Vision 2020

Technology Roadmap for

Combinatorial Methods

September 2001

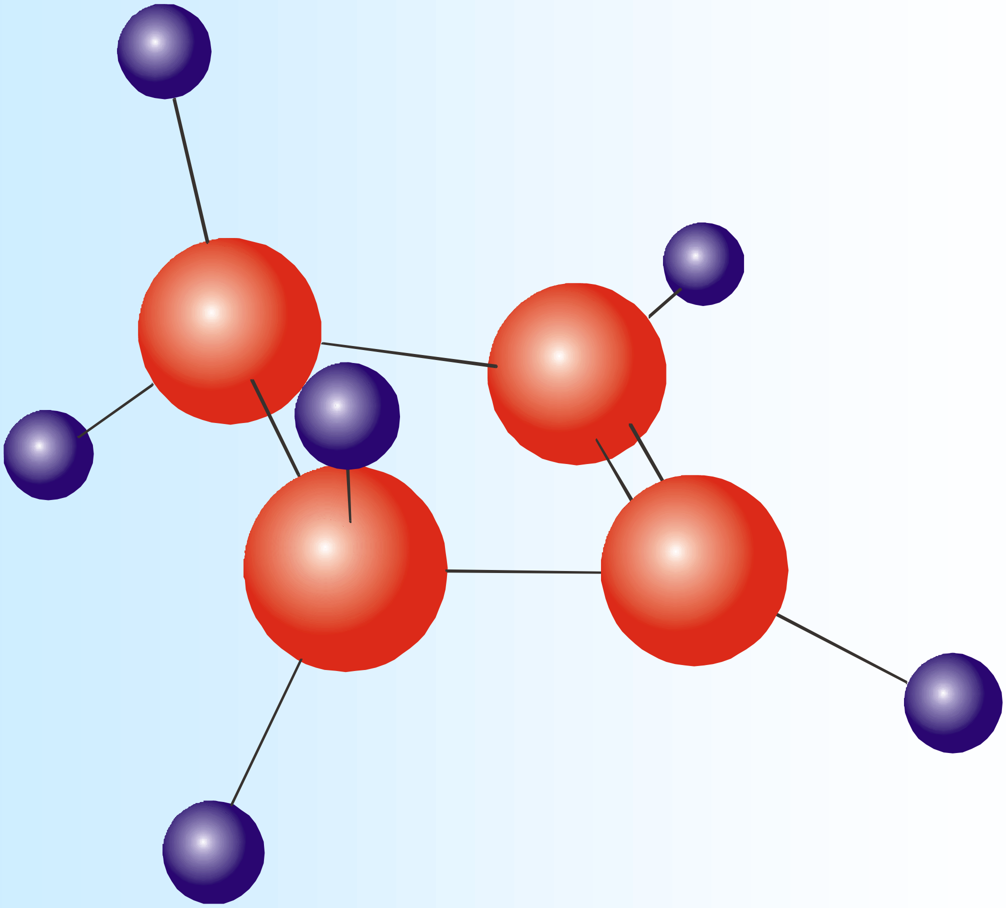


Table of Contents

1 Overview

Meeting the Challenges of the 21^{st} Century

The 21^{st} century brings many economic, environmental and societal challenges to the chemical industry. Major drivers for change include market globalization, societal demand for improved environmental performance, profitability and capital productivity, customer expectations, and changing workforce requirements.

The U.S. chemical industry has outlined a vision of how it can meet the competitive challenges of the future in *Technology Vision 2020: The U.S. Chemical Industry.*[1] *Vision 2020,* published in 1996, was developed in response to a request from the White House Office of Science and Technology Policy (OSTP) in the early 1990's. At that time, OSTP was seeking industry's advice on how to better allocate Federal R&D funds to advance the manufacturing capability and competitive position of the U.S. chemical industry.

Vision 2020 predicts that the next millennium will see a chemical industry that promotes sustainable development by investing in technology that protects the environment and stimulates industrial growth, while balancing economic needs and financial constraints. To accomplish this, *Vision 2020* calls for the industry to set the standard for the efficient use of energy and raw materials, and to work in seamless partnerships. A key tenet of *Vision 2020* is that the chemical industry's growth and competitive advantage *"depends upon individual and collaborative efforts of industry, government, and academe to improve the nation's R&D enterprise."*

These far-reaching goals for improved productivity, cost-effectiveness, energy use, and environmental performance will require that the industry make advances in critical technological areas of new chemical sciences and engineering (see Exhibit 1-1). As part of the strategy for achieving technological change, the chemical industry is developing technology roadmaps in a number of crucial areas. Technology roadmaps link the strategic goals outlined in *Vision 2020* with a detailed research agenda of near-, mid- and long-term technology R&D. Through technology roadmaps, the industry hopes to provide a way for decision-makers to make strategically-driven investments in R&D, and to build the partnerships needed to reach goals for growth and competitiveness.

Chemical Industry Technology Roadmaps (Completed or In Progress)

- Computational Chemistry
- Computational Fluid Dynamics
- Separations
- Materials of Construction in the CPI
- Reaction Engineering
- Biocatalysis
- Materials Technology
- Catalysis
- Process Measurement & Control
- New Process Chemistry

---

[1] Available from the American Chemical Society, Washington, D.C., (202) 452-8917.

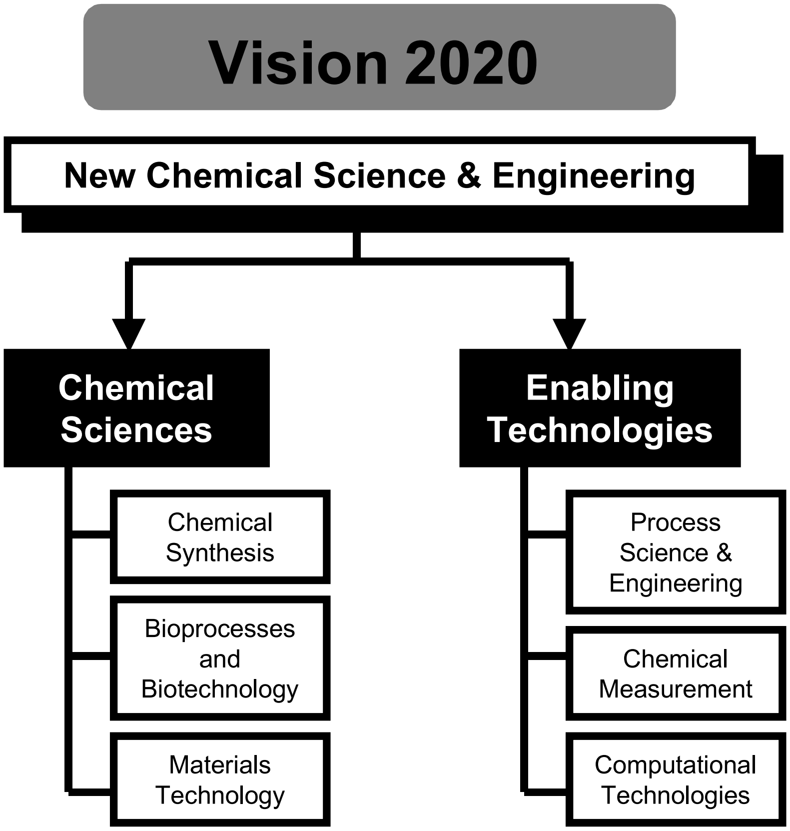


Exhibit 1-1. Selected R&D Areas Identified by Vision 2020 for New Chemical Science and Engineering

The Role of Combinatorial Chemistry

The use of combinatorial chemistry, which is the focus of this report, is cited throughout *Vision 2020* as an area of underlying importance to many areas of new chemical science and engineering. Combinatorial methods will be essential for making advances in chemical synthesis, where a key challenge is the development of new synthetic techniques, and specifically tools that permit the rapid creation of unique molecules. Combinatorial methods are also integral to the development of new catalysts and reaction systems, another area cited by *Vision 2020* where technological advances will be crucial to increased competitiveness. Other important areas in *Vision 2020* that will be impacted by combinatorial methods include the design of new, highly functional materials, biocatalysis, chemical measurement for molecular processes, rapid selection of process chemistry, and complex, multi-site process simulation tools.

Like computational methods, combinatorial chemistry is an enabling tool that can be used to accelerate the discovery process. Over the last ten years it has played an increasingly important role in the screening and discovery of new drugs, and is now beginning to be applied to areas outside the pharmaceutical industry. It is a methodology of great opportunity, and there is growing recognition in the chemical and materials industries that combinatorial and high throughput techniques have the potential to revolutionize the design process for new chemicals, materials, and catalysts. Combinatorial methodologies for discovery could shorten the discovery (e.g., research) phase of R&D from 2 to 3 years to 3 to 6 months; combinatorial methodologies for engineering scale up (not practiced generally today) and accelerated materials aging and validation (not practiced today) could shorten the development phase of R&D that moves from lab success to product on shelves from 2 to 5 years to 6 to 12 months. The advantages are clear—increased competitiveness, reduced time-to-market, and greater potential for the discovery of high performance, revolutionary new products.

In support of *Vision 2020*, the National Institute of Standards and Technology (NIST), in cooperation with the Council for Chemical Research (CCR), held a technology roadmap workshop in June 2000 to address the topic of combinatorial chemistry and how it can be applied to materials design. The workshop was attended by participants from industry, government, academia, and the national laboratory complex (see Appendix A for a complete list of participants). Four separate areas of technology were addressed—library production, library characterization, informatics, and design of experiment and models. The information and ideas generated at the workshop form the basis for this technology roadmap. It is hoped that the priorities identified here for combinatorial chemistry will enable decision-makers to support the critical research needed to help the U.S. chemical industry meet the challenges of the 21st century.

This roadmap is a dynamic document, and will be reevaluated periodically to incorporate new market and technical information, and to ensure that research priorities keep pace with the needs of both the chemical industry and its customers.

