biogeochemical hot spots (upper right quadrant). Plants moderate the subsurface-atmosphere interface by allocating organic carbon or oxygen to below-surface environments, facilitating gas transport to the atmosphere and hosting large and diverse microbial communities. Microbial communities, in turn, benefit the plant host through nutrient acquisition, disease suppression, and the modulation of host immunity.

Microbial activity is fueled by natural organic matter (NOM), which is degraded by enzymes via hydrolytic attack, depolymerization, and oxidation reactions (lower right quadrant). Association of organic matter with reactive minerals such as iron oxides can both accelerate and inhibit organic matter degradation. Different types of microbial metabolisms act in concert to drive these processes, producing CO_2 and other greenhouse gases that can be released to the atmosphere. This microbially driven activity profoundly alters the surrounding geochemical environment.

In **microbial respiration** (lower left quadrant), electrons are transferred to redox-active minerals such as iron and manganese oxides; metal contaminants such as uranium and chromium; and nitrogen- and sulfur-containing compounds such as nitrate, nitrite, sulfate, and elemental sulfur. Organic molecules shuttle electrons and indirectly facilitate electron transfer by modifying redox potentials of metals and stabilizing chemical-reducing conditions. Changes in the oxidation states of metals dramatically modify their speciation, mobility, and bioavailability. For example, U^{6+} adsorbed to surfaces of natural iron oxides may be reduced to U^{4+}, a much less mobile and therefore more desirable form. Electron transfer to solid-phase iron oxides causes them to recrystallize or to dissolve. If the latter occurs, then iron becomes mobile as Fe^{2+}; nutrients and contaminants that often are found adsorbed to iron oxides, such as phosphate and uranium, also become mobilized.

Plant types and communities are distributed at the land surface in response to soil type and nutrient, water, and sunlight availability (top). Biogeochemical hot spots often reside along the margins of stream and river corridors where nutrients and organic matter are relatively abundant and microbial activity drives redox cycling of carbon, nitrogen, iron, sulfur, iron, and other metal(-loid)s.
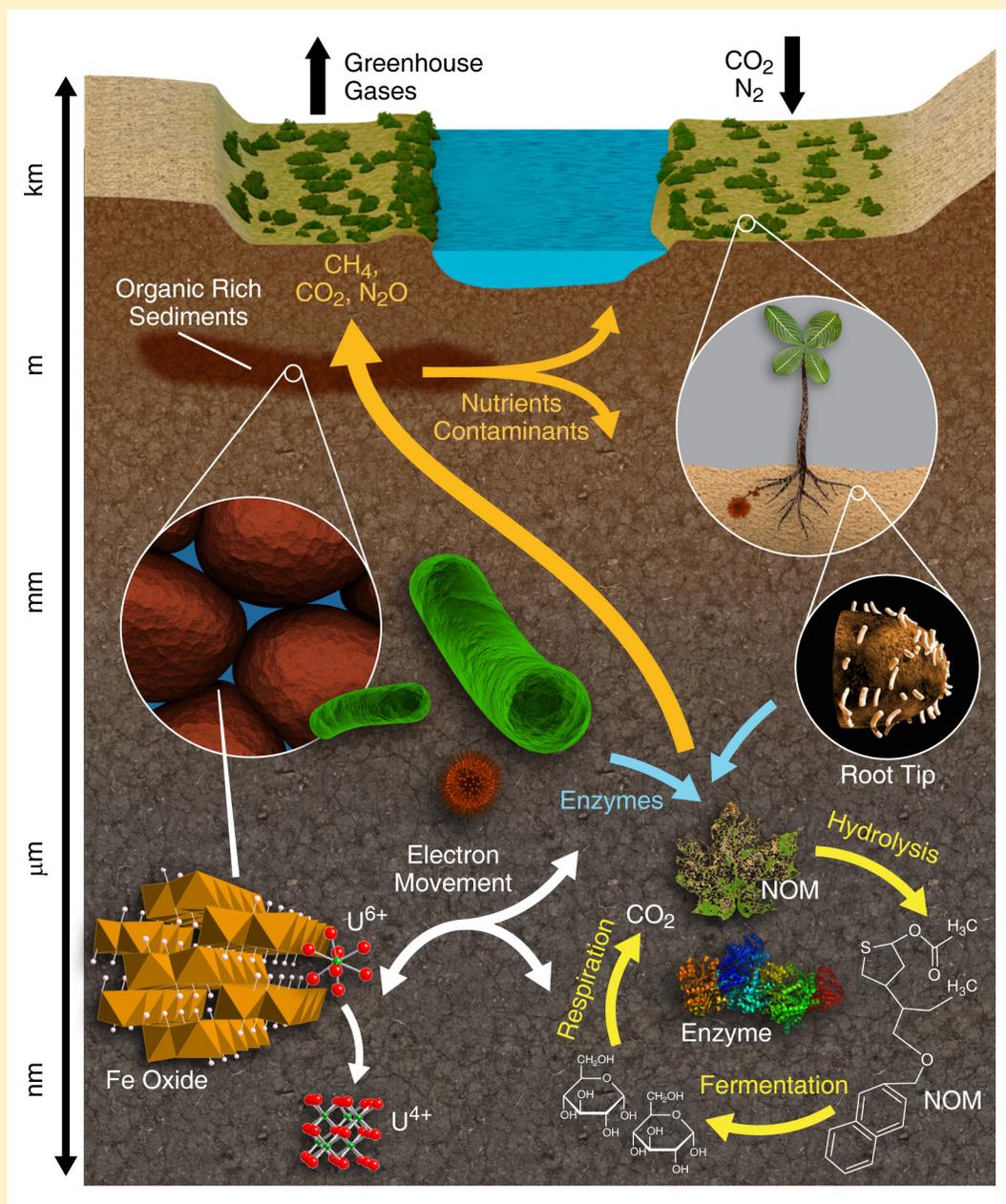
Fig. 6. Molecular-Scale Processes and Biogeochemical Hot Spots. White, yellow, and blue arrows correspond to pore-scale processes. White arrows highlight transfer of electrons between donors and acceptors; yellow corresponds to major microbial metabolic processes and their chemical products; and light blue arrows indicate microbial production of the enzymes that drive these processes. Light orange arrows indicate fluxes of biogeochemical products from hot spots across scales and time to other parts of the ecosystem. [Image courtesy John Bargar, SLAC National Accelerator Laboratory]

technologies to probe the molecular responses of microbial or plant species to dynamic environments. However, current omics technologies and methods of physiological and structural testing, although powerful, are limited when the phenotypes of hundreds to thousands of genetic variants of plants are tested in dynamic environmental conditions. Taking full advantage of genetic diversity would be greatly enhanced by the development of sensitive, high-throughput, and noninvasive technologies that probe an array of phenotypes. These could range from imaging growth and development to monitoring the levels of key metabolites in populations of an organism that differ in genotype, whether they be natural genetic variation or mutants developed in the laboratory. In addition, quantifying the selective pressures that the many relevant environmental factors impose on specific genotypes requires integrating biological and physicochemical data into a comprehensive model.

One particular area in which greater insights are needed is fungal communities because of their centrality to the global carbon cycle and to enzymatic deconstruction of biomass in bioenergy applications. BER programs could benefit from an expanded knowledgebase for fungal genotypes and phenotypes. In the broader sense, discrepancies in the annotations of microbial genomes also limit the understanding of genotype–phenotype linkages that are key to the functioning of ecosystems of concern to DOE.

Developing the instrumentation and technical tools, algorithms, genetic populations, and databases required to build predictive "genotype-to-phenotype" relationships is important and timely for BER investment. Success in these areas will lead to a knowledgebase to link phenotype with genotype in dynamic environments. This grand challenge relates to core BER programmatic goals in energy and builds on BER areas of strength, including data capture and analysis, systems modeling, computation, and facilities and technology development.

Decadal Thrusts

- Develop high-throughput, noninvasive technologies to monitor phenotypes ranging from the molecular (e.g., abundance of specific metabolites and proteins) to whole-organism level (e.g., photosynthetic efficiency) in near real-time.

- Develop the capacity to screen the transcriptomes and metabolomes of microbial communities at high throughput under as close to *in situ* conditions as possible.

- Incorporate probabilistic approaches into modeling genotype-phenotype relationships.

- Expand systems genetics reference populations.

- Improve gene annotation for both microbial and plant genomes.

- Construct a database of understudied fungal genomes and transcriptomes.

- Develop a hierarchy of models—including constraint-based metabolic models of individual microorganisms and predictive metabolome models of microbial communities—to link genotype to phenotype from the molecular to ecosystem scale.

C. Plant-Microbe Interactions

Vision

Plants and their associated microbiomes play central roles in environmental sustainability and serve as biomass feedstocks for biofuels and biomaterials. Plants impact greenhouse gas emissions directly through photosynthesis and respiration and indirectly by acting as sources of organic matter or nutrients for microbes that produce and release greenhouse gases. Hence, the latest climate models for terrestrial ecosystems are beginning to incorporate plant phenotypes and functional types as well as functional guilds of microorganisms. In the context of bioenergy and climate applications, an improved mechanistic understanding of plant-microbe interactions is needed that is site- and organism-specific for climatically sensitive ecosystems.

Plant microbiome research is in its infancy, lagging at least a decade behind research into the human microbiome. Microbes (bacteria, archaea, and fungi) interact with plants to complement plant functional traits and primary productivity. Tens of thousands of microbial species associate with plants, and plant-microbe interactions are crucial to plant health. Broadly defined, plant-microbe associations include the microbes that live inside plant tissues (endophytes) as well as those that grow on the plant surface (see Fig. 7. General Overview of Interactions Between Plants and Microbes, this page). Microbes demonstrate the potential to benefit plants in nutrient acquisition, disease suppression, and the modulation of host immunity (Berendsen, Pieterse, and Bakker 2012). Next-generation sequencing technologies

Induced systemic resistance
Systemic acquired resistance

Epiphytic (phyloplane) microbes

Leaf pathogens
(e.g., *Alternaria brassicicola*)

Carbon root exudates, attracting heterotrophic microbes

Root pathogens
(e.g., *Fusarium oxysporum*)

Plant growth–promoting rhizobacteria and fungi

Microbe-induced priming

Mycorrhizae

Endocytosis of microbes via root hairs

N-fixing bacteria

Endophytes

Uptake of organic carbon, nitrogen, and phosphorus from soil organic matter (e.g., from lyzed microbial cells)

Increased or reduced nutrient availability (e.g., P-solubilizing microbes, siderophores)

Secreted enzymes, competing for nutrients

Fig. 7. General Overview of Interactions Between Plants and Microbes. Environmental transcriptomics is likely to reveal key plant and microbial genes that play important roles during these and other unknown interactions. [Reprinted with permission from Elsevier: Schenk, P. M., et al. "Unraveling Plant-Microbe Interactions: Can Multi-Species Transcriptomics Help?" *Trends in Biotechnology* **30**(3), 177–184. © 2012]

and omics-based approaches have recently enabled some of the first detailed studies of plant microbiomes (Mendes et al. 2011; Lundberg et al. 2012). Metabolic models for microbial symbionts of plants are beginning to be integrated into photosynthetic models for DOE-relevant plants (Weston et al. 2014). However, these studies have been hampered by methodological limitations and applied to relatively few model plant species. Thus, to achieve a predictive, systems-level understanding of the role of plant function in managed or natural ecosystems, an improved mechanistic understanding of plant-microbe associations is required.

In contrast to planktonic environments in which the predominant and active microbial groups have largely been identified, the ecology of plant-associated microorganisms (including their abundance, distribution, diversity, and functioning) remains little studied. To achieve a mechanistic understanding of plant-microbe interactions that contribute to BER programmatic goals over the next 10 years, the following key questions must be addressed:

1. What microbial groups are associated with the tissues of DOE-relevant plants, and where in the plant do these microbial groups reside?

2. How does a plant benefit from its microbiome?

3. How do plant genotype, microbiome, and environment (i.e., climate and soil type) interact to regulate plant growth promotion?

4. What are the signaling or communication pathways between microbiome and plant?

5. Are communication pathways dependent on plant genotype, developmental stage, soil conditions, or attack by insects or microbial pathogens?

6. How do various soil types affect the populations and community structures of microbes residing therein?

7. How can plant-microbe interactions be modeled and scaled to inform ecosystem, landscape, and land-atmosphere models?

To answer these questions, a more comprehensive modeling framework is needed to formalize the understanding and misunderstanding of biology-environment interactions. On the microbial side, the basic ecology of plant-associated microbes must be interrogated first. Subsequently, plant function and the impact of microbes on function must be incorporated into plant models. A hierarchical framework approach could be used to link plant cell and microbial models and then increase the scale to link plant organs and microbial biofilms. Ideally, this framework eventually could enable the full coupling of whole-plant models; microbiome models; and environmental factors such as temperature, water availability, and ultimately the soil habitat, which accounts for site-specific differences based on the resident microbial population with which plants interact.

The ultimate goal will be to identify components of the plant-microbe metabolome and then to link metabolome models to biogeochemical and Earth system models. To build such an understanding, experiments and observations will need to be conducted at high resolution and at high throughput, incorporating experimental molecular and computational tools. The community ecology of plant microbiomes can be elucidated through the use of existing or improved high-throughput omics technologies. With metagenomes and metatranscriptomes as roadmaps, new approaches need to be developed for profiling the metabolomes of plant microbiomes at high resolution in various plant tissues. New technologies are required that can detect microbial and plant metabolites and associate these with specific microbiome cells or populations in living tissues in real time. In addition, high-sensitivity spectroscopic and spectrometric methods and isotope imaging techniques should be improved to determine the concentrations of reactants and products of

metabolic reactions and to locate these metabolites within the plant, microbe, and soil environment with nanometer-scale resolution.