2 Current Capabilities and Trends in Combinatorial Chemistry

Overview

“Combinatorial Methodology” is a set of tools and techniques the chemicals and materials science communities can use to accelerate the methods by which knowledge is discovered and products and processes developed to meet the advanced materials needs of the 21st century. It uses a large number of carefully designed, multi-dimensional experiments that may be performed rapidly or in parallel, on a miniaturized scale using automated instrumentation.

Combinatorial approaches are not new, being traceable to a report by Bruce Merrifield in 1963 on solid-phase peptide synthesis, for which he was awarded the 1984 Nobel Prize in chemistry. In the 80’s, the combinatorial concept was catalyzed by Arpad Furka’s introduction of split and pool concepts and H. Mario Geysen’s report on development of an encoding strategy in solid phase synthesis. These early efforts formed the foundation for the combinatorial methods in use today in the pharmaceutical and agrochemical arenas.

The combinatorial race in the materials development arena was started by an “early” experiment by P. Schultz and X.D. Xiang (1995) that grew out of a curiosity whether combinatorial chemistry could be used for the discovery of interesting inorganic compounds—and perhaps to create a race of super-materials. Thin-film libraries were deposited using precursor oxides or cuprates of the seven elements that made up known high-temperature superconductors. These were then thermally processed to create crystalline films. Superconductivity was observed! This launched combinatorial efforts aimed at a wide variety of materials, including magneto-resistive materials, phosphors, dielectrics, ferroelectrics, polymers, semiconductors, catalysts, and zeolites.

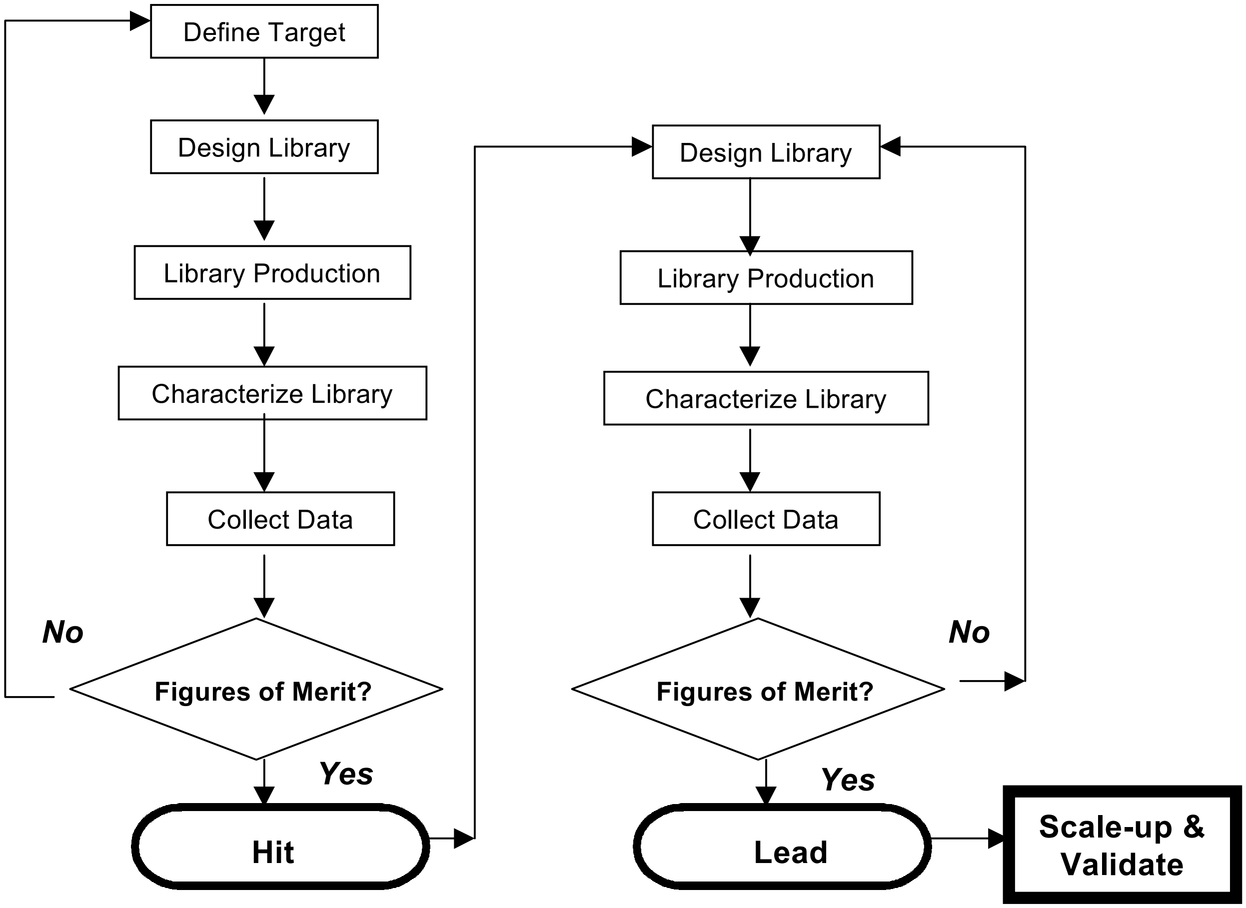
Major forces—globalization of markets and the pace of technology change—continue to drive private sector R&D to narrower, shorter-term investments to maximize returns to the company. Increased pressure on manufacturers to produce “faster/lower-cost/better” has increased the demand for new products made from new materials and/or utilizing new processes while utilizing more environmentally-friendly materials and processes. These market factors are now driving a need for reducing the cycle time for the discovery of new advanced materials and lower-cost chemical products. The technology spill-over from drug discovery has resulted in the embryonic development of combinatorial methodologies and technologies for inorganic materials and non-pharmaceutical organic compounds and materials.

A process flow for high-throughput screening is shown in Exhibit 2-1. As illustrated, the technology associated with combinatorial methods can be divided into four steps that follow target definition:

- Library design with computational inputs such as quantitative structure-property relationships (QSPR), and molecular modeling. **Design of Experiments** (DOE) is required to reduce the number of samples that will be necessary to define sample spaces within the experimental universe or to direct screening to other spaces within the universe. Data from previous screening runs is often used in a feedback loop;

- **Library Production** involves the automated deposition and/or processing of an n-dimensional matrix of physical samples;

Exhibit 2-1. The High Throughput Process Flow



- **Library Characterization** in a parallel or massively parallel mode involves the use of robotics and sensors to rapidly and automatically analyze the library of targets for desired properties;

- Data collection and analysis using the database and artificial intelligence tools—"**informatics**"—expanded into the more complex realm of materials properties.

Design of Experiment and Modeling

Experimental design has traditionally been concerned with experiments that involve a finite number of levels and factors. Methods have been developed that are capable of working with relatively smooth response surfaces, resulting in a widespread use of DOE to improve existing processes and to evaluate developmental aspects of new discoveries.

Coupled to experimental design are models that associate performance criteria with compositional, structural, and process parameters of the chemical systems of interest. By incorporating increasingly powerful and quantitative models at the design phase of the experiments, it will be possible to explore those portions of parameter space with the greatest potential for success.

The opportunity to apply modern design concepts in a combinatorial environment presents a far broader range of variables than in current sequential approaches. Spanning the experimental parameter space with sound sampling methods and valid replication and duplication concepts will require efforts beyond current expertise. Design of experiment methods in a combinatorial approach will not be assured of the smooth response surfaces associated with their success in sequential measurements. In addition, combinatorial designs will have to account for the constraints of the physical measurement approach, such as the achievable compositions, temperature isolation, and degree of mixing. New constraints and limitations may apply in many combinatorial undertakings, such that novel approaches will be essential for the success of this approach.

Current models have limited predictive capabilities. With the emergence of ever-faster computers, it will become possible to include sufficiently detailed models to make confident predictions. Initial approaches will focus on interpolative schemes, but with further development and testing, the extrapolative utility of the models can be exploited.

Library Production

Many methods have used to date in materials and catalyst applications of combinatorial methods. While split and pool approaches have been noticeably absent, the utilization of discrete and gradient libraries have received considerable interest. Within these two library classes, examples have emerged that span compositional variation, thermal treatment variation, and film thickness variation. Preparation techniques have included using chemical vapor deposition, laser ablation, flow coating, temperature gradients, and selective masking. Construction of solid state material libraries currently requires automated deposition onto substrates, using, for example, micro jet, laser ablation, vacuum deposition (PVD, CVD), or micro-fluidics in a lab-on-chip application. Micro-scale methods have already been commercialized for health care diagnostics with liquid samples. Many of these methods are being looked at in terms of their applicability for specific problems in solid-state library synthesis.

Library Characterization

While the methodologies developed for drug discovery can be leveraged for other applications, the complexities of chemicals and materials discovery far out-weigh the current capabilities developed for drug discovery. New drug candidates (“leads”) are tested by chemical means such as gas chromatography, nuclear magnetic resonance spectrometry, mass spectroscopy, and incremental biological activity; biological activity to a single component of a lead mixture is adequate in many cases. Advanced materials, on the other hand, need to be characterized according to their ***performance***, and these analytical techniques are often peculiar to the targeted application area. For example, selectivity and conversion analysis for a new heterogeneous catalyst formulation will require the development of novel microscopic sensors for sampling the complex product/reactant mixtures, possibly in the gas phase, for a large number of (microscopic) metal oxide “samples”.

While organic chemical leads are being successfully analyzed at purity levels of 80% or less, typically advanced materials require very high purities to differentiate their performance. Industrial chemistry typically utilizes more energetic reaction environments than pharmaceuticals, with processes with temperature and pressure requirements of several hundred degrees and thousands of p.s.i. pressure, respectively. Finally, while drug discovery can utilize traditional chemical scale-up processes, advanced materials have been dealt significant challenges in “scalability” from microscopic to lab- or pilot-scale preparation while retaining the discovered properties.