Directly coupling plant or microbial genotype with phenotype or function in the environment remains one of the largest hurdles in environmental science. Existing nano-SIMS and X-ray spectromicroscopy imaging technology can be used to determine reaction rates of carbon and nutrient transformation for microbial cells living inside plant tissues. Coupling of *in situ* hybridization with stable isotope labeling can directly link the phylogenetic populations of plant-associated microbes to biogeochemical function. For example, rates of nitrogen or carbon transformation by microbial cells embedded in plant tissues can be calculated under close to *in situ* conditions. However, existing methods that directly link microbial genotype with phenotype (such as nano-SIMS coupled with hybridization probes) require extensive sample pretreatment, enrichment with an isotopically labeled substrate, and bombardment with secondary ions in a mass spectrometer. For heterotrophic processes (such as polymer hydrolysis) that are catalyzed by extracellular enzymes, techniques that require amendment with a labeled substrate are not effective. Therefore, new tools or approaches are needed to simultaneously characterize the phylogenetic identity of a microbial cell while quantifying its function under *in situ* conditions in living tissue without pretreatment or labeled substrate additions.

Such analyses conducted over numerous sites with sufficiently detailed sampling to be meaningful will generate massive datasets. Terabytes of omics and bio-geochemical data will need to be reduced to tractable equations that can ultimately be incorporated (directly or indirectly) into a modeling framework—preferably one that can be connected to other modeling communities such as Earth system science and hydrology. This connection between empirical plant-microbe experiments and predictive modeling is an important disciplinary gap to be bridged because it is limiting

progress. Although both microbial ecology and computational biology are exploding fields, further integration of these disciplines is needed.

Decadal Thrusts

- Characterize microbial groups associated with the tissues of DOE-relevant plants and determine where in the plant these microbial groups reside as a function of climate and soil type.

- Quantify the genotype-phenotype linkages between plants and microbes, as well as between various microbial groups that reside in or on plants.

- Elucidate mechanisms by which plants benefit from microbiomes.

- Develop conceptual models to describe the interactions between plant genotype, microbiome, and environment (i.e., climate and soil type) that regulate plant growth promotion.

- Characterize signaling or communication pathways between microbiomes and plants. Are these pathways dependent on plant or microbial genotype, developmental stage, soil conditions, or attack by insects or microbial pathogens?

- Develop the capability to detect and quantify plant or microbial metabolites in living tissues in real time without pretreatment or amendment as a top technological objective. These measurements should include the reactants and products of assimilatory and dissimilatory metabolism linked with specific autotrophic and heterotrophic microbes.

- Develop analytical capabilities to scale plant-microbe interactions from molecular to ecosystem, landscape, and land-atmosphere levels.

D. Biogeochemical Cycling of Carbon, Nutrients, and Contaminants

Vision

A decadal challenge for BER is to develop theories to accurately predict nanopore effects on the properties of confined water and reactivity of surface complexes and solutes within nanopores. Mechanistic knowledge of molecular-scale processes occurring in pores also needs to be incorporated into ecosystem-scale models. Such advancements would support sustainable management of the subsurface for biofuel production, CO_2 sequestration, and remediation of contaminated groundwater. In addition, because atmospheric aerosols are one of the largest sources of uncertainty in climate modeling, a better understanding of mineral dusts that significantly contribute to these aerosols could improve climate predictions (Choobari, Zawar-Reza, and Sturman 2014).

The biogeochemical engines (microbes and enzymes) and associated chemical machinery (mineral and microbe surfaces, dissolved species, and water) that carry out critical ecosystem processes reside within the interstitial spaces among soil particles or sediment grains. The processes affected include microbial degradation of organic matter; microbially mediated sulfur, nitrogen, and metal redox cycling; contaminant transport and transformation; mineral dissolution and precipitation; chemical buffering; and filtration. Coatings on mineral grains contain highly reactive oxides, sulfides, and carbonate minerals, as well as biological materials (e.g., polysaccharides, lipids, proteins, and DNA) that may enhance or inhibit these processes. Biogeochemical processes are complex, with numerous feedbacks. For example, soil organic matter (SOM) is critical in water and nutrient retention (i.e., soil quality), and SOM is derived from biological activity that depends on water and nutrient availability.

D.1 Organo-Metallic Interactions

Natural organic matter (NOM) is partially degraded detrital plant, animal, and microbial biomass and their decomposition products such as recalcitrant macromolecules and low-molecular-weight fatty acids. NOM is the dominant fuel for microbial metabolism in below-surface environments and interacts strongly with metal ions. These chemical interactions profoundly influence the biogeochemical cycling of major and minor elements (e.g., carbon, nitrogen, and sulfur), biologically essential metals (e.g., iron, manganese, zinc, and molybdenum), and metal(-loid) contaminants (e.g., chromium, arsenic, selenium, mercury, and uranium) and also affect the production of greenhouse gases such as CO_2, nitrous oxide (N_2O), and methane (CH_4).

Goal: Develop quantitative molecular-scale mechanistic understanding of NOM-metal interactions. In particular, determine how organic carbon degradation couples to metal redox cycling in below-surface environments.

Metal ions influence organic matter decomposition and fate by forming metal-NOM complexes in near-surface and subsurface environments. The association of NOM with metal oxides and other mineral surfaces is believed to confer resistance against microbial degradation (Eusterhues et al. 2014). Iron- and manganese-respiring bacteria oxidize organic carbon to CO_2. For example, abundant Fe(III) favors the growth of metal-reducing bacteria in Arctic coastal plain soils (Lipson et al. 2013; 2012), providing a path for CO_2 production. Changes in climate will drive changes in the saturation state; the redox conditions of these soils; and, consequently, the molecular-scale chemical and physical forms of iron and iron-NOM complexes. These changes, in turn, are expected to perturb the biogeochemical iron cycle, causing direct impacts on greenhouse gas emissions. The converse is also true: NOM decomposition profoundly influences metal behavior. NOM can directly mediate metal reactivity through the formation of metal-organic complexes that

increase metal solubility and mobility and modify the energetics of metal redox couples. Moreover, soluble organic molecules can act as electron shuttles, facilitating metal redox transformations.

The mechanistic linkages between organic matter decomposition and contaminant metal redox transitions are poorly understood. Under anaerobic conditions, a consortium of syntrophic organisms is required to degrade particulate organic carbon, specifically fermentative bacteria and dissimilatory respirers. Fermenters, iron reducers, and sulfate-reducing bacteria are generally expected to be present. If uranium is the contaminant, all members of the consortium are potentially capable of driving metal reduction. Establishing a physical model that links syntrophic carbon cycling to uranium reduction requires linking the locations and identities of bacteria to the location of uranium, its quantity and oxidation states, and carbon and nitrogen utilization at the pore scale.

Advancing the state of knowledge will require addressing several major challenges:

1. **Determine the detailed structures of natural organic molecules, at the molecular scale, that occur in near- and below-surface environments.** NOM exhibits a high diversity of molecular weights, compositions, and chemical properties. Moreover, individual pore water samples containing NOM will consist of thousands of different molecules, according to recent high-resolution mass spectroscopic analysis. New tools are needed to simultaneously characterize the molecular structure and composition of individual organic molecules in pore waters at high throughput (a capability currently not possible). Dealing with the abundance of structures requires developing high-throughput technologies that can select and purify organic molecules on the basis of size and composition for subsequent characterization. Alternatively, techniques with low detection limits that can distinguish molecular mass need to be combined with those that determine molecular structure. Utilizing the large amount of information obtained from these approaches will require developing spectrum and chemometric approaches that can simplify, visualize, and manipulate large populations of molecules.

2. **Determine the proton- and metal-binding affinities of functional groups present in natural organic molecules.** Accomplishing this goal will first require detailed structural models provided by developments outlined in the preceding paragraph. Developing quantitative structure-activity relationships probably will be important to this effort.

3. **Determine the molecular structures and chemical properties of metal-NOM complexes, such as Fe(III)-NOM.** Accomplishing this goal will first require detailed knowledge of the structures of NOM molecules being used.

4. **Determine the detailed structural mechanisms—down to the atomic scale—by which natural organic molecules coordinate with metal ions and mineral surfaces over the range of chemical variables.** This is a challenging objective because of the difficulty in imaging the noncrystalline surface structure of nanoparticles. Major breakthroughs will be achieved when lensless imaging techniques can obtain atomic-scale resolution. Until that time, progress can be made using carefully selected model studies and extending surface-sensitive X-ray techniques with vibrational spectroscopy, nuclear magnetic resonance, and density functional theory numerical approaches.

5. **Develop the capability to image, simultaneously, the oxidation states and speciation of metals, light elements (carbon and nitrogen), and light-element isotopes, as well as the taxonomic identities of bacteria (e.g., using FISH) at the pore scale (nanometers to hundreds of microns).** Although it does not currently exist, this capability could emerge from the development of synchrotron fluorescence-yield X-ray microscopes with mass spectroscopy and FISH capabilities.

D.2 Pore-Scale Biogeochemistry

The subsurface contains a hierarchical network of pores ranging from nanometers to millimeters in diameter. Pore size and connectivity influence water structure, dynamics, and movement (Cheng, Hu, and Hu 2012) that affect the chemistry occurring at solid-water interfaces (Brown and Calas 2012) and mass transfer. Biological components and processes are intertwined with this pore network (Bengough 2012). Effects range from enzymatic exclusion in nanopores that causes reduced bioavailability to specialized biogeochemistry near root-bacteria-fungi communities (Philippot et al. 2013). These phenomena must be quantified to solve for macroscopic system response based on molecular information.

Goal: Parameterize biogeochemical processes as a function of pore size and interface chemistry to inform larger-scale modeling of subsurface processes and subsequently connect with Earth system models.

The influence of pore size on the biogeochemical thermodynamics and kinetics of biogeochemistry is not known. The physicochemical properties of water are a strong function of pore size (Michot et al. 2002) and will affect those processes, but data are not available that can be used to model these changes and their impacts on subsurface behavior. Because nanometer pores have near-zero advection and slow water diffusion, any addition of solute to these pores will have a major impact on the thermodynamic saturation state. Furthermore, because the dielectric constant of water decreases near a surface, the "aqueous" biogeochemistry in narrow pores is dramatically different from the bulk water solutions used in most laboratory experiments. Deviations from bulk behavior occur because the structure and diffusivity of water are different in pore environments (Köfinger, Hummer, and Dellago 2009; Bonnaud, Coasne, and Pellenq 2010; Solveyra et al. 2013).

Details of the physicochemical properties of water confined in pores can be analyzed by molecular-based studies focused on water structure, diffusivity, and dynamics. Such studies can be accomplished through a multifaceted approach that investigates processes through a combination of theory and experiment. This approach includes the integration of molecular dynamics, quantum chemical studies, and state-of-the-art spectroscopy and more classical batch studies of kinetics. The classical studies measure an average value, but the tails of the putative normal distributions may actually control water flow in the environment. Methods that can probe individual molecules through isotope labeling and on fast timescales may be particularly useful in this regard (Park, Kwak, and Fayer 2007). Also, the structural properties and the movement of water in confined nanopores can be probed by neutron scattering and imaging techniques. The goal of these studies is to better understand the property of water in pores and the transport of material within this water layer as a function of pore size. The integration of data from new approaches with those from classical studies will potentially result in unexpected findings that were unattainable without these new techniques and methodologies.

Organisms exude compounds and carry out chemical reactions in small pore spaces separated from the average chemical conditions of the macroscopic system. Within small pores, exudates may more effectively alter pH, redox potential, and equilibrium states. One example is iron acquisition. In the subsurface, with oxidized iron minerals, ligands such as oxalate and siderophores may be used to solubilize low-solubility Fe(III) (Lemanceau et al. 2009). The smaller the space in which these molecules are released, the higher the effective concentration of the ligands and the greater the efficiency of nutrient acquisition. Pore size exerts a strong influence on the physical, chemical, and biological processes of the subsurface; and these processes must be studied under *in situ* conditions that replicate the subsurface.

The suggested strategy to incorporate pore-scale information into larger scales is to design molecular-level studies within the context of macroscale studies and allow for easier upscaling in space and time (Section II.E, Linking Subsystems and Processes Across Scales to Describe Ecosystem Behavior, p. 30). A useful tactic would be to support molecular-level researchers to work with scientists examining larger-scale processes (mesoscale to macroscale) to ensure that information transfer occurs among individuals and that results from one level are translatable to larger scales. An example is studies of heterogeneous nucleation within micropores. Thermodynamic and kinetic information for this nucleation process could be incorporated into reactive transport models to calculate changes in porosity and permeability more accurately in real time during the simulation. Another example is the transport of matter with nanoparticles. Incorporation or adsorption of insoluble elements into or onto nanoparticles can increase the transport of these elements dramatically and move matter over long distances through advection. Knowledge of how these nanoparticles nucleate, grow, and adsorb elements could become part of fluid flow simulations that account for this mass transfer process and support studies of aerosol chemistry.

D.3 Plant-Soil Interactions

Goal: Improve knowledge of molecular-scale processes occurring at the plant-soil interface to inform macroscale and global-scale processes.

Interactions among plants, microbes, minerals, and organic carbon are critical components of plant productivity. Plant inputs to soils in the form of organic matter fuel heterotrophic microorganisms, which in turn influence the formation and mineralization of SOM. From a climate perspective, plant-microbe-mineral interactions influence the production of greenhouse gases, including N_2O, CO_2, and CH_4. The relative distribution between storage of carbon and nitrogen in soils and their release as atmospheric greenhouse gases is regulated in part by the coupled biology of belowground plant allocation and the microbial metabolism of substrates. Biomolecules produced by microorganisms in response to plant inputs can interact with mineral surfaces to influence soil aggregation, which physically protects soil carbon. Together, soil aggregation and root architecture influence soil porosity, soil water-holding capacity, and organic matter accrual, which can generate positive feedbacks for plant productivity (see Fig. 8. Research Integration, p. 28).

Engineering plants to maximize productivity without considering plant-microbe-soil interactions and mineralogical factors can decouple the soil ecosystem. This decoupling could result in biogeochemical shifts that cause the loss of carbon and nitrogen to the atmosphere or aquatic systems. Furthermore, reducing microbial diversity can limit the processing of plant inputs to the soil, limiting decomposition, soil mineral aggregation, and the physical and chemical protection and stabilization of carbon and nitrogen in soils. For this reason, an explicit understanding of plant-microbe-soil interactions is required for designing plants with improved phenotypic traits and predicting ecological changes among multiple plant species in complex hydrobiogeochemical soil systems.

As indicated earlier, to achieve a mechanistic understanding of the plant-microbe environment in the context of BER programmatic goals over the next 10 years, the ultimate technological objective should be to develop the capability to detect and quantify plant or microbial metabolites in soil in real time without pretreatment or amendment. These metabolites should include inorganic or organic constituents that are the reactants and products of assimilatory and dissimilatory metabolism in autotrophic and heterotrophic organisms. With this information, models of plant-microbe-soil interactions can be built and incorporated into land process models to improve the ability to predict the interactions of plant- and root-driven

biogeochemical cycles and climate. Coupling experimentation and computational research also offers the ability to elucidate the unifying principles present in the vast array of complex soil ecosystems and illuminate emergent soil properties.

D.4 Microbe-Mineral Interactions

Goal: Develop detailed, mechanistic understanding of the molecular-scale and surface-mediated processes that control microbe-mineral interactions, including electron transfer, sorption and desorption, and mineral growth and dissolution.

By producing and consuming minerals, microbes profoundly influence Earth's soils and sediments. Biogenic minerals precipitated in near- and below-surface environments are often poorly crystalline and reactive (e.g., Nelson et al. 2002). Because of their disordered structure and large specific surface areas, these inorganic phases serve as sorption substrates for both organic and inorganic soil components. Biogenic minerals sequester contaminants, provide reactive surfaces that accelerate reactions, and oxidize SOM. Microbes mediate mineral dissolution by modifying pH, donating or accepting electrons to and from minerals, and producing metal-complexing ligands. These reactions support microbial metabolism; affect the redox potential and pH of pore waters; and release stored reactants, nutrients, and contaminants. Mineralization of bacterial cells, necromass, and extracellular biomass on mineral surfaces produces reactive high-surface-area mineral coatings. These activities bestow an outsized chemical influence upon biogenic inorganic precipitates that is disproportionate to their small
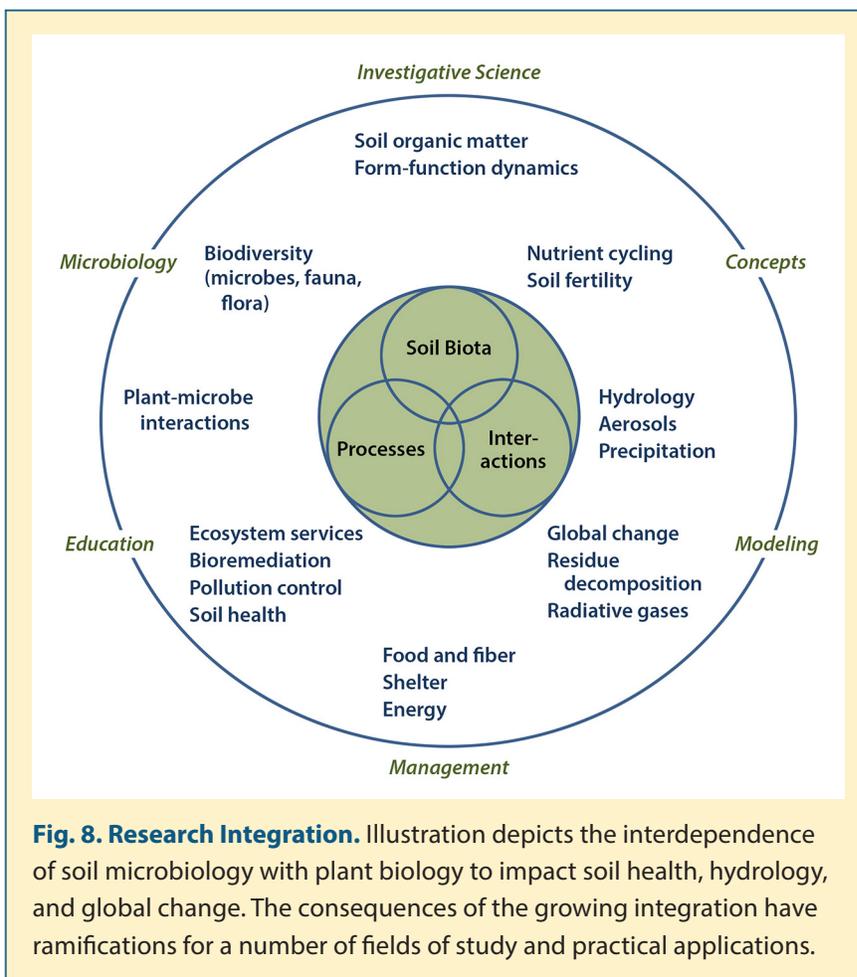
Fig. 8. Research Integration. Illustration depicts the interdependence of soil microbiology with plant biology to impact soil health, hydrology, and global change. The consequences of the growing integration have ramifications for a number of fields of study and practical applications.

mass fraction. They substantially alter the adsorption properties of the underlying soil grains, especially with regard to hydrophobic compounds (Murphy, Zachara, and Smith 1990). Organic-inorganic interactions have a particularly profound influence on soil carbon turnover rates (Torn et al. 1997; Koven et al. 2013). Developing an understanding of the molecular-level processes that govern microbe-mineral interactions directly supports the prediction of below-surface climate interactions (Koven et al. 2013).

Major questions that must be addressed to advance this field include:

1. What limits the growth and controls the morphology of biogenic nanominerals?

2. What kinetic and thermodynamic factors allow nanobiogenic minerals to persist?

3. How do SOM and metal ions bond to biogenic nanominerals?

4. How does the biological matrix mediate biomineral properties?

5. What are the structures of biogenic nanomineral surfaces?

6. How are electrons transferred from enzymes and electron shuttles to and from mineral surfaces?

7. What are the rates of these electron transfers and what controls the rates?

Answering these questions will require applying time-resolved spectroscopy tools to follow the sequential reactions. Time-resolved optical absorption, vibrational, and X-ray absorption techniques exist. However, developing appropriate triggers to initiate reactions and optimizing experiments will be challenging and will require significant investments.

Decadal Thrusts

- Incorporate plant-soil interaction processes into land process models.

- Assemble atomic-scale structure libraries for common types of natural organic molecules found in pore waters and mineral soils.

- Assemble databases of proton- and metal-binding affinities or quantitative structure-activity relationship models of functional groups in natural organic molecules found in pore waters and mineral soils.

- Characterize the mechanistic molecular-scale linkages between carbon cycling and metal redox transformations at the pore scale and incorporate these data into land process and climate land models.

- Characterize the structures, compositions, and chemical reactivity (protonation and charging behavior) of mineral surfaces exposed inside pores.

- Characterize the thermodynamics and kinetics of aqueous and mineral-water reactions as a function of pore size.

- Measure fluid flow rates through nanopore networks as a function of pore size and network topology.

- Characterize the physical and electronic structures of nanomineral surfaces and interfacial species.

- Resolve individual steps in interfacial reactions at biogenic nanomineral surfaces, including electron transfer.

- Characterize molecular mechanisms by which plants recruit or select for beneficial microbial communities.

- Identify microbial compositions and functions that promote plant productivity.

- Develop a library of biomolecules produced by plants and microbes that contribute to soil aggregation and physical protection of carbon and nutrients in SOM.

- Generate models to inform empirical laboratory and field studies designed to characterize relationships among plants, microbes, and soils important for the storage and release of carbon, nutrients, or contaminants.

To be useful, molecular modeling will need to approximate the actual systems as closely as possible. The number of atoms necessary to generate models similar in concentration and complexity to real-world situations is on the order of 10^5. Simulations of systems of this scale with quantum mechanical calculations are currently improbable, but they may be possible in the next decade with advances in hardware, software, and theory (Bylaska, Weare, and Weare 2013; Shimojo et al. 2014). Hence, classical force field simulations will probably be the dominant mode of tackling this challenge. A focus on developing accurate force fields with the ability to model chemical reactions is probably the most efficient route to success in this area. By working with larger-scale models running reactive transport or lattice Boltzmann codes, molecular-level

information can be parameterized and incorporated into larger-scale systems. Again, information technology will be imperative to making these results available and integrating them with the experimental and analytical data discussed earlier. Tools for checking the consistency of models compared with observations should be developed to allow modelers to verify results and to encourage experimentalists to interpret data with current models.

E. Linking Subsystems and Processes Across Scales to Describe Ecosystem Behavior

Vision

A grand challenge for the next decade of BER science is to incorporate molecular-scale data accurately into ecosystem and Earth system models. Transforming molecular-scale information into ecosystem-scale knowledge is a formidable challenge that requires an unprecedented breadth and depth of understanding among fields. A major milestone in this process, integrating molecular data into Earth system models, requires identifying keystone biogeochemical processes and characterizing their behavior at many scales. Multiscale perspectives will play important roles in illuminating these keystone microprocesses. This knowledge will accelerate the development of sustainable bioenergy production systems, land-use management strategies, and more accurate carbon cycling models.

Molecular-scale and microscale biogeochemical interactions have large-scale repercussions, as evidenced by aquatic blooms, the fate of pollutants (Azam and Smith 1991; Paerl and Pinckney 1996; Stockdale, Davison, and Zhang 2009), and transfer of greenhouse gases and mineral aerosols to the atmosphere (Bagley et al. 2014; Engelbrecht and Derbyshire 2010; see sidebar, Computing Biological and Climatological Interactions Across Scales, p. 31). Environmental subsystems residing at different locations in an aquifer are biogeochemically linked through advective and diffusive exchange of solutes. We are on the verge of incorporating conceptual models into reactive transport models to quantitatively predict solute exchange between eco-subsystems. However in some cases, the presence, direction, and magnitude of microscale responses can change when aggregated to the ecosystem and beyond (Young et al. 2008; Ehrenfeld, Ravit, and Elgersma 2005; Bridgham et al. 1995; Holmes et al. 2005). A network of laboratory and field experiments must be designed to characterize molecular mechanisms *in situ*, test scaling relationships, identify sensitive parameters, and generate model-relevant data. In addition to this network of coupled laboratory-to-field experiments, databases and computational, statistical, and modeling efforts need to be expanded to facilitate access and analysis of multiscale data. To ensure the rapid success of scaling systems science, tools for assembling, visualizing, and analyzing deep reductionist science need to be developed and deployed with an awareness of the interactive nature of biogeochemical reactions. Incorporating the cascading scaled responses—both biological and geochemical—into larger systems will enable the identification of key reactions (requiring high-resolution data) or regions (biogeochemical hot spots, see Section E.1 below) that most strongly influence global elemental cycling and Earth's climate. We envision using microscale and mesoscale models in conjunction with replicated laboratory-to-field experiments to quantify how molecular interactions influence large-scale responses and vice versa.

E.1 Biogeochemical Hot Spots and Hot Moments

There are physical *locations* in an ecosystem where biogeochemical reactivity is high compared with the surrounding soil or subsurface matrix. The coupling of biogeochemical cycles is relatively intense at these sites, and reactions proceed at relatively fast overall rates. These are termed biogeochemical hot spots (Vidon et al. 2010; McClain et al. 2003). Examples of hot spots include microsites and aggregates in

Computing Biological and Climatological Interactions Across Scales

Imagine a computational resource allowing us to move seamlessly across biological or atmospheric scales. From an initial view of molecules inside a cell nucleus, we would continue outward to see the cellular components, then the whole cell, the cell in the context of a tissue interacting with adjacent cells, the tissue as part of an organism (e.g., a tree in a forest, the forest as part of a geographical region, and then as part of a continent), and finally to a view of the whole Earth (see Fig. 9. Investigating Biological and Environmental Interactions Across Scales, below). (A related example might be the rhizosphere associated with a tundra biome that has carbon and climate impacts.) For atmospheric science, we might envision seeing the data taken within a column above an Atmospheric Radiation Measurement Climate Research Facility site placed in context within the broader image of the region surrounding the site and its evolution over time.