The current technology relies on contact and non-contact methods of characterization and external control of process conditions. The chemical process industry (CPI) sector is moving toward smaller, more integrated sensing devices in process control of manufacturing sites. Automated library processing is especially challenging for new materials development since samples within a library may require different or non-equilibrium processing parameters across the library.

Informatics

Compared to the use of combinatorial methods for drug discover, materials science applications require substantially different capabilities. Consider, for example, the data management needs for effectively sampling the complex, multi-dimensional formulation/processing space inherent in materials science, and then correlating the sample characteristics selected for evaluation with the critical properties of the material in the target end use. Correlative methods integrate information and materials technologies, and are much more challenging than ever encountered in drug design. Additional challenges that differentiate materials science applications from those in the pharmaceutical industry include:

- more complex experimental parameter space;
- practical limits based on experimental data, repeatability and scalability;
- experimental design limits based on availability of reliable data, and
- more complex quantitative structure-property/activity/selectivity relationships

Integration

There are significant scientific and technological challenges to implement HT methods in advanced materials. These challenges can borrow from advances made in other areas such as drug discovery and bioinformatics, however in most cases the technology transfer is not straightforward. The transfer of traditional *serial* research methodologies to *parallel* methodologies will require the integration of otherwise diverse computational and characterization tools for library design (statistics, design of experiments), library fabrication (sample deposition, automation), characterization (sensors, automation, data acquisition), and informatics (data handling, visualization, decision tools). The scalability of the results obtained will require scientific and technological advances in the areas of interfacial effects and structure-property prediction.

Materials processing typically involves more energetic reaction environments than pharmaceutical processing, with temperature and pressure requirements as high as several hundred degrees and hundreds of Pascals. Micro-scale solid-state samples may also be subject to significant influence from the substrate onto which they are deposited—interfacial effects can produce phenomena that are not reproducible in bulk samples of manufacturing scale. Thus, advanced materials suffer from “scalability”, or differences in micro-scale to lab- or pilot-scale properties. Finally, solid-state compositions may develop into different (kinetically-controlled) metastable structures depending on processing and testing conditions. Therefore, sample libraries must undergo validation at every process step. The basic underlying software technology must be capable of defining profitable experimental spaces and visualization of complex data relationships, and of correlating target materials with properties to permit database queries from a broad spectrum of data mining engines and the development of structure-property relationships. This requires interfacing with data visualization tools at the back end and statistical experimental design engines on the front end while remaining compliant with enterprise-wide systems for knowledge management and maintaining control of experimental hardware.

3 Market Drivers and Key Applications

Market Drivers

As shown in Exhibit 3-1, the increased use of combinatorial libraries in commercial applications will be driven by factors affecting competitiveness (profit margins, process flexibility, time to market) as well as many external factors (consumer demand, societal concerns, regulatory climate). Competition in the global marketplace will drive the need to create cheaper, more efficient, and more flexible manufacturing processes, along with products that are higher in quality and performance. Companies may ultimately recognize combinatorial methods as a means to move products to market quicker, to make products cheaper, and to create entirely new and revolutionary products.

Exhibit 3-1. Market Drivers

Market Expansion/Competitiveness	**Technical**
• Globalization • Competition driving need for cheaper, more efficient and flexible manufacturing processes and better products • Financial expectations for profits • Opportunity of small companies to leap-frog • Ability of large companies to shift gears • Diversity of and shifts in markets • Decreasing percent of sales going to R&D • Cost of raw materials and waste handling • Getting products to market faster and cheaper • Faster generation of niche products • Defensive/offensive intellectual property factor • Cost of combinatorial equipment • Reduced human cost of experience • Customization of products • Utilization of unique technology • Creation of high-end products • Discovery of entirely new materials • Increased outsourcing • Ability to reduce amount of off-shore characterization and testing	• Need for multi-functional design tools • Equipment development and breakthroughs • Ability to improve quality and reduce quality control costs • Ability to deal with the large quantity of information • Miniaturization requirements • Cover a large number of expectations at once **Societal** • Greenhouse gas reduction/Kyoto protocol • Government regulations (environment and safety) • Demand for improved environmental performance • Increasing demand for materials that can be recycled • Increasing population • Increasing standard of living worldwide • Preservation of jobs • “Born with” abilities of coming generations • Generation of knowledge/education • Increased utilization of biomass • Shift from fossil fuels

Technical issues will also drive the use of combinatorial methods. Breakthroughs in instruments and equipment for combi chemistry will facilitate its use in more applications, making it routine rather than novel. Advances in characterization techniques could lower the costs of quality control and drive the use of combi methods by providing cheaper, precise measurements. Developing the capability to deal with large quantities of data, and successfully meeting the unique miniaturization requirements for combi chemistry will also help to drive its commercial use.

Societal factors could have a significant impact on the use of combi methods. For example, more stringent environmental regulations could drive the need to have more options for “green” materials and processes. Consumers could drive demand for new materials that can be recycled, but exhibit the same quality and performance of existing conventional materials. In future materials will increasingly be made from alternative raw materials, such as biomass, and processing may rely less on the use of fossil fuels. Combinatorial libraries can provide a cost-effective means of creating more choices for companies and consumers in meeting changing demand.

Key Applications

There are a multitude of applications where combinatorial methods can make an impact on the design of new materials (see Exhibit 3-2). These include the development of new types of materials, as well as

Exhibit 3-2. Markets and Key Applications

Materials	• Formulation/process interaction • Non-lead solder • Superconductors • Modifying cement properties • Ability to better match material to application • Materials interaction • Composites/smart materials • Magnetic materials	• Anisotropic materials • Sensors • Materials by design • Ceramics • Improved/alternative materials • Material substitution • Finding the DNA of materials • Nano-scale materials • Polymers • Hybrid products—bi-functional
Electronics	• Materials for nano technologies in electronics • Home appliances • High speed • Flexible tooling	• Miniaturization • Dilectric constant • Fiber optics • Breakthrough technologies • Multifunctional materials
Pharmaceutical/ Medical/Biotech	• Genetic-based drugs and diagnostics • Lab-on-a-chip • Custom-designed/patient-specific materials • Proteomics • Overcoming liability issues • Virtual clinical studies/drug design	• Hi-thru put blood, air, breath samples • Medical and biomaterials • Biodegradable • Materials to serve aging population • Biomimetics • Biometric materials
Photonics	• Multifunctionality • Nonlinear optics • Optical computers	• High band width • High speed • Optical networking
Coatings, Adhesives & Lubricants	• Automotive industry - plastics, solvent-based aluminum • Multifunctional coatings to impart variety of qualities	• Shift from solvent-based coatings (non-aqueous to aqueous) • New adhesives
Catalysis	• Faster, cheaper • Selectivity/Yield	• Heterogeneous/homogeneous catalysts • Alternative fuels
Chemical Process/Product Design	• Directed molecular evolution • Sustainable, more profitable routes to commodity chemicals • More efficient, more selective, higher yield processes with fewer byproducts • Third-world infrastructure for commodity chemicals • Final applications (e.g., 3D travel, colonization of space) • Developing different pathways for same problem (alternatives) • Refining models to predict kinetics	• Dream reactions (e.g., methane to methanol) • Extending spin-offs to process control applications • Product diversification/ customization • Combinatorial factory • Troubleshooting/failure analysis • Process optimization • Absorption/separation processes • Raw materials selection • Gas to liquids • Combination production on large scale
Environmental Performance	• Environmentally-safe materials • Design for recycle or absorption - industrial ecology	• Waste minimization • Remediation and environmental applications (sensing, mining)
Energy Technology	• Energy conversion and storage	• Fuel cell catalysts/components

innovations in the materials design process. Opportunities exist to use combinatorial methods to develop superconductors, smart materials, magnetic materials, anisotropic materials, sensors, ceramics, polymers, and a host of other products. Combi methods can also facilitate the formulation process for polymers, assist in evaluating materials interaction, and enable discovery of material substitutions or alternatives (e.g., additives).

There are many newly emerging opportunities for the design of materials for pharmaceutical and medical applications. Combi chemistry could be used to design unique materials for individual patients, or materials especially suited for an aging population. These methods are already being used in the synthetic drug discovery process, and could also be used to create genetic-based drugs and diagnostics, or for virtual clinical studies for drug design. In the medical arena, high throughput methods could be extended to the analysis of blood, air and breath samples.

Another area with large potential impact is electronic materials, where technological advances are being made at a rapid pace and where the time expended in developing new materials is of crucial importance. A closely related and equally important field is photonic materials, which are needed to fuel the revolution in optical computing and networking. Requirements for new photonic materials include high band width and speed, and multifunctionality.

There is a continual need for improved coatings, adhesives, and lubricants, particularly in the automotive industry. Combinatorial methods are already beginning to make inroads in this broad opportunity area. Demand for these materials is often driven by consumer expectations for higher performance, better quality, and environmental sensitivities. Combinatorial methods could provide an opportunity to better meet consumer demand and expectations. For example, combinatorial methods could be used to facilitate the discovery of alternative coatings that require little or no chemical solvents, which would reduce the amount of volatile compounds released to the environment.

The design of new catalysts represents a particularly promising opportunity area. Combinatorial libraries can substantially decrease the amount of time and money spent searching for catalysts, and facilitate the design of catalysts with tailored properties.

Chemical process design is an area where combinatorial methods could potentially revolutionize materials manufacturing. It could be used, for example, to create more sustainable and profitable processes for manufacturing commodity chemicals, which are typically low margin, high volume products. By generating many different choices, combi methods could identify many options for creating the same product, allowing the selection of the most optimal. With combi methods, it could be possible to identify routes for making “dream” reactions possible (e.g., methane to methanol). There also opportunities to use combi chemistry to optimize current processes, making them more efficient and selective, with higher yields and fewer byproducts.