Computational modeling at multiple scales is required for atmospheric systems and understanding biological systems within a bacterium or in the context of environmental and ecological interactions that ultimately impact the carbon cycle and climate. Whatever the scale of study, each discipline has existing models and will continue to refine and improve them as new experimental data become available and new analysis methods are developed.

Models could interact at each scale, but today, there is no continuity from models of microbial systems to those that predict global climate. Researchers in each discipline are focused on improving current scientific understanding in their domain and generally view today's models as crude. As models become refined and increasingly precise at each scale, the question of how a model can interact meaningfully with those at adjacent scales will present opportunities to gain further insight.

Fig. 9. Investigating Biological and Environmental Interactions Across Scales. Understanding how complex biological and environmental systems will respond to and affect critical Earth system processes requires measuring, simulating, and integrating biological, chemical, and physical components and their interactions across vast spatial and temporal scales—from subnanometers to kilometers and nanoseconds to millennia. Shown here is the example of moving from an initial view of molecules inside a cell, to cellular components, whole cells, the cells of a tissue interacting with adjacent cells, the tissue as part of an organism (e.g., a tree), the organism within a geographical region of a continent, and finally to a view of the whole Earth. As computational models become refined and increasingly precise at each scale, the question of how a model can interact meaningfully with those at adjacent scales will present opportunities to gain a deeper system understanding. [Image from p. 46 of BERAC 2010.]
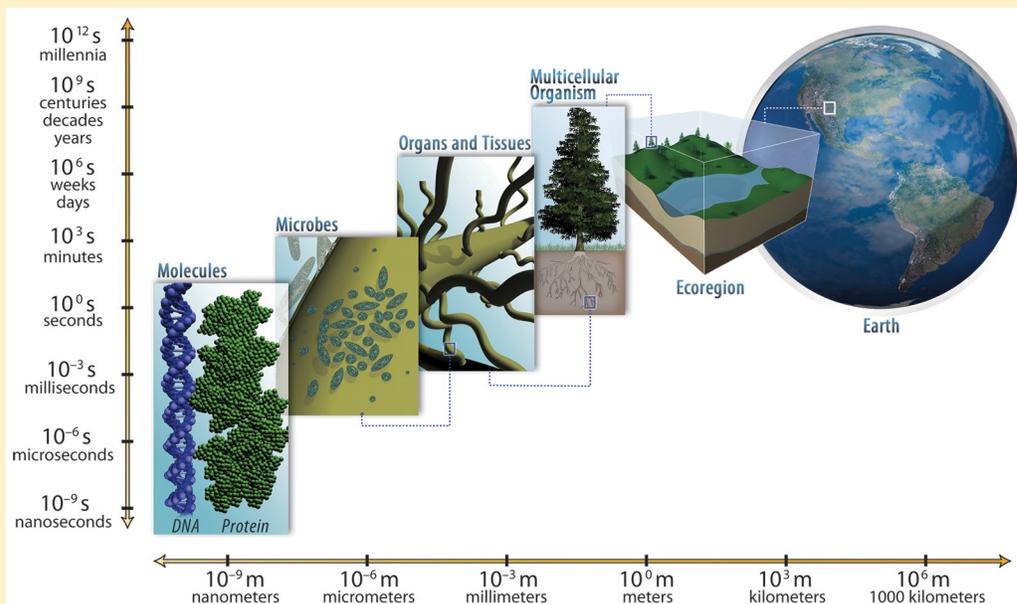
soils (McClain 2003; Six, Elliott, and Paustian 2000; von Lützow et al. 2006), organic-rich sediment zones (Hill et al. 2000), hyporheic zones (Fuller and Harvey 2000), soil-stream interfaces (Hedin et al. 1998), and plant roots (Schade et al. 2001). Hot moments are periods of *time* when biogeochemical activity is high (McClain et al. 2003). Hot moments often are associated with changing surface water conditions that occur following rain, irrigation, or snowmelt events (Boyer et al. 1997). Similarly, seasonal variations in water table elevations can bring groundwater into contact with nutrient-rich sediments, resulting in flushing of accumulated soluble nutrients and the initiation of (short periods of) intense biogeochemical redox activity.

Hot spots and hot moments are important to ecosystems because they host a disproportionately large amount of biogeochemical activity. For example, nutrient cycling in soils is relatively intense following wetting events, particularly so in desert soils where rainfall is infrequent (Fliessbach, Sarig, and Steinberger 1994; Zaady, Groffman, and Shachak 1996). There may be little biogeochemical activity in between these events. Organic-rich sediments in the aquifer at the Rifle, Colorado, DOE field research site (Campbell et al. 2011) act as biogeochemical hot spots within an otherwise nutrient-poor aquifer. Redox reactions hosted at these locations are believed to regulate the mobility of macro- and micronutrients—including oxygen, nitrogen, and iron—in the surrounding floodplain, as well as the long-term mobility of uranium. In this fashion, hot spot and hot moment biogeochemistry is expected to control fluxes of nutrients and contaminants in terrestrial environments and between terrestrial-aquatic and terrestrial-atmosphere systems. Biogeochemical reactive transport models that fail to account for hot spots and hot moments or underestimate their abundance in an ecosystem are likely to perform poorly in predicting carbon fluxes, nutrient cycling, organism phenotype, contaminant mobility, ecosystem response to changing environmental conditions, and resiliency.

Hot spots and moments emerge from the coupling of specific suites of chemical and biological reactions in time and space. Molecular-scale information is vitally important to developing quantitative models to understand these complex systems and their responses to environmental perturbations, which may include contaminant (im-)mobilization or changes in nutrient and greenhouse gas fluxes. Molecular-scale knowledge about dissolved species, such as the chemical form of nitrogen in pore waters, provides information about the types of microbial functions and biogeochemical processes occurring in a component within an ecosystem. Information about the solid phase—such as the oxidation state of sulfur or iron, or the chemical composition of soil aggregates—provides insight into the reservoir of reactive species and thus the *potential* for these and other processes to occur and the possible *duration*. The ultimate goal is to collect molecular-scale information in four dimensions, that is, in space and time, and to use this information to reconstruct the ecosystem biogeochemical function. Ultimately, information provided by such studies is expected to improve predictive ecosystem-scale models profoundly by focusing attention on zones and reactions of greatest importance.

Decadal Thrusts

- Support ongoing discovery of new types of hot moments and hot spots by developing techniques to identify the three-dimensional (3D) distributions of hot spots and moments at field sites over scales from microns up to kilometers.

- Characterize the activity and biogeochemical signatures of essential molecular-organismal-mineral processes at the pore scale that drive the function of "hot" systems in the field.

- Develop a network of intertwined laboratory studies and carefully designed field observations to address molecular- to pore-scale coupling between major elemental cycles such as organic carbon degradation and metal redox cycling.

- Discover and elucidate the molecular-scale mechanisms driving the rates, extents, and directions of greenhouse gas fluxes in response to environmental transitions (e.g., thawing Arctic tundra).

- Integrate molecular mechanisms to build upon, and project beyond, the confines of simple reaction networks used in large-scale quantitative Earth system and reactive transport models.

- Use laboratory-to-field experiments to parameterize the molecular-scale controls of ecosystem behavior so that ecosystem responses to climatic perturbations and other disturbances can be forecast.

- Map the biogeochemical activity (e.g., processes, microbial functions, and reactions) of hot spots and hot moments in ecosystems.

Meeting these goals will require developing techniques with high spatial and temporal resolution to identify hot spots and moments in the field (e.g., satellite and geophysical). New capabilities are needed to image the 3D distributions of hot spots and biogeochemical activity on the micron to millimeter scale. Such capabilities (which do not yet exist) need to be coupled with *in situ*, real-time detection of biogeochemical reactions and physicochemical properties (e.g., redox potential, carbon and nutrient species, rhizodeposition quantity and quality, and plant or microbe communication molecules).

Linking processes from the molecular to ecosystem scale requires new instrumentation to quantify fungal, bacterial, and archaeal biomass and capture extracellular enzymes in the environment. Similar to rhizotron technology, *in situ* imaging of soil and sediment structure and porosity would enable the detection of biogeochemical connectivity in transition zones or in response to perturbation events (e.g., thawing or flooding).

Most important, extensible and scalable software that can mix models and data in a modular fashion is urgently needed. This software will arbitrate the mixing of modules that cross multiple scales of time,

space, biological organization, or new dimensions to be discovered. Such software also will provide a framework for biogeochemical reactive transport predictions and for predictions based on integration of available multi-omic data over time (phenological responses) to inform community- and ecosystem-scale process measurements.

F. Technology and Capabilities Needed

Linking genotypes with phenotypes and identifying biogeochemical cycling of carbon, nutrients, and contaminants will require comprehensive and creative approaches that use existing knowledge and technologies in new ways, such as combining nano-SIMS, X-ray nanoprobe, FISH, and Chip-SIP approaches. Existing analytical capabilities will need to be significantly improved or adapted from other fields.

- Detection limits and spatial resolution for synchrotron-based X-ray techniques and mass spectroscopy need to be advanced by orders of magnitude.

- Entirely new techniques and approaches will need to be developed to couple plant or microbial genotype directly with phenotype or function in the environment.

- Omics approaches will have to be applied to plants and scaled for high-throughput application, and the resulting datasets will need to be reduced to tractable algorithms that can be visualized or incorporated into models. Molecular and geochemical modeling approaches must be significantly advanced.

- The detection and quantification of plant and microbial metabolites in living tissues in real time without pretreatment or amendment will be necessary for modeling whole cells to decipher molecular changes during growth and evolution. Imaging techniques should be improved to determine the concentrations of reactants and products of metabolic reactions down to the nanometer scale.

- A catalog of SOM components is needed, as well as the ability to predict compositional changes in SOM from known biogeochemical parameters. Determining the relationship between SOM composition and soil quality will be necessary.

- *In situ* imaging of soil and sediment structure and porosity is needed to enable detection of biogeochemical connectivity in transition zones or in response to perturbation events.

- In this era of open and rapid science, empiricists lack the means of organizing, archiving, and distributing data to the scientific community. Deploying computational tools, including database platforms, extensible and scalable software, and novel statistical and computational approaches, will allow scientific discovery to match the pace of rapid data generation.

- Capabilities for workforce development must complement all aspects of this scientific frontier.

III. Synthetic and Genomic Bioscience

A. Decadal Overall Vision

In the 16th and 17th centuries, Galileo (and other scientists) observed motions across multiple scales, from pebbles on Earth to planets in the solar system. These researchers tried to understand and predict the basis of motion as well as reconstruct three-dimensional (3D) pictures of the solar system with time-resolved accuracy. With the data available at the time, Newton took on the challenge of formulating the fundamental laws of motion and converted descriptive science to predictive physics. The resulting laws set the foundation for understanding, simulating, and predicting all interactions in the physical world above atomic scales. Those laws have stood the test of time and today enable the understanding of a wide range of biological phenomena, spanning from molecular dynamics of proteins to water pumping in plants.

Similarly, during the 20th century, a revolution in the biological sciences led to the derivation of the "laws of life." These resulted in a mechanistic understanding of inheritance and evolution of traits, the biomolecular functions that form the basis of the modern life sciences. Discoveries of the double helix, the "central dogma" of DNA-RNA-protein information transfer, and energy transduction across membranes are examples of key mechanistically driven studies.

Now, in the 21st century, the research community is confronted with a fusion of biology with the information revolution. In the past decades, various data-producing technologies, particularly bioinformatics and DNA sequencing, have advanced at breathtaking speed, pushing the envelope toward an era of systems biology. These data have enabled the re-engineering of biosystems for various purposes and resulted in several notable success stories from the laboratories of Jay Keasling and James Liao, as examples (Ro et al.

2006; Higashide et al. 2011). However, this inundation of biological data has led to the question, "How do we proceed from here?"

B. Toward Newtonian Biology

What should be initiated now that will ultimately serve those in the biological sciences in the 22nd century? What is needed to advance biology and biosystems design beyond the current status? Both hypothesis- and discovery-driven approaches have made significant contributions to the understanding and re-engineering of biology. However, the influx of data leads to particular challenges. To meet this massive data challenge, the trend in many disciplines is to turn to computational science to "integrate" disparate detailed information into working models capable of providing causative explanations of the behavior of complex systems. In turn, well-designed models will inform experimentalists as to missing data needed for model improvement.

A biological "parts list" defines the components of working cells under various conditions, generating a descriptive or a structured reconstruction of cells and organisms. However, although computer models generated on these bases may describe features of biological systems behavior, their predictive power is limited unless the parameters and parts lists are updated for each system. Even without complete knowledge of all interacting bodies in any given system, simplified computational models often capture the essence of behavior. Based on these results, engineering decisions can be made for design purposes. Increasing knowledge then sharpens the resolution of prediction. In such systems, conceptual models that capture the principal mechanism of interaction sometimes are even more powerful than a detailed description of reality. How can that concept be applied to biology?

The interplay between data-driven computation and experiments in biology will need to go beyond reconstruction and short-range predictive capability. What is needed is the derivation of new underlying biological principles, analogous to the interplay between Galileo and Newton. Newtonian physics applies to all levels of complexity. The biological field would benefit enormously from the discovery of rules that underlie the interaction and evolution of biomolecules and processes. These principles should be more general than individual mechanisms and be guided by the laws of physics and chemistry. They may drastically reduce the complexity of analyzing enormous datasets and serve as guiding rules for the redesign and construction of biological systems. One biological example that reinforces the existence of such simplifying principles is the case of enzyme activity. At one time, enzymes were considered almost infinitely variable, but now they are understood to obey a limited number of well-described chemistries that allow binning of the majority of enzymes into six major classes.

The principles of chemiosmosis and electron bifurcation have changed and probably will continue to change the field of bioenergetics, especially in regards to how biological systems conserve energy. The latter has profound implications for biofuel production. In 1967, Peter Mitchell hypothesized that as electrons are spontaneously transferred from a low-potential donor to a higher-potential acceptor within membrane protein complexes, the free energy released is conserved by the pumping of protons across the membrane (Mitchell and Moyle 1967; see Fig. 10. Known and Unknown Principles of Energy Conversion Coupled to Electron

Transport, this page). This chemiosmotic principle of transferring an electron motive force into a proton motive force (that can then be used to drive thermodynamically unfavorable chemical reactions) has been a foundation of biology ever since. It represents the second mechanism by which life conserves energy, in addition to the direct chemical synthesis of so-called high-energy phosphate compounds such as adenosine triphosphate (ATP).

More recently, in 2008, Wolfgang Buckel and Rolf Thauer (Herrmann et al 2008; Li et al. 2008) proposed a third mechanism of energy conservation in certain enzymes. In this case, as electrons are spontaneously transferred from a donor to an acceptor in some enzymes, the free energy released is directly coupled to the reduction of that same acceptor by another donor that has a more positive potential than the acceptor. Electron bifurcation enables much

Fig. 10. Known and Unknown Principles of Energy Conversion Coupled to Electron Transport. Examples of electron donors (D) and electron acceptors (A) are indicated. [Image courtesy Michael Adams, University of Georgia]
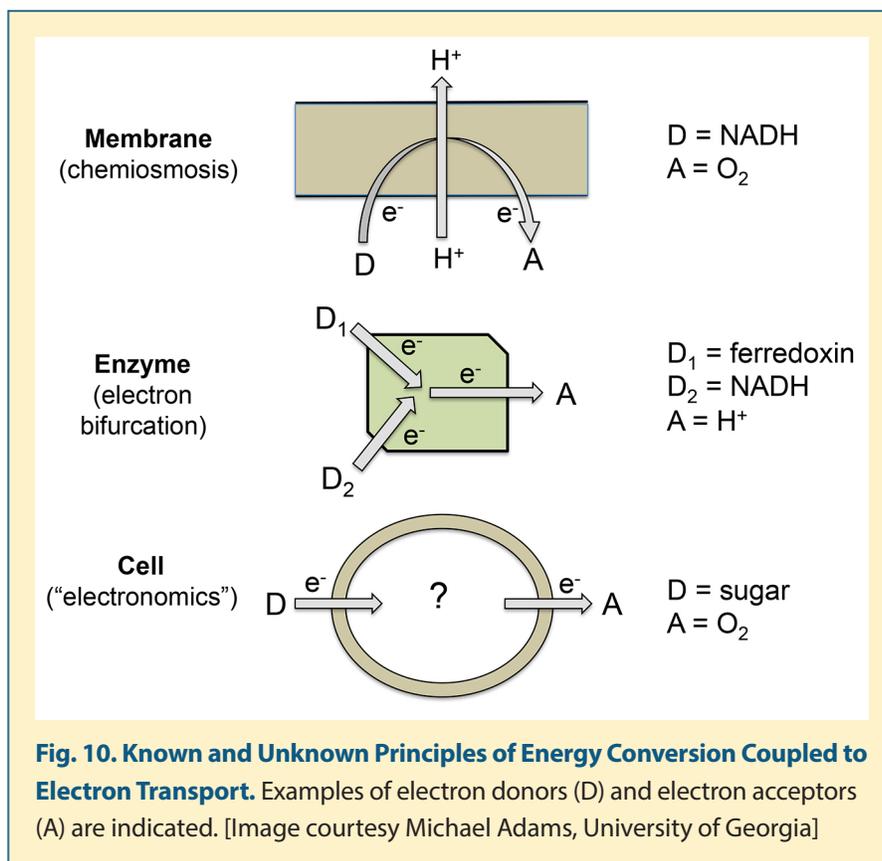
smaller amounts of energy to be conserved and used than the amount required to generate ATP. Although this mechanism of conserving energy is still poorly understood, primary metabolic processes, particularly in anaerobic environments, are now being viewed from a completely new perspective (Buckel and Thauer 2013). Moreover, electron bifurcation is proposed to be an ancestral form of energy conservation that in part may have driven the origin of life (Nitschke and Russell 2012).

ATP is often used to drive thermodynamically unfavorable and kinetically challenging reactions. ATP may serve both as an energy source and a kinetic driving force; in such cases, energy expenditure is favored in exchange for kinetic advantage. Moreover, if organisms are engineered to perform a non-native task, such as the production of biofuel, is ATP conservation a desirable goal? Examples have shown that ATP draining may serve as a driving force for faster substrate consumption, particularly in engineered organisms. Therefore, the interplay between energy conservation and kinetic expediency is complex and may differ in organisms engineered for non-native purposes.

If the principles of energy conservation by a membrane and an enzyme are known and understood, what might be the principles for a whole cell based on a systems biology viewpoint? How does a cell use substrates and manage energy transduction while achieving a native or non-native function? How is electron flow regulated at the cellular level? Can systems biology data be used to derive principles of whole cell "electronomics"?

Applying the Newtonian approach, we most likely do not need the complete parts list of the cell. Nonetheless, the biofuels industry would be revolutionized by the ability to apply such principles to direct both carbon and electron flow in a cellulose-digesting microbe so that it produces a reduced carbon compound of choice. What are the "Newtonian principles" of cellular energy conversion that can guide the production of a sustainable energy supply?

C. Toward a Multiscale 3D View of Cells

Although a full parts list may not be needed to discover the underlying principles of biological systems, 3D time-resolved observation of biological events could help identify the key players. Current experimental and computational tools examine only narrow windows of spatial, temporal, or chemical information, leading to a fragmented vision. The Office of Biological and Environmental Research (BER) is supporting ongoing efforts to collect and interpret omics-based information to define molecular systems for metabolism, regulation, and signaling. This support includes BER investment in the Department of Energy (DOE) Systems Biology Knowledgebase (KBase), a computational platform for integrating and analyzing such research and the disparate data types that result. However, this computational strength must be complemented with capabilities to visualize, conceptualize, and test these molecular networks with physical models on relevant timescales. Furthermore, the ability to observe and measure the impact of these molecular systems at successively linked system scales will be required.

An important foundation to advance Newtonian biology is the knowledge of macromolecular structure and dynamics, the positions of subcellular structures, the positions of macromolecules with respect to those expressed structures, and the distribution of metabolites and ions. Advanced experimental imaging techniques need to be developed to enable the capturing of this information. Such knowledge could then aid in identifying the governing principles of cellular functions. Molecular-level simulation can play a crucial role, as illustrated by the 2013 Chemistry Nobel Prizes (Smith and Roux 2013). However, computational methods still are needed that enable reliable understanding of the mechanisms of enzyme functions and complex, multimolecular machines. Also required are methods that efficiently and fully leverage available computational power to understand folding, allostery, binding, and reaction. Moreover,

knowledge of thermodynamics and energetics at the single-protein level is critical to understanding cellular energy flow. The research community needs to master the structural control of electron donor and acceptor redox potentials to refine methods for calculating the free energy profiles of enzymatic reactions. Ultimately, methods need to be developed to predict protein structures reliably (in the absence of a close structural homolog) and protein:protein and protein:ligand associations. Understanding these mechanisms will permit deeper insights into and derivation of principles of biological functions, from the molecular to whole-organismal level.

The understanding of molecular mechanisms and cell functions may benefit from supercomputing, which now permits the molecular dynamics (MD) of systems up to 100M atoms in size to be simulated in atomic detail and on a microsecond timescale. Until recently, the starting structure for such a simulation would have been thought unattainable. However, a recent synapse study using an integrative approach provides a glimpse of the future. The study combines quantitative immunoblotting and mass spectrometry to determine protein numbers; electron microscopy to measure organelle numbers, sizes, and positions; and super-resolution fluorescence microscopy to localize the proteins (Wilhem et al. 2014). With these data, a 3D model of an "average" synapse was generated, displaying 300,000 proteins in atomic detail. The challenge remains to derive fundamental principles that ultimately lead to the "laws of the cell," much like deriving Newtonian laws of motion from a large amount of space- and time-resolved observations.

Deriving the principles of biology requires distilling data from myriad sources to arrive at simple expressions that define and predict observables. Some of these defining principles already exist. In fact, many scientific disciplines are gathering data and seeking to understand biological systems and principles. Such disciplines include microscopy, spectroscopy, genetics,

omics, computational simulation, and, notably, systems biology. As the basis of life, biological systems are complex and self-organizing and push the limits of thermodynamics, maintaining far-from-equilibrium states and processes at the expense of energy dissipation and entropy creation. The critical processes are molecular and electronic at their most basic and detailed level, and they combine to produce increasingly higher levels of complexity with accompanying emergent behaviors that are far simpler than might have been predicted. At each level of complexity, principles must be derived to describe and define—in the most meaningful and simplest way—biological behaviors and emergent properties so that function can be predicted, ultimately, on a systems level.

On one hand, the overarching principles of biology can be a combination of existing principles. These overarching principles may simplify special cases, much the same as Newton's laws are the special limiting case of quantum mechanics. For example, collision theory and quantum mechanics can be used to describe a chemical reaction rate, but having a simple second-order rate equation that captures all the behavior and physics of lower-level theories is much more useful. It is important to note that the goal is not to fit some functional form to a behavior, but rather to find the most appropriate form that reflects the underlying mechanisms. Even if the underlying mechanisms are not known, the simplest form with fewest parameters is most likely the one that reflects the underlying mechanisms.

With the "laws of the cell" and advanced computing, the wealth of systems biology data will play a significant role in understanding, predicting, and redesigning biological systems for advanced medicine, renewable energy sources, and environmental sustainability. Currently, the level of detail that must be observed to establish biological principles remains an open question. Techniques are being developed that should enable complex events such as energy and electron flow in four dimensions to be addressed. Ultimately,

accumulated data can be used to investigate such questions at multiple scales ranging from a cell to organisms to communities and eventually the planet. Similar to Newton's laws of motion that apply from the atomic to planetary scale, what are the principles of biology that can be applied to, for example, a bifurcating enzyme to the consequences of global warming?

D. Technology and Capabilities Needed

D.1 Three-Dimensional Systems Biology

Systems biology goes beyond the consideration of single macromolecules to obtain holistic information about interacting biological systems, such as metabolic networks, genomics, and proteomics. However, an important element missing from these descriptions is a *3D, time-dependent picture of the systems involved*. This 3D view requires knowledge of macromolecular structure and dynamics; the positions of subcellular structures and the macromolecules with respect to those expressed structures; and the distribution of smaller species such as solvents, metabolites, and ions. Obtaining such views would advance Newtonian biology.

D.2 Experimental Molecular-Scale Research

Biology originates from the molecular scale, and rational engineering of biological systems in the energy and environmental biosciences will need to drill down to this level. DOE is ideally equipped to support molecular-scale research at its large-scale facilities for next-generation synchrotron radiation (including free-electron laser instruments), neutron scattering, and supercomputing. Optimally using these facilities for molecular-scale research will be increasingly valuable. Synchrotron radiation will be useful for determining the 3D structures of macromolecules and their complexes. Further development is needed to fully leverage small-angle X-ray scattering (SAXS) and extended X-ray absorption fine structure (EXAFS) to answer structural biological

questions. Neutron scattering provides complementary techniques over a range of space and time scales—from Ångströms to micrometers and from femtoseconds to microseconds. The ability to locate hydrogen atoms in enzyme active sites using neutron crystallography accompanied by high-level quantum mechanics and molecular mechanics computations promises to provide unprecedented detail on enzyme function. Small-angle neutron scattering needs to be coupled with SAXS and computation to understand structures of large molecules and complexes in solution. Although unique in its ability to probe the dynamics of molecular systems, neutron scattering has been largely unexploited in research sponsored by BER. Knowledge of dynamics on the femtosecond to microsecond timescale is also fundamental to understanding the properties of protein machines and membranes.

Unstructured elements of biological systems also are likely to be increasingly important in the future. These elements resist structural determination but have been identified as playing important roles in biological function. Examples include disordered linker regions in multidomain proteins and assemblies in organisms relevant to energy and subsurface science. Little is known about the structure, dynamics, and function of these disordered systems. The integration of solution neutron and X-ray scattering with supercomputing may provide missing information.

D.3 Computational Simulation

Enzyme reactions are critical to molecular systems biology. Techniques that combine quantum and molecular mechanics to compute reaction mechanisms and associated energetics need to be improved and accelerated to provide missing information in metabolic pathway maps and guide rational engineering of enzymes. Through efficient and full use of available computational power, advances are being made in understanding protein folding, allostery, binding,

and reaction. This understanding is rate-limiting for synthetic biology and should be stimulated. Methods also should be developed for reliably predicting protein:protein and protein:ligand associations in a high-throughput manner.