Combinatorial methods can potentially make an impact on environmental performance and the use of energy resources. There are opportunities to design materials that are environmentally safe, or materials that are easily recycled or absorbed, eliminating potential release to the environmental. Combi could be used to identify ways to minimize waste and reduce byproducts, or for remediation and other environmental end-of-pipe options. Combi could also provide a means for exploring new, cleaner ways of converting and storing energy (e.g., fuel cells).

There are a number of applications that are unique to individual aspects of combinatorial methods, specifically library characterization and informatics (see Exhibit 3-3). Characterization methods could

Exhibit 3-3. Unique Applications for Specific Combi Methods

Measurement/ Characterization	• Obtaining physical properties • Combi and MEMS systems • Obtaining kinetic numbers for plugging into models • Corrosion, degradation, stability, and effects on human health • Characterization tools, high-throughput screening for nano technology	• Performance prediction • Directed molecular evolution (Cal Tech work) • Non-robotic high-throughput measurements • Non-optical based methods (10^9 mechanical property tests/year) • Processing assays (rheological, etc.)
Informatics	• Value maximization • Globalization • Business development • Linking technical, business, other decision factors • Industrial power management • Safety • Good citizenship/public image—incorporate all variables/ constraints • Increasing R&D productivity	• Reducing massive amounts of information to manageable levels • Places where knowledge must be preserved • Academic, government, industry integration • New knowledge generation • Novel algorithm development • Distance research (collaborative) • Training and academic use

yield a wide diversity of useful data, such as physical properties, reaction kinetics, materials performance, mechanical properties, corrosion and degradation of materials, and toxicity. This information can provide a basis for new models, for materials design, screening for safety and reliability, and many other applications.

Informatics, which is essentially an organized system of collecting, interpreting, and storing the data obtained from combinatorial library screening, is essential for combinatorial techniques but can also provide important information for management and operation. If developed effectively, informatics could be used to maximize value from products, to build global infrastructures, and for business development. Informatics could provide a tool for linking technical, business and other factors needed for the corporate decision-making process.

The advantage of informatics is that it can reduce massive amounts of information to the point where it is more easily managed and used. It also provides the means for storing or archiving data, preserving it for future use. There are opportunities to use informatics to integrate the knowledge generated from different scientific communities (industry, academia, government) and create a basis for collaborative research and information sharing.

4 Design of Experiment and Models

Vision for 2020

By 2020, experiments and models will be integrally linked with the combinatorial process, and will be designed to effectively obtain the desired results.

> *By 2020, the knowledge base for designing experiments and models will be exponentially larger, as will computing capability. Databases will be highly sophisticated, and economic and business models will be routinely used for decision-making.*
>
> *In the future, there will be close integration between experimentation and modeling capabilities, with the ability in place to re-evaluate accepted models and premises. Models will be able to distinguish and consider practical operating conditions versus simulated conditions. Experiments will be better designed at the front end to obtain the desired output.*
>
> *By 2020, many experiments will be done virtually. Robots will run experiments, and they will conform to established standards. Negative as well as positive results will be stored to facilitate the learning process.*
>
> *Enabling technologies, such as computational chemistry, fluid dynamics, and process models, will be in place to support better design of experiments. Through more advanced experiments, the design of "super" materials will be possible.*
>
> *By 2020 experimental design will be based on known reaction mechanisms and pathways—especially pathways that result in specific materials microstructures.*

Performance Goals

To achieve the vision for 2020, a number of goals have been identified for better design of experiments and models. An overarching goal is to advance computational and combinatorial methods to the point where they ultimately reduce the volume of experiments required for designing new compounds. Technical goals in support of this are the ability to use experiments to effectively evaluate the impacts of different micro structures on materials performance, and the ability to solve a chemical field.

Design of Experiment and Models Goals for 2020

- Ability to consider/evaluate micro structures
- Ability to "solve" a chemical field
- Fewer experiments through the use of computational and combinational methods

Barriers to Entry

The **technical infrastructure** for integration of experiments and computational capability is currently limited in several critical ways (see Exhibit 4-1). The most critical barrier is the lack of integration between the equipment used to perform experiments and the theory available for interpreting the results. This lack of integration creates a disconnect in experiment design as well as utility. Another critical problem is that current models cannot span multiple scales (i.e., translate results from atomistic scale to

Exhibit 4-1. Critical Barriers to Achieving Goals for Design of Experiment and Models

Technical Infrastructure	**Lack of integration of equipment and theory capabilities** **No models that span multiple scales; lack of interface between micro-meso-macro scales** **Lack of knowledge representation standard** Lack of data and standardized data Lack of common language (e.g., a "web-site" language to describe and communicate materials across multiple scales) Lack of standards for instrumentation Lack of connectivity Software/hardware tools for pharmacology not applicable No models matched to the experiments being done No scale-up models Lack of sufficiently flexible databases and queries Difficulty in translating software tools to new applications
Other Technical Issues	**Paradigm shift needed in designing experiments** – **Ability to process multiple experiments and results** – **Make the first experiment right** – **Include failures** Sample preparation and product analysis Lack of an assay and associated equipment Weak understanding of the correlation between process micro-structure and performance
Cultural	**Cultural perceptions of combinatorial methods** Intellectual property issues and lack of communication Lack of knowledge sharing Lack of rewards for working in teams
Economics	**Cost of entry** Insufficient funding (public, private) No off-the-shelf technology Heterogeneity of markets and applications and products Existing structures (e.g., capital equipment, geography, sales)

Note: **Bold = top priority;** Regular type = high priority

mesoscale to bulk, etc.), and there is no adequate interface between models that are simulating properties at different scales. Models are currently not well-matched to the experiments being performed, and models are not available to scale-up results to practical levels.

Standards and common interfaces between equipment is another area of importance where current systems are inadequate. There is no standard for knowledge representation, little standardized data (or any data), and no standards for instrumentation. No common language exists to interpret data, and there is little connectivity between databases.

Other technical issues exist in the current mind set for designing experiments. A paradigm shift will be needed to change the way experiments are now designed, to enable the ability to process multiple experiments and results, to design experiments correctly from the beginning, and to include and learn from failures. Other technical limitations exist in capabilities for sample preparation and product analysis. While technology exists in the pharmaceutical field, many of these cannot be easily or successfully applied to materials design.

Cultural issues include misconceptions about the use of combinatorial methods in general. Not enough sharing of knowledge takes place, and there are real concerns about protection of intellectual property rights. Part of the problem is that not enough information about combinatorial methods is being communicated, and it is poorly understood except within a small community. The benefits of working through consortia or other collaborative efforts is also poorly communicated, and not many such efforts are taking place. However, there is an increasing trend toward collaboration in the pharmaceutical arena, and some of these are extending to other areas of research such as catalyst and coating design.

A critical **economic** barrier is the high cost of entry into combinatorial techniques. Existing equipment is expensive, and scientists trained in combinatorial methods are few. Little off-the-shelf technology exists that can be adapted to materials design, and market potential must be significant for vendors to design entirely new equipment. Markets and applications are very diverse, and equipment must be designed to meet specific requirements in each. Existing structures to support interest in new markets are simply not well-established (e.g., capital equipment, demonstrated sales records). An underlying problem is the lack of funding from both the public and private sectors for combinatorial experimentation and theory. Without this support, some of the costly pre-competitive research needed to make the necessary advances in equipment and theory is not being conducted by companies or universities.

Research Needs

Experimental Methods—There is an immediate need to design experimental methods that produce accurate and most important, repeatable results (see Exhibit 4-2). To support this, methods are needed for effective tracking of samples, as well as protocols and standards for interpreting and reporting experimental data and results. One of the top priorities, and one that is achievable over the next 10 years, is the ability to design experiments based on reaction pathways. This would facilitate the ability to design materials based on existing or newly formulated synthetic techniques and processes. Another mid-term priority is developing the capability to conduct high throughput analysis of assay data.

Modeling Tools—Over the near-term, a top priority is the development of algorithms to enable rapid searching of experimental data. To support rapid search capability, technology is needed to perform rapid database queries, and to integrate database content. In the mid-term, a top priority is to conduct the experiments necessary to enable theory development for new, critical models. High priorities for model development include those that consider time-dependent parameters. Examples are stability over the material life cycle, accelerated testing applied to chemistry, and standards testing (ASTM-type). Ultimately, the long-term top priority is to utilize mid-term results to develop and perfect models that span multiple scales, which are now currently very limited. These would allow the prediction of materials properties at one level, and scale-up of those properties to the next level (e.g., from atomistic to mesoscale).

Exhibit 4-2. Research Needs for Design of Experiment and Models

EXPERIMENTAL METHODS

Near-Term (0-3 yrs)	*Mid-Term (>3 -10 yrs)*	*Long-Term (>10 yrs)*
H Experimental methodologies that are repeatable M Tracking mechanism for samples M Set up protocols and standards for experimental data	T Experiment design based on reaction pathways H High throughput analysis of assay data	

MODELING TOOLS

Near-Term (0-3 yrs)	*Mid-Term (>3 -10 yrs)*	*Long-Term (>10 yrs)*
T Algorithms to enable rapid search capabilities M Technology to query databases M Integration of content M Accessible tools for modeling computational fluid dynamics (CFD)	T Experiments to enable theory development for new models H Models that consider time dependence, including 1) accelerated studies, 2) life-cycle stability, 3) accelerated testing applied to chemistry, and 4) ASTM-type testing. H Integrated system of modeling tools	T Models that span multiple scales, allowing prediction at one level and scale-up to another M "Fuzzier" search capabilities (e.g., learning engines)

HARDWARE/INSTRUMENTATION

Near-Term (0-3 yrs)	*Mid-Term (>3 -10 yrs)*	*Long-Term (>10 yrs)*
	T New hardware, including 1) micro-reactors, 2) faster analytical methods, 3) micro-channels, and 4) expert system product to guide new users, order of process used. T Central nervous system with flexible architecture	

ADVANCED COMPUTING CAPABILITY

Near-Term (0-3 yrs)	*Mid-Term (>3 -10 yrs)*	*Long-Term (>10 yrs)*
H Problem taxonomy H Capability to handle larger models M Pattern recognition M Models that are scalable	M New, efficient descriptor generator M Advanced capabilities for properties that experiments measure	H New/improved computational algorithms

T = Top priority; H = High priority; M = Medium priority

Hardware/Instrumentation—The development of new hardware is a top priority. Specifically, needs include micro-reactors, faster analytical methods, instruments with micro-channels, and expert systems for new users. An overarching top priority is the development of a central nervous system with a flexible architecture to guide experimentation and interpretation of results.