Supercomputing now permits MD simulation at atomic detail of systems up to 100M atoms in size and on a microsecond timescale (see Fig. 11. Three-Dimensional Illustration of Lignocellulose Meshwork, this page). Hence, atomistic MD has moved well beyond the single-molecule level to permit systems-level simulation of hundreds of interacting biological macromolecules, such as those involving the transport of chemical signals across the cellular membrane. Indeed, an extrapolation of current performance at the petascale to the exascale indicates that researchers will ultimately be able to perform MD simulations of systems consisting of $\sim 10^{11}$ explicit interacting atoms (approximately the number of atoms in a bacterial cell) for about 10 microseconds. Atomistic simulations will provide information on the response of macromolecules to ligand binding and the diffusion and transport of metabolites and proteins in crowded cellular environments. Cellular events on a millisecond timescale or longer and system sizes beyond 100M atoms call for simulation methods more simplified than atomistic MD, averaging out the unimportant degrees of freedom to preserve long timescale properties. This "coarse-graining" will lead to a multiphysics description of biological phenomena. The challenge is to filter phenomena on short time and length scales that have mesoscopic consequences so that both important and trivial data are preserved in coarse graining, all while maintaining self-consistency. Coarse-grain methods scale efficiently on a variety of supercomputers and will permit cell-scale simulations on timescales up to one second. With this extended time, tracing the diffusion of macromolecules and metabolites across the cell is feasible, including in the crowded cellular environment, providing information on system-dependent diffusion constants and associations between multiple molecules in the cytoplasm and at membranes.

Handling spatial heterogeneity and efficiently simulating timescales on the order of the cell cycle (minutes to hours) requires stochastic modeling techniques for systems of biochemical reactions inside a cell. These techniques should allow scientists to address complex events such as signaling cascades,

Fig. 11. Three-Dimensional Illustration of Lignocellulose Meshwork. Researchers are using computational modeling to gain a molecular-level understanding of the plant cell wall and its major components, including cellulose fibers (green), lignin molecules (brown), and hemicellulose (light green). [Image courtesy Thomas Splettstoesser, www.scistyle.com, for Oak Ridge National Laboratory]

transcription, translation and degradation, biofilm formation, and cell division. Once a molecular-level 3D model of the cell is established, the evolution of this system will need to be followed over time, requiring multiresolution imaging capabilities. Techniques that can identify individual macromolecular and small-molecule species in the cell are needed, as is the ability to follow the distribution of these components with a high level of temporal resolution. Also required are techniques capable of following and detecting macromolecular interactions and the structures of their complexes *in vivo*.

IV. Cross-Cutting Themes

The Department of Energy's Office of Biological and Environmental Research (BER) is tasked with advancing knowledge of biological and environmental systems and providing facilities to support missions in scientific discovery and innovation, energy security, and environmental responsibility. More specifically, BER research includes disciplines that span many spatial and temporal scales—biology and physics interfaces, subsurface biogeochemistry affecting contaminant and nutrient cycling, sustainable biofuels development through the power of genomics and systems biology, and climate science research to develop powerful predictive models of energy use and climate change. These areas provide the research space for BER-relevant questions about Earth's systems that should be prioritized for study (see Fig. 12. Integration of Research and Department of Energy Facilities, this page).

Inherent in all the types of science discussed in this report are issues associated with crossing scales—spatial and temporal factors, discipline-unique assumptions, and data acquisition. The major impediment to parameterizing results from one discipline as it interfaces with another is the heterogeneity existing at every level: particle size and composition in aerosols, compositions and functions of microbial communities, geochemistry and weathering, genome dynamics, and evolutionary influences, to mention a few. If a simple average of normal distributions could accurately represent disciplines, precisely predictive models of Earth's systems would be completed. In reality, it is clear that steady state is an illusion of limited data. Kinetic and similar models rarely address the exceptions and "tails" of the data that might be essential for tipping points.[1] For example, when do single-site or small-site events need to be examined explicitly to understand the changes in the system? Data must be evaluated with full knowledge of sampling density, space, and frequency and

Fig. 12. Integration of Research and Department of Energy Facilities.

---

[1]"Tipping point" in this context is broadly defined as a region in time and space at which changes from one state to another state become important to the science under investigation.
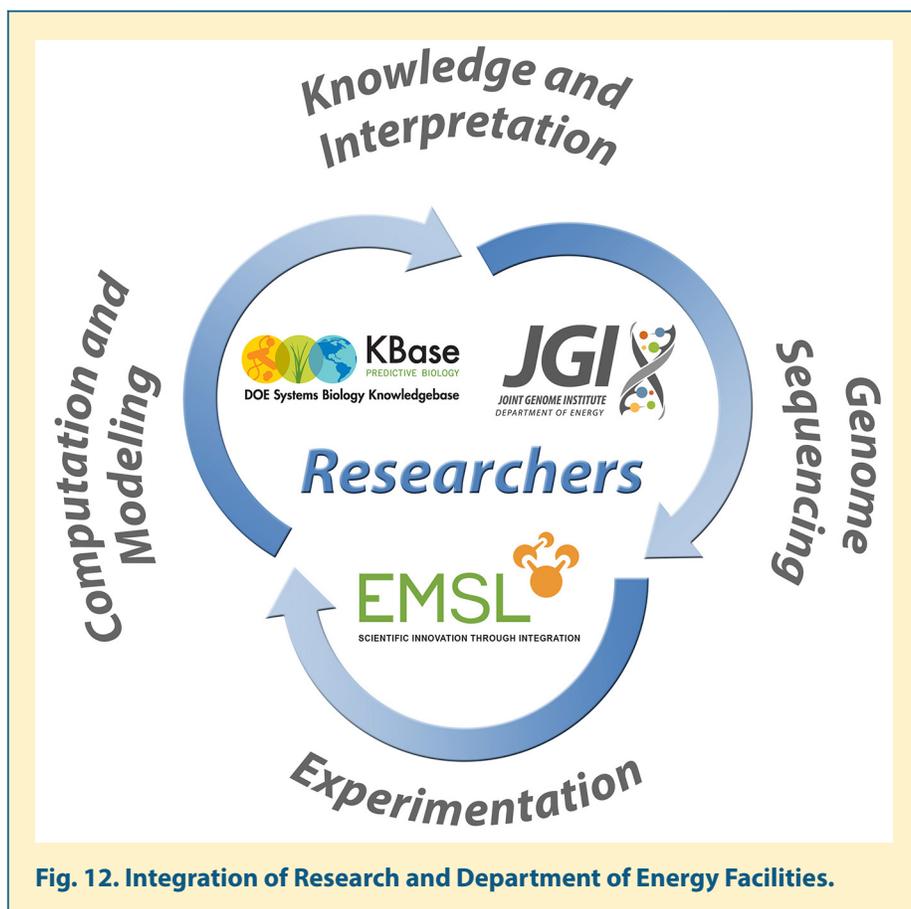
the recognition that nonequilibrium events and non-linear processes are abundant in the environment. Each discipline should inform those with which it interfaces, and the interfaces should be reciprocal and dynamic.

A theme that developed from workshop discussions is that it is neither possible nor practical to attempt to represent the full heterogeneity of the various scales in parameterizing models to predict interactions. Rather, to understand and predict how events at one scale affect those at a larger, encompassing scale often requires obtaining a bigger overview of the subjects. Such an approach helps elucidate the biogeochemical principles at play in the system. Moreover, insights derived from studies at much smaller or fundamental scales provide valuable clues to system behavior that facilitate robust parameterization of models at larger scales. These smaller-scale studies involve, for example, individual molecules, particles, or pores (e.g., ion speciation, rate constants, or the nature of mineral-microbe aggregates).

Of paramount importance to a discussion about crossing scales is an understanding of which processes are linked across scales, how they are linked, and the fundamental factors that govern them. Many processes operate only for short periods of time, such as during rainfall events, and may be highly localized. Thus, understanding the spatial and temporal behavior of cross-scale linkages also is essential. Based on these linkages, the critical physical, chemical, biogeochemical, and biological components can be assessed to determine which data must be included for predictive model development. What level of information is really required to address the scientific question being asked? Some questions will involve macromolecular information, such as large-scale eddies or plumes of contaminants from anthropogenic sources, but sometimes they will involve knowing the molecular interactions to understand tipping points in the environment.

As an example of seeking linkages, consider changing ecosystems, which provide an excellent opportunity to examine tipping points. Local shifts in soil hydration induced by climate change could affect the emission of chemicals into the atmosphere, either directly from the soil, microbes in the soil, or plant matter. These emissions could then impact the formation of aerosols and, eventually, cloud condensation nuclei. Note that the definition of "local" is vague and leads to additional questions. Does it refer to a state, a county, or a farm field? The emissions could be quite different depending on soil makeup, crops grown, and anthropogenic effects (such as oil drilling). Other questions include determining the smallest scale that must be examined to understand the tipping points and their effects on the environment. If molecular information is required, is such information needed for the full system or only for critical regions of time and space?

Another significant challenge is to make current and newly generated information and data from each domain (either as a discipline or spatial/temporal scale) available to other researchers in a facile manner that allows rapid data integration and knowledge generation.

The classical example of difficulty in integrating across scales is that of extrapolating between laboratory results and field studies. Addressing the fidelity between laboratory and real-world environments necessitates two types of research thrusts that are synergistic. The first thrust is to use the best available field knowledge to select relevant parameters that can mimic the complexity of the real-world environment in the laboratory with sufficient fidelity to reproduce behavior across controlled and uncontrolled environments. The second thrust is to develop methodology, frequently based on laboratory-derived insights, that enables qualitative and quantitative analyses in the native environmental setting.

At all levels of workshop discussion, there was complete consensus that predictive models are essential for integrating the masses of data that have and will certainly

continue to accumulate and likely increase in size and complexity. Thus, serious investment must be made in computational capacity to maximize the knowledge gained and to allow synergies to emerge from data gathered at different spatial and temporal scales.

Finally, to ensure that the achievements of this era of science will be the foundation of future discoveries, interdisciplinary education and workforce training for the next generation of early career scientists must be given a very high priority. Identifying talented and passionate scientists who will be the innovators and leaders of environmental research in the coming

decade already is hampered by the loss of individuals to less challenging careers. Opportunities for focusing on the scientific questions, and collaborative interactions to achieve integration across disciplines, should be designed to attract the best and brightest. Many young students want to "cure a disease" that will improve the lives of perhaps a few million people. However, if they help discover responsible and innovative pathways to sustainable energy, climate change adaptation and mitigation, or environmental remediation, they will contribute to a better world for all future generations.

V. References

Azam, F., and D. C. Smith. 1991. "Bacterial Influence on the Variability in the Ocean's Biogeochemical State: A Mechanistic View." In *Particle Analysis in Oceanography.* Springer-Verlag: Berlin Heidelberg. 213–36.

Bagley, J. E., et al. 2014. "The Biophysical Link Between Climate, Water, and Vegetation in Bioenergy Agro-Ecosystems," *Biomass and Bioenergy* **71**, 187–201. DOI: 10.1016/j.biombioe.2014.10.007.

Bengough, A. G. 2012. "Water Dynamics of the Root Zone: Rhizosphere Biophysics and Its Control on Soil Hydrology," *Vadose Zone Journal* **11**(2). DOI: 10.2136/vzj2011.0111.

BERAC. 2013. *BER Virtual Laboratory: Innovative Framework for Biological and Environmental Grand Challenges: A Report from the Biological and Environmental Research Advisory Committee*, DOE/SC-0156. U.S. Department of Energy, February 2013. 52 pp.

BERAC. 2010. *Grand Challenges for Biological and Environmental Research: A Long-Term Vision*: *A Report from the Biological and Environmental Research Advisory Committee March 2010 Workshop*, DOE/SC-0135. Biological and Environmental Research Advisory Committee, Steering Committee on Grand Research Challenges for Biological and Environmental Research. 82 pp.

Berendsen, R. L., C. M. Pieterse, and P. A. Bakker. 2012. "The Rhizosphere Microbiome and Plant Health," *Trends in Plant Science* **17**(8), 478–86. DOI: 10.1016/j.tplants.2012.04.001.

Bonnaud, P. A., B. Coasne, R. J.-M. Pellenq. 2010. "Molecular Simulation of Water Confined in Nanoporous Silica," *Journal of Physics: Condensed Matter* **22**(28), 284110–25. DOI: 10.1088/0953-8984/22/28/284110.

Borch, T., et al. 2010. "Biogeochemical Redox Processes and Their Impact on Contaminant Dynamics," *Environmental Science & Technology* **44**(1), 15–23. DOI: 10.1021/es9026248.

Boyer, E. W., et al. 1997. "Response Characteristics of DOC Flushing in an Alpine Catchment," *Hydrological Processes* **11**(12), 1635–47. DOI: 10.1002/(SICI)1099-1085(19971015)11:12<1635::AID-HYP494>3.0.CO;2-H.

Bridgham, S. D., et al. 1995. "Potential Feedbacks of Northern Wetlands on Climate Change," *BioScience* **45**(4), 262–74. DOI: 10.2307/1312419.

Brown, G. E., Jr., and G. Calas. October 2012 and January 2013. "Mineral-Aqueous Solution Interfaces and Their Impacts on the Environment," *Geochemical Perspectives* **1**(4–5), 276 pp. DOI: 10.7185/geochempersp.1.4.

Brown, G. E., Jr., and G. Calas. 2011. "Environmental Mineralogy – Understanding Element Behavior in Ecosystems," *Comptes Rendus Geoscience* **343**(2–3), 90–112. DOI: 10.1016/j.crte.2010.12.005.