Advanced Computing Capability—There is an immediate need for advanced computing capability that can handle problem taxonomy and classification. Other near-term priorities include computing systems that can handle much larger models than what is currently available, and have pattern recognition capability. This includes models that are scalable from micro-macro levels. In the mid-term, research is needed to create new, efficient descriptor generators, and advanced capabilities for a very wide range of materials properties. Over the next 10 years and beyond, better algorithms will be needed to handle theory developments coming out of combinatorial science.

Priority Non-research Activities/Strategies—To really make advances in the way experiments and models are designed, it will be critically important to pursue multi-disciplinary research activities (see Exhibit 4-3). This can best be accomplished through collaborative efforts, such as consortia where data and results are collected and shared to mutual benefit. To reach goals for combinatorial methods, it will also be important to explore ways to take advantage of breakthroughs in computing capability and other areas of science (e.g., parallel computing). Financial modeling tools will be an essential element in the road toward the future—investors must be able to see and measure the value of combinatorial methods. Finally, to be truly effective, combinatorial chemistry must increasingly become an on-line process through the greater use of computers and visualization software and hardware.

Exhibit 4-3. Priority Non-Research Activities/Strategies for Design of Experiment & Models

- **Identify/develop multi-disciplinary research needs through collaborative efforts, consortiums to collect/share data and results**
- **Explore analogous applications for paradigm shifts (e.g., parallel computing)**
- Develop financial modeling tools
- Develop more of an on-line process through increased use of computers, visualization software and hardware

Note: **Bold face** = top priority; Regular type = high priority.

5 Library Production

Vision for 2020

By 2020, it is envisioned that technical advances will establish library production as an integral part of the design process for many kinds of materials, from polymers to coatings to catalysts.

> *In the future, libraries will be easily integrated to the characterization method. For example, libraries will be produced onto a sensor, or onto an array. This will provide the capability to characterize libraries as they are created.*
>
> *The process for producing libraries will be robust, enabling easily reproducible libraries. Library production will be automated to make libraries from both liquid and solid raw materials, and techniques such as split and pool will be readily applied to materials. Off-the-shelf equipment will be available to process the library, with a standard footprint/universal standards in place for processing and calibration.*
>
> *A variety of methods will be available to create very diverse libraries, including libraries that explore both compositional and process variables, and libraries for very sensitive materials (e.g., produce the container). Synthesis methods will advance to the point where they are not limited by combinatorial precursors, or rate-limited by accessibility to raw materials. Non-destructive testing techniques will enable the production of micro-scale libraries.*
>
> *It will be possible to readily archive libraries, using a standard footprint (microtitre plate equivalent). Informatics, design, and production will be linked and web-accessible, enabling the seamless movement of information. Virtual libraries will be in place with the capability to accurately predict the performance of materials.*

Performance Goals

A number of goals have been identified for library production for the year 2020. All the goals support the overarching concept that the use of combinatorial libraries could potentially revolutionize the way materials are designed, as well as the time required for the discovery process. A specific goal was set to reduce the time of development, from test tube to pilot scale, by 10-fold, with the understanding that more time may be required to develop entirely new materials. To achieve this in a cost-effective and efficient way, goals were set to increase the capability for high throughput screening by 100-fold, and to ultimately reduce the costs required to run samples by 100-fold.

In support of the state-of-art envisioned for library production in 2020, goals were set to create libraries that are web-accessible and networked, and accurate and scalable in terms of providing leads for promising new compounds. An ultimate goal is that the discovery of materials be integrated with process development.

Library Production Goals for 2020

- Reduce development time by 10-fold
- Improve throughput by 100-fold
- Reduce the cost to run samples by 100-fold
- Web-accessible/networked libraries
- Accurate, scalable libraries
- Balance process development with discovery

Critical Barriers to Entry

The capability of currently available **equipment** for producing combinatorial materials libraries is very limited (see Exhibit 5-1). While many of the instruments and systems developed for pharmaceuticals can potentially be adapted for materials research, only a few have been. Most that have been adapted are being used by individual companies for specialized purposes, and are not available commercially, although a number of equipment manufacturers are beginning to enter this market. Equipment for processing and handling samples, particularly gases, is very limited. The extreme diversity of reaction conditions makes it difficult to develop equipment that can be universally applied to a wide range of compounds. Currently available equipment also lacks the capability to purify and characterize materials at the rate at which they are being synthesized.

Exhibit 5-1. Critical Barriers to Achieving Goals for Library Production

Equipment Limitations	**Lack of equipment for processing/ handling library samples, especially gases** Extreme diversity of reaction requirements Inability to purify and characterize materials at the same rate as they can be synthesized
Raw Materials Synthesis	**Lack of easy and affordable access to raw materials needed for libraries** Lack of understanding of how pure raw materials must be
Sample Handling	**Solids handling/ deposition is difficult** High viscosity liquid deposition is difficult Lack of confidence that sample is properly mixed
Scalability (Performance & Process)	**Lack of control of form/morphology of material sample** Sample size needed for testing is too big Lack of confidence in performance scalability
Strategic Issues (Non-Technical)	**Intellectual property issues** **Conservative business management** Lack of multi-skilled groups; inadequate integration of multiple skills and dynamics
Education/ Awareness	**Lack of academic awareness and understanding of industrial impact** **Understanding of industrial impact** Lack of properly trained students

Note: **Bold = top priority;** Regular type = high priority

To make materials library production a realistic design tool, an easily accessible, affordable source of **raw materials** must be available for synthesis. Currently available sources are often not affordable or available in the range of compounds needed, and information on where to obtain raw materials is not well-publicized. Another issue is lack of demonstrated scientific knowledge behind library production for materials rather than pharmaceutical purposes—i.e., how pure must the starting materials be to obtain the desired answers?

Issues also arise in the physical **handling of samples** of some compounds. The deposition of solids and high viscosity liquids is a particularly difficult operation. Another limitation is that it is often hard to tell if samples are properly mixed, which reduces confidence in the validity of results.

An important part of finding useful, viable compounds through combinatorial libraries is **scalability—**i.e., the ability to scale results at the micro-scale up to a larger scale. A scalable compound is one that retains the desirable properties exhibited at the lab scale when it is produced on a larger scale (e.g., pilot scale). One of the barriers currently limiting scalable of new materials is the inability to control the form or morphology of the sample. Another key issue is that the sample sizes needed for testing to ensure scalability are too large. Given the current methods of library production, there is a general lack of confidence in the scalability of performance of newly discovered materials.

In addition to technical limitations, there are a number of **strategic issues** that will influence the development and adoption of library production for materials design. Intellectual property rights is an issue in both development of new equipment and methods for library production, and in the sharing of knowledge gained by individual companies. Conservative business philosophies can be a critically limiting factor. In a conservative business climate, funding is limited for enabling tools like combinatorial methods unless profit motivation is clear. Another limiting factor is that groups currently working on combinatorial techniques are often isolated, and do not include multi-disciplinary teams. In general, in the combinatorial field there is inadequate integration of multiple skills and the dynamics needed to accelerate advances in technology.

An important barrier is the general lack of awareness in both academia and industry of the potential impacts of combinatorial methods on materials development. This is due in large part to the limited understanding at the present time of the real impacts, especially in terms of the cost of development of new products. In academia, few students are being trained in combinatorial methods, although a few universities have established combinatorial centers.

Research Needs

A number of important research areas have been identified that are necessary to make advances in library production. Priority topics include equipment for library production and handling/storage of samples, raw materials, and scalability of systems. Exhibit 5-2 illustrates specific priority research within these topics, and the time frame needed to provide useful results.

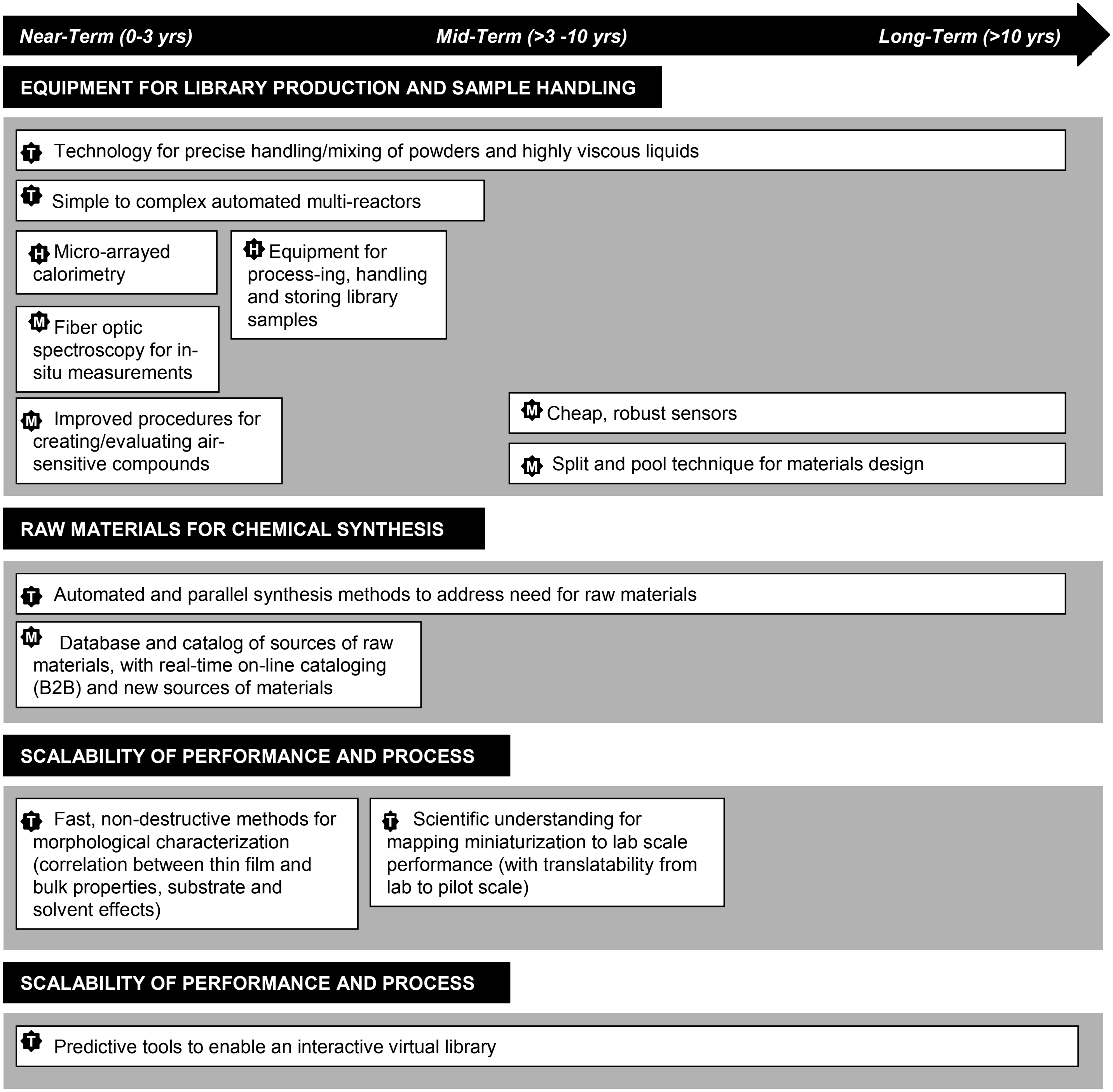
Equipment for Library Production and Sample Handling—Reliable automated systems are essential for handling the millions of reactions needed for the production of large libraries. A top priority near-term research need is the development of a series of automated reaction devices with a range of capabilities, from very simple to highly complex, multiple reactions. Other important near-term priority research that will help to make the needed advances in equipment includes micro-arrayed calorimetry, and *in situ* measurement techniques such as fiber optic spectroscopy.

In addition to library production, equipment is needed to process, handle and store library samples. One of the highest priorities is the development of technology for the precise handling and mixing of powders and very viscous liquids at small scales. System requirements are the capability to handle wide temperature, pH, and pressure ranges. While there is an immediate, near-term need for handling and storage technology, developments will continue to be necessary over the next ten years and into the longer-term to keep pace with advances in library production.

Over the longer-term, a priority is the development of robust, cost-effective sensors for use in conjunction with library synthetic techniques. Another important mid-long term research need is the extension of existing split and pool techniques (developed for pharmaceuticals) to the production of libraries for materials design.

Raw Materials /Chemical Synthesis—A top priority is the development of both automated and parallel synthesis methods that can supply a readily-available source of raw materials for library production, with commercial technology in place by the mid-term (within 10 years). To support library production in the near-term, a database or catalog of sources of available raw materials is needed. Capabilities might include real-time on-line cataloging e.g., on-line combi store) and updates on new sources of materials (e.g., from universities).

Exhibit 5-2. Research Needs for Library Production



Near-Term (0-3 yrs) ***Mid-Term (>3 -10 yrs)*** ***Long-Term (>10 yrs)***

EQUIPMENT FOR LIBRARY PRODUCTION AND SAMPLE HANDLING

- T Technology for precise handling/mixing of powders and highly viscous liquids
- T Simple to complex automated multi-reactors
- H Micro-arrayed calorimetry
- H Equipment for process-ing, handling and storing library samples
- M Fiber optic spectroscopy for in-situ measurements
- M Improved procedures for creating/evaluating air-sensitive compounds
- M Cheap, robust sensors
- M Split and pool technique for materials design

RAW MATERIALS FOR CHEMICAL SYNTHESIS

- T Automated and parallel synthesis methods to address need for raw materials
- M Database and catalog of sources of raw materials, with real-time on-line cataloging (B2B) and new sources of materials

SCALABILITY OF PERFORMANCE AND PROCESS

- T Fast, non-destructive methods for morphological characterization (correlation between thin film and bulk properties, substrate and solvent effects)
- T Scientific understanding for mapping miniaturization to lab scale performance (with translatability from lab to pilot scale)

SCALABILITY OF PERFORMANCE AND PROCESS

- T Predictive tools to enable an interactive virtual library

T = Top priority; H = High priority; M = Medium priority

Scalability—Development of fast, non-destructive methods for characterization of morphology is a top priority in scalability of library results, especially in the near term. The most important criteria in new characterization methodology will be the capability to draw correlations between thin films and bulk properties. Methods should also be able to characterize substrate and solvent effects on samples.

Another key research area for scalability is the capability to map miniaturization to performance of materials at the lab scale. This will require new scientific understanding of the process, and the inherent limits of scalability. Most important and of high priority over the longer term is the ability to further transfer and extend this knowledge from the lab to the pilot scale.

Virtual Library—New and improved predictive tools are needed to enable the development of virtual libraries with interactive capabilities. Development of these tools will require research to develop a better fundamental understanding of the performance of materials.

Non-research Activities/Strategies—Exhibit 5-3 illustrates some of the critical non-research activities that could help to accelerate development an4d use of combinatorial libraries for materials design. Increasing awareness in industry, government and academia of the benefits and applications of combinatorial chemistry is a key issue. A strategically important action is obtaining Federal support for pre-competitive research programs that focus on advances in combinatorial chemistry.

Exhibit 5-3. Priority Non-Research Activities

- **Increase exposure of combi chem in materials/ chemical trade associations (MRS, ACS, AIChE, SMAPE, APS, ACerS)**
- **Lobby government for combi chem funding and support (both state and Federal)**
- **Introduce formal programs in universities (NSF funding)**
- Secure funds for joint industry/ academia programs
- Improve patent office understanding of combi chem claims
- Encourage application-oriented patents
- Use consortia approach to facilitate sharing of information
- Sponsor inexpensive workshops on combi chem to increase participation of academics

Note: **Bold face** = top priority; Regular type = high priority.

6 Library Characterization

Vision for 2020

By 2020, it is envisioned that the ability to rapidly and accurately characterize libraries for materials design will be a reality.

> *In the future, characterization methods will rely on the use of rapid, inexpensive sensors to measure composition and physical properties. Through advanced measurement techniques, it will be possible to realistically map ternary material phase diagrams (composition versus structure versus properties). Systems will also have the capability to relate micro-structural to non-micro properties and conditions.*
>
> *By 2020, combinatorial characterization will be completely automated, with a hands-off system for screening and organizing data. Characterization will be fully integrated within the combinatorial process, from discovery through commercialization.*

Performance Goals

Characterization Goals for 2020

- Prepare, screen and characterize a trillion experiments a year for a 1/10th of a penny each (industry-wide) and archive the data
- Ability to characterize 1,000s of compounds a day—structure, surface, composition (e.g., rheology, performance)
- Reliable, accurate, precise high throughput capability
- Be able to correlate performance and properties
- Be able to scale-up micro measurements to performance level
- World-wide materials data bank accessible by all
- Combi lab in every company/university (effective, fast, routine, standard part of R&D), with shareable data
- Output measure(s)
  - Reduce time-to-market to 1-2 years (break-through discovery); 3-6 mos (incremental)
  - Increase number of products to market

Rapid, accurate, flexible characterization methods are essential to the future acceptance and more widespread use of combinatorial methods for materials design. Specific goals were set to achieve the capability to prepare, screen and characterize a trillion experiments a year, at a cost of 1/10th of a penny. The goals are to be able to characterize thousands of compounds a day in terms of structure and composition, be able to relate properties to performance, and be able to scale up measurements.

Ultimately, the data generated through combinatorial methods would be used to create a world-wide materials data bank that is accessible to all scientific communities and companies involved in combinatorial research and development. A supporting goal is to have a combinatorial lab in every company and university, and to advance this tool to the point where it is considered an effective, routine, standard part of research.

With these goals in mind, it should be possible to reduce the time-to-market for new products to 1-2 years for scientific breakthroughs, and as little as 3-6 months for incrementally improved products. The result will be a significant increase in the amount of new products entering the market.

Barriers to Entry

Currently existing **measurement tools and techniques** are significantly limited or non-existent for the characterization of libraries produced for materials design. One of the most critical barriers is the current lack of technology for making high-throughput measurements at the needed time and scale (see Exhibit 6-1). Another very limiting factor is that the characterization of the properties of small samples does not necessarily translate well to the larger scale, particularly in terms of performance. In general, current methods for measuring performance characteristics, which largely determine whether a compound is desirable (a "hit"), are very inadequate for materials design. The tools and instruments available today cannot provide the rapid, high quantity measurements needed to make combi a really useful tool. The capability for parallel measurements, for example, is virtually non-existent for these types of compounds. Current equipment can be used to characterize from ten to hundreds of compounds, not the thousands needed.

Current tools are also limited in terms of the range of material surfaces that can be characterized. For example, existing instruments cannot characterize materials that are irregular or non-ideal. A rough surface on a sample may diminish the ability to distinguish composition or other properties. Detector technology in use today cannot be easily applied to materials applications, and is currently too expensive or difficult to adapt. In general, most of the existing techniques have been developed for pharmaceutical applications, and have not been tailored to materials design. This situation will only improve if more vendors become interested in the market potential of new applications.

Advances in **infrastructural technology** (models, information technology) may not be sufficient to support new combinatorial methods. For example, models may not keep pace with new developments in combinatorial methods, and could ultimately place limits on how far combi can advance. An area where serious limitations currently exist is technology for data treatment, storage and handling (often referred to as "informatics"). Informatics and related issues are dealt with in more detail in Chapter 7.