Buckel, W., and R. K. Thauer. 2013. "Energy Conservation via Electron Bifurcating Ferredoxin Reduction and Proton/Na$^+$ Translocating Ferredoxin Oxidation," *Biochimica et Biophysica Acta (BBA) – Bioenergetics* **1827**(2), 94–113. DOI: 10.1016/j.bbabio.2012.07.002.

Bylaska, E. J., J. Q. Weare, and J. H. Weare. 2013. "Extending Molecular Simulation Time Scales: Parallel in Time Integrations for High-Level Quantum Chemistry and Complex Force Representations*," Journal of Chemical Physics* **139**(7), Article Number: 074114. DOI: 10.1062/1.4818328.

Campbell, K. M., et al. 2011. "Oxidative Dissolution of Biogenic Uraninite in Groundwater at Old Rifle, CO," *Environmental Science & Technology* **45**(20), 8748–54. DOI: 10.1021/es200482f. Erratum in: *Environmental Science & Technology* 2012 Aug. 21, **46**(16), 9199.

Cheng, H., E. Hu, and Y. Hu. 2012. "Impact of Mineral Micropores on Transport and Fate of Organic Contaminants: A Review," *Journal of Contaminant Hydrology, Special Issue* **129–30**, 80–90. DOI: 10.1016/J.JCONHYD.2011.09.008.

Choobari, O. A., P. Zawar-Reza, and A. Sturman. 2014. "The Global Distribution of Mineral Dust and Its Impacts on the Climate System: A Review," *Atmospheric Research* **138**, 152–65. DOI: 10.1016/j.atmosres.2013.11.007.

Chorover, J., et al. 2007. "Soil Biogeochemical Processes Within the Critical Zone," *Elements* **3**, 321–26. DOI: 10.2113/gselements.3.5.321.

DiChristina, T. J., J. K. Fredrickson, and J. M. Zachara. 2005. "Enzymology of Electron Transport: Energy Generation with Geochemical Consequences," *Reviews in Mineralogy & Geochemistry* **59**(1), 27–52. DOI: 10.2138/rmg.2005.59.3.

Donahue, N. M., et al. 2006. "Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics," *Environmental Science & Technology* **40**(8), 2635–43. DOI: 10.1021/es052297c.

Ehrenfeld, J. G., B. Ravit, and K. Elgersma. 2005. "Feedback in the Plant-Soil System," *Annual Review of Environment and Resources* **30,** 75–115. DOI: 10.1146/annurev.energy.30.050504.144212.

Engelbrecht, J. P., and E. Derbyshire. 2010. "Airborne Mineral Dust," *Elements* **6**(4), 241–46. DOI: 10.2113/gselements.6.4.241.

Eusterhues, K., et al. 2014. "Biodegradation of Ferrihydrite-Associated Organic Matter," *Biogeochemistry* **119**(1–3), 45–50. DOI: 10.1007/s10533-013-9943-0.

Fliessbach, A., S. Sarig, and Y. Steinberger. 1994. "Effects of Water Pulses and Climatic Conditions on Microbial Biomass Kinetics and Microbial Activity in a Yermosol of the Central Negev," *Arid Soil Research and Rehabilitation* **8**(4), 353–62. DOI: 10.1080/15324989409381409.

Fuller, C. C., and J. W. Harvey. 2000. "Reactive Uptake of Trace Metals in the Hyporheic Zone of a Mining-Contaminated Stream, Pinal Creek, Arizona," *Environmental Science & Technology* **34**(7), 1150–55. DOI: 10.1021/es990714d.

Gilbert, B., and J. F. Banfield. 2005. "Molecular-Scale Processes Involving Nanoparticulate Minerals in Biogeochemical Systems." Ed. J. E. Banfield, J. CerviniSilva, and K. H. Nealson, *Reviews in Mineralogy and Geochemistry* **59**(1), 109–55. DOI: 10.2138/rmg.2005.59.6.

Gottlieb, B., L. K. Beitel, and M. Trifiro. 2014. "Changing Genetic Paradigms: Creating Next-Generation Genetic Databases as Tools to Understand the Emerging Complexities of Genotype/Phenotype Relationships," *Human Genomics* **8**, 1–9. DOI: 10.1186/1479-7364-8-9.

Guenther, A., et al. 2012. "The Model of Emissions of Gases and Aerosols from Nature Version 2.1 (MEGAN2.1): An Extended and Updated Framework for Modeling Biogenic Emissions," *Geoscientific Model Development* **5**, 1471–92. DOI: 10.5194/gmd-5-1471-2012.

Guenther, A., et al. 2006. "Estimates of Global Terrestrial Isoprene Emissions Using MEGAN (Model of Emissions of Gases and Aerosols from Nature)," *Atmospheric Chemistry and Physics* **6**, 3181–210. DOI: 10.5194/acp-6-3181-2006.

Guenther, A., et al. 1995. "A Global Model of Natural Volatile Organic Compound Emissions," *Journal of Geophysical Research* **100**(D5), 8873–92. DOI: 10.1029/94JD02950.

Hedin, L. O., et al. 1998. "Thermodynamic Constraints on Nitrogen Transformations and Other Biogeochemical Processes at Soil-Stream Interfaces," *Ecology* **79**(2), 684–703. DOI: 10.1890/0012-9658(1998)079[0684:TCONAO]2.0.CO;2).

Herrmann, G., et al. 2008. "Energy Conservation via Electron-Transferring Flavoprotein in Anaerobic Bacteria," *Journal of Bacteriology* **190**(3), 784–91. DOI: 10.1128/JB.01422-07.

Higashide, W., et al. 2012. "Metabolic Engineering of *Clostridium cellulolyticum* for Production of Isobutanol from Cellulose," *Applied Environmental Microbiology* **78**(19), 7171 [corrects **77**(8), 2727–33. 2011. DOI: 10.1128/aem.02454-10]. DOI: 10.1128/aem.02335-12.

Hill, A. R., et al. 2000. "Subsurface Denitrification in a Forest Riparian Zone: Interactions Between Hydrology and Supplies of Nitrate and Organic Carbon," *Biogeochemistry* **51**(2), 193–223. DOI: 10.1023/A:1006476514038.

Hinckley, E.-L., et al. 2014. "Digging into the World Beneath Our Feet: Bridging Across Scales in the Age of Global Change," *Eos, Transactions American Geophysical Union* **95**(11), 96–97. DOI: 10.1002/2014EO110004.

Hoffmann, T., et al. 1997. "Formation of Organic Aerosols from the Oxidation of Biogenic Hydrocarbons," *Journal of Atmospheric Chemistry* **26**, 189–222. DOI: 1023/A:1005734301837.

Holmes, K. W., et al. 2005. "Multi-Scale Variability in Tropical Soil Nutrients Following Land-Cover Change," *Biogeochemistry* **74**(2), 173–203. DOI: 10.1007/s10533-004-3544-x.

Jelier, R., et al. 2011. "Predicting Phenotype Variation in Yeast from Individual Genome Sequences," *Nature Genetics* **43**, 1270–74. DOI: 10.1038/ng.1007.

Köfinger, J., G. Hummer, and C. Dellago. 2009. "Macroscopically Ordered Water in Nanopores," *Proceedings of the National Academy of Sciences USA* **105**(36), 13218–22. DOI: 10.1073/pnas.0801448105.

Koop, T., et al. 2011. "Glass Transition and Phase State of Organic Compounds: Dependency on Molecular Properties and Implications for Secondary Organic Aerosols in the Atmosphere," *Physical Chemistry Chemical Physics* **13**(43), 19238–55. DOI: 10.1039/c1cp22617g.

Koven, C. D., et al. 2013. "The Effect of Vertically Resolved Soil Biogeochemistry and Alternate Soil C and N Models on C Dynamics of CLM4," *Biogeosciences* **10**, 7109–73. DOI: 10.5194/bg-10-7109-2013.

Laanbroek, H. J. 2010. "Methane Emission from Natural Wetlands: Interplay Between Emergent Macrophytes and Soil Microbial Processes. A Mini-Review," *Annals of Botany* **105**(1), 141–53. DOI: 10.1093/aob/mcp201.

Lehner, B. 2013. "Genotype to Phenotype: Lessons from Model Organisms for Human Genetics," *Nature Reviews Genetics* **14**, 168–78. DOI: 10.1038/nrg3404.

Lemak, S., et al. 2013. "Toroidal Structure and DNA Cleavage by the CRISPR-Associated [4Fe-4S] Cluster Containing Cas4 Nuclease SSO0001 from *Sulfolobus solfataricus*," *Journal of the American Chemical Society* **135**, 17476–487. DOI: 10.1021/ja408729b.

Lemanceau, P., et al. 2009. "Iron Dynamics in the Rhizosphere as a Case Study for Analyzing Interactions Between Soils, Plants and Microbes," *Plant and Soil* **321**(1–2), 513–35. DOI: 10.1007/s11104-009-0039-5.

Li, F., et al. 2008. "Coupled Ferredoxin and Crotonyl Coenzyme A (CoA) Reduction with NADH Catalyzed by the Butyryl-CoA Dehydrogenase/Etf Complex from *Clostridium kluyveri*," *Journal of Bacteriology* **190**(3), 843–50. DOI: 10.1128/JB.01417-07.

Lipson, D. A., et al. 2013. "The Contribution of Fe(III) and Humic Acid Reduction to Ecosystem Respiration in Drained Thaw Lake Basins of the Arctic Coastal Plain," *Global Biogeochemical Cycles* **27**(2), 399–409. DOI: 10.1002/gbc.20038.

Lipson, D. A., et al. 2012. "Water-Table Height and Micro-topography Control Biogeochemical Cycling in an Arctic Coastal Tundra Ecosystem," *Biogeosciences* **9**(1), 577–91. DOI: 10.5194/bg-9-577-2012.

Lundberg, D. S., et al. 2012. "Defining the Core *Arabidopsis thaliana* Root Microbiome," *Nature* **488**(7409), 86–90. DOI: 10.1038/nature11237.

McClain, M. E., et al. 2003. "Biogeochemical Hot Spots and Hot Moments at the Interface of Terrestrial and Aquatic Ecosystems," *Ecosystems* **6**(4), 301–12. DOI: 10.1007/s10021-003-0161-9.

Mendes, R., et al. 2011. "Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria," *Science* **332**(6033), 1097–1100. DOI: 10.1126/science.1203980.

Michot, L. J., et al. 2002. "Water Organisation at the Solid-Aqueous Solution Interface," *Comptes Rendus Geoscience* **334**(9), 611–31. DOI: 10.1016/S1631-0713(02)01801-1.

Mitchell, P., and J. Moyle. 1967. "Chemiosmotic Hypothesis of Oxidative Phosphorylation," *Nature* **213**(5072), 137–39. DOI: 10.1038/213137a0.

Murphy, E. M., J. M. Zachara, and S. C. Smith. 1990. "Influence of Mineral-Bound Humic Substances on the Sorption of Hydrophobic Organic Compounds," *Environmental Science & Technology* **24**(10), 1507–16. DOI: 10.1021/es00080a009.

Nelson, Y. M., et al. 2002. "Effect of Oxide Formation Mechanisms on Lead Adsorption by Biogenic Manganese (Hydr)oxides, Iron (Hydr)oxides, and Their Mixtures," *Environmental Science & Technology* **36**(3), 421–25. DOI: 10.1021/es010907c/.

Nitschke, W., and M. J. Russell. 2012. "Redox Bifurcations: Mechanisms and Importance to Life Now, and at Its Origin: A Widespread Means of Energy Conversion in Biology Unfolds," *Bioessays* **34**(2), 106–09. DOI: 10.1002/bies.201100134.

Odum, J. R., et al. 1997. "The Atmospheric Aerosol-Forming Potential of Whole Gasoline Vapor," *Science* **276**(5309), 96–99. DOI: 10.1126/science.276.5309.96.

Odum, J. R., et al. 1996. "Gas/Particle Partitioning and Secondary Organic Aerosol Yields," *Environmental Science & Technology* **3**(8), 2580–85. DOI: 10.1021/es950943+.

Paerl, H. W, and J. L. Pinckney. 1996. "A Mini-Review of Microbial Consortia: Their Roles in Aquatic Production and Biogeochemical Cycling," *Microbial Ecology* **31**(3), 225–47. DOI: 10.1007/BF00171569.

Park, S., K. Kwak, and M. D. Fayer. 2007. "Ultrafast 2D-IR Vibrational Echo Spectroscopy: A Probe of Molecular Dynamics," *Laser Physics Letters* **4**(10), 704–18. DOI: 10.1002/lapl.200710046.

Philippot, L., et al. 2013. "Going Back to the Roots: The Microbial Ecology of the Rhizosphere," *Nature Reviews Microbiology* **11**(11), 789–99. DOI: 10.1038/nrmicro3109.

Prather, K. A., C. D. Hatch, and V. H. Grassian. 2008. "Analysis of Atmospheric Aerosols," *Annual Review of Analytical Chemistry* **1**, 485-514. DOI: 10.1146/annurev.anchem.1.031207.113030.

Ro, D.-K., et al. 2006. "Production of the Antimalarial Drug Precursor Artemisinic Acid in Engineered Yeast," *Nature* **440**, 940943. DOI: 10.1038/nature04640.

Schade, J. D., et al. 2001. "The Influence of a Riparian Shrub on Nitrogen Cycling in a Sonoran Desert Stream," *Ecology* **82**(12), 3363–76. DOI: 10.1890/0012-9658(2001)082[3363:TIOARS]2.0.CO;2.

Schenk, P. M., L. C. Carvalhais, and K. Kazan. 2012. "Unraveling Plant-Microbe Interactions: Can Multi-Species Transcriptomics Help?" *Trends in Biotechnology* **30**(3), 177–184. DOI: 10.1016/j.tibtech.2011.11.002.

Shimojo, F., et al. 2014. "A Divide-Conquer-Recombine Algorithmic Paradigm for Large Spatiotemporal Quantum Molecular Dynamics Simulations," *The Journal of Chemical Physics* **140**, 18A529. DOI: 10.1063/1.4869342.

Six, J., E. T. Elliott, and K. Paustian. 2000. "Soil Macroaggregate Turnover and Microaggregate Formation: A Mechanism for C Sequestration Under No-Tillage Agriculture," *Soil Biology and Biochemistry* **32**(14), 2099–2103. DOI: 10.1016/S0038-0717(00)00179-6.

Smith, J. C., and B. Roux. 2013. "Eppur si muove! The 2013 Nobel Prize in Chemistry," *Structure* **21**(12), 2102–05. DOI: 10.1016/j.str.2013.11.005.

Solveyra, E. G., et al. 2013. "Structure, Dynamics and Phase Behavior of Water in TiO_2 Nanopores," *Journal of Physical Chemistry C* **117**(7), 3330–42. DOI: 10.1021/jp3079900q.

Stockdale, A., W. Davison, and H. Zhang. 2009. "Micro-Scale Biogeochemical Heterogeneity in Sediments: A Review of Available Technology and Observed Evidence," *Earth-Science Reviews* **92**(1), 81–97. DOI: 10.1016/jearscirev.2008.11.003.

Torn, M. S., et al. 1997. "Mineral Control of Soil Organic Carbon Storage and Turnover," *Nature* **389**(6647), 170–73. DOI: 10.1038/38260.

Vidon, P., et al. 2010. "Hot Spots and Hot Moments in Riparian Zones: Potential for Improved Water Quality Management," *Journal of the American Water Resources Association* **46**(2), 278–98. DOI: 10.1111/j.1752-1688.2010.00420.x.

von Lützow, M., et al. 2006. "Stabilization of Organic Matter in Temperate Soils: Mechanisms and Their Relevance Under Different Soil Conditions – A Review," *European Journal of Soil Science* **57**(4), 426–45. DOI: 10.1111l/j.1365-2389.2006.00809.x.

WEF. 2014. *Global Risks 2014: Ninth Edition.* World Economic Forum, Geneva, Switzerland. 59 pp. ISBN-13: 92-95044-60-6. ISBN-10: 978-92-95044-60-9. www3.weforum.org/docs/WEF_GlobalRisks_Report_2014.pdf.

Weston, D. J., et al. 2014. "*Sphagnum* Physiology in the Context of Changing Climate: Emergent Influences of Genomics, Modeling and Host-Microbiome Interactions on Understanding Ecosystem Function," *Plant, Cell, & Environment*, Early View. DOI: 10.1111/pce.12458.

Wilhelm, B. G., et al. 2014. "Composition of Isolated Synaptic Boutons Reveals the Amounts of Vesicle Trafficking Proteins," *Science*, **344**(6187), 1023–28. DOI: 10.1126/science.1252884.

Williams, K. H., et al. 2013. "Bioremediation of Uranium-Contaminated Groundwater: A Systems Approach to Subsurface Biogeochemistry," *Current Opinion in Biotechnology* **24**(3), 489–97.

Young, I. M., et al. 2008. "Microbial Distribution in Soils: Physics and Scaling," *Advances in Agronomy* **100,** 81–121. DOI: 10.1016/S0065-2113(08)00604-4.

Yvert, G. 2014. "'Particle Genetics': Treating Every Cell as Unique," *Trends in Genetics* **30**(2), 49–56. DOI: 10.1016/j.tig.2013.11.002.

Zaady, E., P. M. Groffman, and M. Shachak. 1996. "Litter as a Regulator of N and C Dynamics in Macrophytic Patches in Negev Desert Soils," *Soil Biology and Biochemistry* **28**(1), 39–46. DOI: 10.1016/0038-0717(95)00097-6.

Appendices

Appendix A: Purpose

U.S. Department of Energy Office of Science
Office of Biological and Environmental Research (BER)
Proposed BER Molecular Science Challenges Workshop

BER research, which spans biological, environmental, and climate sciences, has evolved over recent years to require a much more robust understanding of the molecular systems and processes that underpin program goals. This workshop will seek to identify science and technology challenges and opportunities relevant to BER's mission. Overcoming these challenges and exploring the opportunities will expand our ability to understand, describe, and model molecular-scale processes based on a synergistic and multidisciplinary approach that is relevant to a wider set of BER challenges.

Workshop participants will represent the major program elements of BER that depend on molecular science and include leading scientists with relevant expertise who are not associated with BER. Participants will identify major BER challenges in molecular science and develop high-level progressions of scientific objectives that address each challenge. The workshop will develop an expert assessment of these challenges, objectives, and research pathways to overcome barriers in BER-relevant molecular science. The workshop report will be used by BER to plan its long-term investment strategy over a time horizon from 2014 through 2024.

BER Interests

Progress in research across the BER programs will require significant advances in molecular science. For example, characterizing the spatiotemporal expression, structure, and function of biological molecules and macromolecules in cells and among communities of cells, as well as the computational modeling of such systems, is an essential foundation for understanding basic cellular and organismal processes. This understanding, in turn, will be the critical element in achieving meaningful success in biosystem design.

Similarly, molecular-scale knowledge of biotic and abiotic environmental factors impacting genotype-to-phenotype linkage in plants will be essential for predicting plant growth in changing environments, designing plants with improved phenotypic traits, and predicting ecological changes among multiple plant species in complex hydro-biogeochemical soil systems.

Experimental and modeling capabilities in molecular science will enable the elucidation of the physical, chemical, and biogeochemical processes that govern the formation and evolution of aerosol particles and their interactions with clouds. These advances in molecular science also will provide predictive insights into the interactions among heterogeneous populations of aerosols and cloud droplets.

Progress in understanding carbon cycling will require knowledge of the biogeochemical interactions between microbes, fungi, and plant roots and macro- and micro-nutrients, organic constituents, and inorganic elements under varying conditions in surface soils, hyporheic zones, and the deeper subsurface.

These four topical areas have served as BER's historical paradigm for investing in molecular science research. However, new challenges involving, for example, molecular-scale processes that govern interdependencies between biogenic aerosol formation and ecological functioning, also will need to be explored in the workshop.

The workshop will be expected to define a set of molecular science questions and priorities for BER to consider in its future multidisciplinary investment strategies.

Appendix B: Agenda

U.S. Department of Energy (DOE) Office of Science
Biological and Environmental Research (BER) Workshop
BER Molecular Science Challenges
May 27–29, 2014

Tuesday, May 27

6:30 p.m.	Buffet dinner in hotel
7:00 p.m.	Welcome: Judy Wall (University of Missouri, Columbia), James Liao (University of California, Los Angeles)
7:10 p.m.	Welcome and BER organization: Paul Bayer (BER)
7:20 p.m.	Welcome and charge: Roland Hirsch (BER)
7:30 p.m.	Agenda and mechanics of the workshop: Judy Wall and James Liao
7:45 p.m.	Introductions: Participants (name, institution, title, and area of expertise or research)

Wednesday, May 28

7:30 a.m. – 7:45 a.m.	Transport from hotel to DOE Germantown building
7:45 a.m. – 8:15 a.m.	DOE security check
8:15 a.m. – 8:30 a.m.	Break
8:30 a.m. – 8:40 a.m.	**Welcome**: Sharlene Weatherwax, Associate Director of Science for BER; BER Division Directors, Todd Anderson and Gary Geernaert
8:40 a.m. – 9:10 a.m.	**Keynote: Carbon and Contaminants in the Critical Zone** Jon Chorover, University of Arizona
9:10 a.m. – 9:15 a.m.	**Review of breakout instructions and assignment of groups:** Judy Wall, James Liao (see Appendix C: Breakout Groups, p. 57)
9:15 a.m. – 10:15 a.m.	**Breakout 1: Identify key needs** **a.** What are three key unanswered questions in your research area? **b.** What are two or more scientific areas that scale up or down from your own? **c.** What new capabilities are needed to address your key questions?
10:15 a.m. – 10:30 a.m.	Break
10:30 a.m. – 12:00 p.m.	**Breakout 1 continues**
12:15 p.m. – 1:15 p.m.	Lunch at cafeteria
1:15 p.m. – 2:00 p.m.	**Reports from Breakout 1**
2:00 p.m. – 3:45 p.m.	**Breakout 2: Organize ideas within and across disciplines and groups**
3:45 p.m. – 4:00 p.m.	Break
4:00 p.m. – 5:00 p.m.	**Reports from Breakout 2**
5:00 p.m. –5:30 p.m.	**General Discussion:** Common interests and ideas across the groups
Dinner	**Small Group Discussions:** Continue discussions to develop ideas and consider links among topics. Multiple restaurants, on your own or by breakout group.

Thursday, May 29

7:30 a.m. – 7:45 a.m.	Transport from hotel to DOE Germantown building
7:45 a.m. – 8:15 a.m.	DOE security check
8:15 a.m. – 8:30 a.m.	Break
8:30 a.m. – 10:00 a.m.	**Breakout 3: Develop plans for sections of the report**
10:00 a.m. – 10:15 a.m.	Break
10:15 a.m. – 10:45 a.m.	**Reports from Breakout 3**
10:45 a.m. – 11:30 a.m.	**Summary discussion of key workshop findings**: Co-Chairs
11:30 a.m. – 12:55 p.m.	**Writing Breakout Session**
	Prepare summary chapters for report and drafts of potential journal papers
12:55 p.m.	**Concluding Remarks:** Co-Chairs
1:00 p.m.	Adjourn
1:15 p.m.	Lunch and continued writing as travel plans permit

Appendix C: Breakout Groups

One **Atmosphere-Land Surface Interactions Involving Molecular Science (Room A184/6)**
Discussion Lead: Vicki Grassian
Rapporteur: Scott Bridgham
Karl Booksh
Rick Flagan
Mary Gilles
Sean McSweeney
Theresa Windus

Two **Below-Surface Interactions Involving Molecular Science (Room E301)**
Discussion Lead: Michael Thomashow
Rapporteur: John Bargar
Kirsten Hofmockel
Joel Kostka
James Kubicki
Albert Valocchi
Judy Wall

Three **Synthetic Science and Engineering Involving Molecular Science (Room F441)**
Discussion Lead: Norman Dovichi
Rapporteur: Michael Crowley
Michael Adams
James Liao
Stephen Long
Jeremy Smith
Ganesh Sriram

Appendix D: Participants

Michael Adams
University of Georgia

John Bargar
SLAC National Accelerator Laboratory

Karl Booksh
University of Delaware

Scott Bridgham
University of Oregon

Jon Chorover
University of Arizona

Michael Crowley
National Renewable Energy Laboratory

Norman Dovichi
University of Notre Dame

Rick Flagan
California Institute of Technology

Mary Gilles
Lawrence Berkeley National Laboratory

Vicki Grassian
University of Iowa

Kirsten Hofmockel
Iowa State University

Joel Kostka
Georgia Institute of Technology

James Kubicki
Pennsylvania State University

James Liao
University of California, Los Angeles

Stephen Long
University of Illinois, Urbana-Champaign

Sean McSweeney
Brookhaven National Laboratory

Jeremy Smith
Oak Ridge National Laboratory and
    University of Tennessee, Knoxville

Ganesh Sriram
University of Maryland Clark School of
    Engineering

Michael Thomashow
Michigan State University

Albert Valocchi
University of Illinois, Urbana-Champaign

Judy Wall
University of Missouri, Columbia

Theresa Windus
Iowa State University

Acronyms and Abbreviations

3D	**three dimensional**
ATP	**adenosine triphosphate**
BER	**DOE Office of Biological and Environmental Research**
BERAC	**Biological and Environmental Research Advisory Committee**
BVOC	**biogenic volatile organic compound**
CH$_4$	**methane**
Chip-SIP	**phylogenetic microarray stable isotope probing**
CO$_2$	**carbon dioxide**
DOE	**U.S. Department of Energy**
EXAFS	**extended X-ray absorption fine structure**
Fe	**iron**
FISH	**fluorescence *in situ* hybridization**
H$_2$S	**hydrogen sulfide**
KBase	**DOE Systems Biology Knowledgebase**
LAI	**leaf area index**
MD	**molecular dynamics**
MEGAN	**Model of Emissions of Gases and Aerosols from Nature**
micro-XANES	**microprobe X-ray absorption near edge structure**
N	**nitrogen**
nano-SIMS	**nanometer-scale secondary ion mass spectrometry**
N$_2$O	**nitrous oxide**
NOM	**natural organic matter**
P	**phosphorus**
PFT	**plant functional type**
SAXS	**small-angle X-ray scattering**
SOA	**secondary organic aerosol**
SOM	**soil organic matter**
U	**uranium**
WEF	**World Economic Forum**