The cost of entry for a combinatorial system is currently very high, and investment in these techniques are not warranted unless significant **economic** benefits can be identified. Market segments must be large enough to attract the interest of instrument manufacturers, as each will have different measurement requirements calling for individualized equipment. Another factor is the limited amount of funding (private or public) for combinatorial techniques, although interest is growing, primarily due to the increasing use of combi in drug discovery.

An underlying problem is the **cultural** perception of combinatorial methods as a scientific novelty rather than as a means to more successful product development. The value of combi techniques has not been clearly demonstrated to decision-makers, and continues to be viewed with skepticism because of the lack of proven successes in the market place. There is little understanding in the corporate boardroom for the true vision of combinatorial methods and their potential to revolutionize the discovery process. Contributing to the problem is the lack of champions lobbying for government funding for combi research. Because of its growing use in pharmaceuticals, some may view it a strictly commercial enterprise—but a tremendous amount of research is still needed to successfully extend these techniques to other fields. Without government funding, not enough academic (pre-competitive) research of the type needed to make breakthrough advances is being pursued. Those involved in the traditional sciences of chemical synthesis and materials research are inherently opposed to combinatorial methods, and only a few groups are pursuing new characterization techniques.

Exhibit 6-1. Critical Barriers to Achieving Goals for Library Characterization

Current Measurement Tools/Techniques	**Technology to make measurements at time and scale needed do not exist for high throughput** **Small samples don't necessarily tell you how the "whole" will perform** **Inadequate methods for measuring performance characteristics (balanced hits versus non-hits)** Current tools are not robust enough— equipment limited to finite number of samples Instruments and techniques for parallel measurements are virtually non-existent Inability to characterize materials that are irregular, i.e., non-ideal samples Current detector technology is not cheap or broad enough for multiple application High throughput characteristic techniques exist, but are not tailored for materials needs
Infrastructural Technology	**Modeling development may not keep pace with what is needed to support combi** Non-existent data treatment, storage, and handling technology (informatics)
Market/ Economic Issues	**Cost of entry** Lack of funding Inadequate development by analytical instrument manufacturers All market segments will have different measurement goals
Cultural Issues	**Lack of champions to lobby for government support of combi** **Lack of acceptance in industry and academia; lack of education** **– Few demonstrated successes** **– Value not sufficiently demonstrated** Lack of vision for the potential of combi Inherent opposition to combi techniques in traditional chemical/materials sciences Too few groups are pursuing characterization techniques

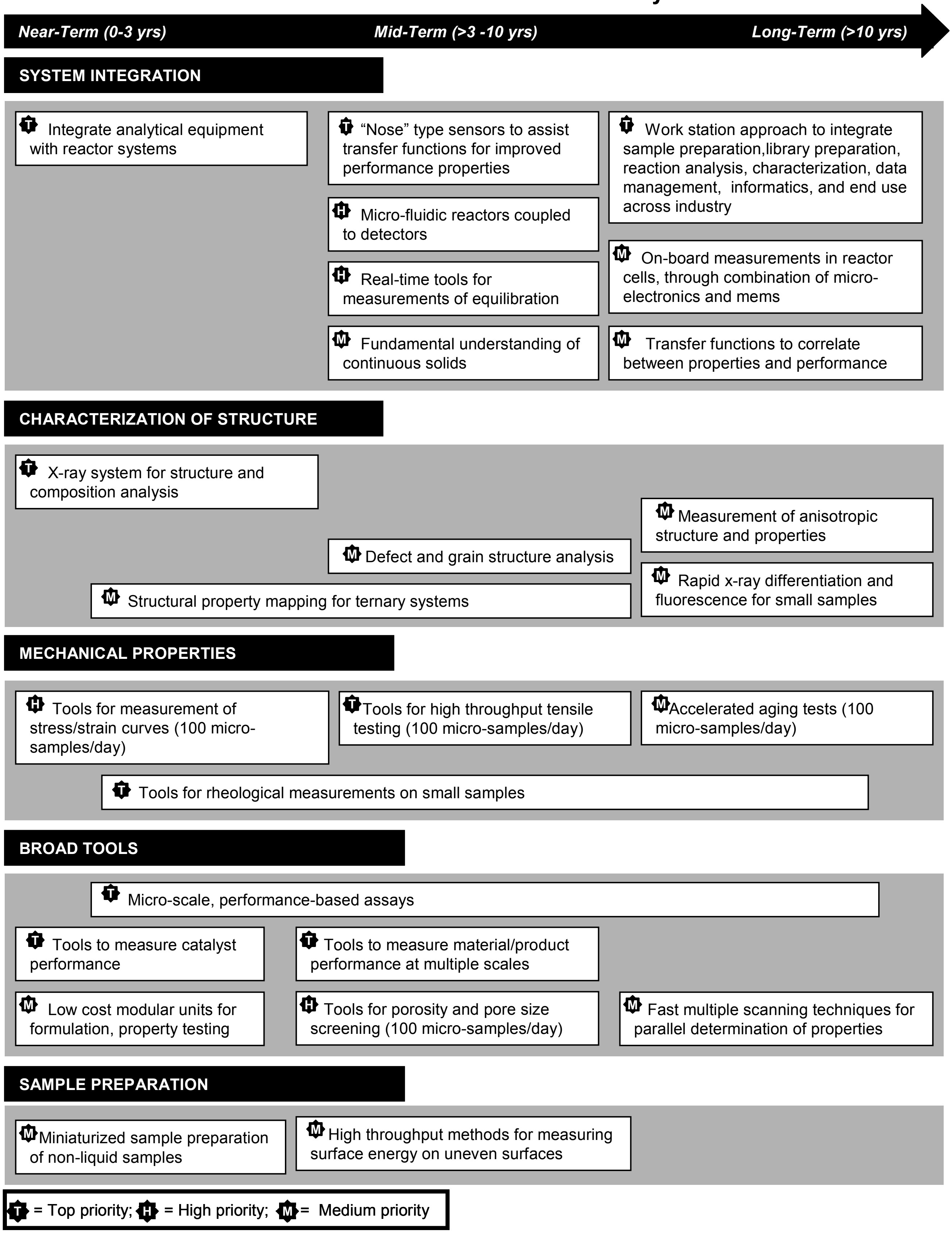
Note: **Bold = top priority;** Regular type = high priority

Research Needs

Research needed for more accurate, useful characterization of libraries is illustrated in Exhibit 6-2.

System Integration—One of the top near-term priorities for library characterization is the ability to integrate analytical equipment with the reactors performing chemical synthesis. Integrated systems would enable more rapid, accurate characterization of larger numbers of samples. Over the next ten years, research is needed to develop a variety of sensors and measurement tools to enable better characterization and support rapid analysis. These include "nose" type sensors to assist transfer functions for improved performance properties, relying on continuous rather than batch reactions. Other important concepts include real-time tools to measure equilibrium conditions, microfluidic reactors that can be coupled to detectors, and on-board measurements within individual reactor cells. Ultimately, a

Exhibit 6-2. Research Needs for Library Characterization



Near-Term (0-3 yrs) — *Mid-Term (>3 -10 yrs)* — *Long-Term (>10 yrs)*

SYSTEM INTEGRATION

- T Integrate analytical equipment with reactor systems
- T "Nose" type sensors to assist transfer functions for improved performance properties
- H Micro-fluidic reactors coupled to detectors
- H Real-time tools for measurements of equilibration
- M Fundamental understanding of continuous solids
- T Work station approach to integrate sample preparation,library preparation, reaction analysis, characterization, data management, informatics, and end use across industry
- M On-board measurements in reactor cells, through combination of micro-electronics and mems
- M Transfer functions to correlate between properties and performance

CHARACTERIZATION OF STRUCTURE

- T X-ray system for structure and composition analysis
- M Measurement of anisotropic structure and properties
- M Defect and grain structure analysis
- M Rapid x-ray differentiation and fluorescence for small samples
- M Structural property mapping for ternary systems

MECHANICAL PROPERTIES

- H Tools for measurement of stress/strain curves (100 micro-samples/day)
- T Tools for high throughput tensile testing (100 micro-samples/day)
- M Accelerated aging tests (100 micro-samples/day)
- T Tools for rheological measurements on small samples

BROAD TOOLS

- T Micro-scale, performance-based assays
- T Tools to measure catalyst performance
- T Tools to measure material/product performance at multiple scales
- M Low cost modular units for formulation, property testing
- H Tools for porosity and pore size screening (100 micro-samples/day)
- M Fast multiple scanning techniques for parallel determination of properties

SAMPLE PREPARATION

- M Miniaturized sample preparation of non-liquid samples
- M High throughput methods for measuring surface energy on uneven surfaces

T = Top priority; H = High priority; M = Medium priority

top priority goal is to achieve integration of the system from sample preparation to end-use. This will require longer term research to develop a work station approach that will integrate reactor sample preparation, reaction analysis and characterization, management of the data produced (storage, interpretation, informatics), and end-use across industry.

Characterization of Structure—Sensitive, accurate tools for structural analysis are needed to enable effective high-throughput screening of libraries and optimize promising leads for new compounds. A top priority in the near-term is the development of an x-ray system for analyzing structure and chemical composition. Over the next ten years, research is needed to develop tools for better characterizing defect and grain structure, anisotropic structure and properties, and mapping of ternary systems. Fast, accurate analytical techniques such as rapid x-ray diffraction and fluorescence that can be readily applied to small samples are also needed.

Mechanical Properties—To keep pace with library development and use for materials design, tools that can measure 100 micro samples per day will be needed to measure a variety of mechanical properties. In the immediate future, the most important tool will be the measurement of material stress/strain curves. In the mid-term, as libraries become more sophisticated and routinely used, tools will be needed for high throughput testing of tensile strength and accelerated aging of materials. An on-going priority research need is the capability for making rheological measurements on small samples, particularly for high throughput screening.

Broad Tools—Better measurement tools are needed to characterize a broad range of chemical and physical properties. In general, new tools must be cost-effective, rapid and able to handle 100 samples per day. The most important properties are catalyst performance (yield, selectivity, kinetics), material performance at multiple scales (particularly the micro-scale), and porosity and pore size. Rapid multiple scanning techniques are also needed to enable parallel determination of physical properties. A near-term need is the development of low-cost modular units for formulation, mixing, and multiple property testing (gloss, hardness).

Sample Preparation—To broaden the utility of combi libraries for materials design, a high priority will be the development of high throughput methods for measuring the surface energy of uneven surfaces, such as those found in powders. A near term priority is the development of miniaturized techniques for the preparation and mixing of non-liquid samples.

7 Informatics

Vision for 2020

By 2020, informatics will have advanced to a stage where it enables the global distribution and use of information. By 2020, the world will exchange information and knowledge seamlessly and transparently using the same language for the generation of new knowledge and products.

> *In 2020, informatics will maximize the value of information. As a tool it will turn information into knowledge using visualization and modeling as analysis techniques, and will employ artificial intelligence to help solve problems. It will allow easy storage and access to both raw and refined data through a simple end-user interface. With informatics, users will be able to integrate old and new data with simulation techniques. Data will be verifiable, and users will have confidence in the quality and accuracy of results.*
>
> *Informatics in 2020 will enable a paradigm shift, changing the mind set of researchers and accelerating a trend towards combinatorial approaches to solve all problems. Researchers will move from a hypothetical/deterministic approach to solving problems to a more probability-based approach. Informatics will be a pervasive, routinely accepted tool and will provide a decision-making and knowledge infrastructure for R&D. It will facilitate and enhance the entire process from discovery through commercial product development, and will be recognized as a critical step in R&D to generate value and provide a competitive advantage.*
>
> *Supporting informatics will be a common framework and architecture to integrate applications, data and instrumentation encompassing a variety of sources and disciplines. This will foster an environment of open collaboration through standard data structures and platform independence, with data stored in a central repository.*

The key components of future informatics systems are shown in Exhibit 7-1. Attaining these components will require advances in current technology, as well as the development of entirely new technology.

Exhibit 7-1. Key Components of Informatics

- Quality of data/knowledge (quality assessment)
- Databases
- Data warehousing
- Data mining
- Data fragmentation
- Data transformation
- Visualization
- Data fusion
- Data integration (imagery with spectroscopic)
- Data acquisition
- Functional databases versus phenomonological databases
- Data-to-knowledge conversion
- Data representation
- Data access
- Local versus global informatics (distributed)
- Artificial Intelligence
- System asks questions of human and itself
- Query optimization
- Resolution of semantic discrepancies (e.g., thesaurus)
- Modeling
- Mathematics/statistics
- Information tracking
- Clarity of goals
- Experimental control systems
- Analytical interfaces
- Human interface with data
- Data analysis feeds control systems in real time
- Interpretation

Performance Goals

Quantitative goals for informatics focus primarily on revolutionizing the process and costs involved with new product research and development. Specifically this means reducing the cost of manufacturing, the time associated with discovery and R&D decision-making, and effectively integrating downstream manufacturing information with product development. Through informatics, products will get to market faster, they will be cheaper to develop, and markets will expand.

To support the successful use of informatics, goals were identified to make the data within informatics readily accessible based on a reasonable economic structure (paying a subscription, fees for data), and to double the amount of data publically available.

Many of the goals identified for informatics can potentially be achieved within the near-mid term, or within the next 10 years. As these goals are achieved, the value of combinatorial informatics in the research process will be increased. The goal is for informatics to be recognized as 80% of the value of total R&D programs. By 2020, it is hoped that all winners of R&D 100 Awards for materials will be achieved through the use of combinatorial informatics in the development process.

Informatics Goals

Near-Mid Term
- Reduce manufacturing costs by 50% by 2005 through the use of informatics
- Reduce the time to make a decision to pursue a product (stage gate decision) by a factor of 10
- Accelerate the time to market for materials by 2-fold by 2005
- Reduce the needed market size for product launch by 50% by 2010 through cost reduction
- Informatics becomes the physical basis (tool) for the physical science information infrastructure (similar to the medical field information sources) by 2005

Long Term (2020)
- Combi-fed informatics is recognized as 80% of the value of the total R&D program
- All winners of the R&D 100 Awards for materials are developed via the combinatorial approach using informatics
- Efficiency of using informatics for discovery increases

Improving the efficiency of using informatics (ratio of discoveries/ number of experiments or dollars) is a long term goal. Today, the number of experiments is rising. Informatics will allow the rise in number of experiments to stop by 2010 and start a downward trend in the number of experiments. Ultimately, informatics will allow researchers to do only one experiment.

Barriers to Entry

One of the most critical **technical barriers** to the use of informatics is current database design, which was developed to store information, not necessarily retrieve and use it (see Exhibit 7-2). The lack of new algorithmic approaches to information handling and database design is another limiting factor—many problems are currently solved by brute force. The capability of current visualization technology is significantly limited as a tool for informatics, and human interface with visualization media is poor. Vendors have not shown much interest in developing different techniques and tools for visualization.

Exhibit 7-2. Critical Barriers to Achieving Goals for Informatics

Technical	**Databases are designed to store information, not necessarily retrieve it** **Lack of new algorithmic approaches** **Much more can be done with 2D than with current visualization** – **Inadequate visualization media** – **Lack of interest in different visualization approaches and tools** – **Visualization should be quantitative** **Lack of error bars/confidence thresholds on data** Disjointedness of materials databases Lack of standards for data structure, systems, interfaces for analytical instruments Inadequate/non-existent integration of length and time scales Inadequate decision-making capabilities for computers, ineffectiveness of current artificial intelligence Lack of standard problem formulation Materials system complexity Lack of predictive experimental tests Lack of semantics Current process of database design and implementation
Cultural and Economic Issues	**Importance of informatics is underestimated** – **Insufficient understanding of IT by materials scientists and vice versa** – **Lack of clarity about the nature of informatics** – **Little recognition that informatics is necessary** – **Informatics is not seen as a science** **Lack of drive from upper management to pursue combi methods** **Poor clarity of relationship of informatics to the bottom line** Diverse goals of user community Intellectual property issues
Infrastructure	**Lack of informatics in education programs** Shortage of skilled people to educate students in the field (e.g., bioinformatics) Combi mathematics has no reference to materials, and no linkages

Note: **Bold = top priority;** Regular type = high priority

In terms of data, a significant barrier is the lack of error bars or confidence thresholds. This would add to the credibility of results and improve the ability to select promising leads. In general, the quality and quantity of databases available for materials is inadequate and lacks cohesiveness. Exacerbating this problem is the lack of standards for data structures and interfaces with analytical instruments. Researchers often must deal with data obtained from many different sources and with no common structure.

There are significant limitations in some of the current information technology that could be applied to informatics. Some examples of this are inadequate decision-making capabilities in computing platforms, and the ineffectiveness of existing artificial intelligence systems. Another issue is that there are no standards for formulation for problems.

A number of **cultural and economic barriers** also inhibit the development and use of informatics. The most critical of these is that the importance of informatics in the research process is seriously underestimated, and there is a general distrust of the field in the traditional sciences. Part of the problem is that information scientists have a good understanding of their field, but not of materials science, and vice versa. Generally, many in the scientific community do not have a clear understanding of informatics and its role in the research process, and do not view it as a science *per se.* Many do not understand the basic need for informatics, and that organization of data is as important as collection of data, and is also critical to interpretation of data. At the management level, there is little motivation to pursue combinatorial methods in general, including informatics, because the economic benefits are not clear and the field is not well known.

A critical barrier to informatics is the lack of an **infrastructure** for providing scientists skilled in the field. There are few programs in universities dealing with informatics. This is primarily due to the serious shortage of skilled people who can educate students in the field (e.g., bioinformatics), and the lack of funding from both private and public sources for this type of research in academia. Another factor is that current theory is not well-linked to some of the physical sciences. Combinatorial mathematics, for example, currently have no linkage or reference to materials science.

Research Needs

The priority research needed to make advances in informatics is shown in Exhibit 7-3. A major challenge for informatics will be developing more effective database design and methods for interpreting data, along with enabling models and theoretical tools.

Database Design and Interpretation—In the near-term, developing alternatives to existing data visualization technology is one of the most critical areas of need. Improved visualization will enhance the ability to identify promising leads and desirable compositions. Another immediate top priority is the development of tools that can integrate and analyze both structured and unstructured data. Supporting research is needed to develop new techniques for database aggregation and federation, rather than the conventional approaches which rely on integration of various data types. Systems are needed to enable textual data mining and compression, and data mining in relational databases. In the mid-term, to facilitate library characterization, methods are needed to enable comparison and contrast of model libraries with those created through experiment. Over the long term, a top priority is the development of alternatives to current relational database technology. An important component of an alternative would be architectures that can be updated in real-time. New technology should also reduce the work required for database design, implementation and modification.

Exhibit 7-3. Priority Research Needs for Informatics

DATABASE DESIGN AND INTERPRETATION

Near-Term (0-3 yrs)	*Mid-Term (>3 -10 yrs)*	*Long-Term (>10 yrs)*
T Alternatives to existing data visualization technology T Tools to integrate and analyze structured and unstructured data H Textual data mining and compression H Better enabling of data mining in relational databases H New methods for database aggregation and federation, rather than integration for various types of data	M Better methods to compare and contract model libraries vs experimental libraries	T Alternatives to current relational database technology, including 1) dynamically up-datable architectures, 2) systems that reduce work in database design, implementation and modification.

ENABLING MODELS/THEORY

Near-Term (0-3 yrs)	*Mid-Term (>3 -10 yrs)*	*Long-Term (>10 yrs)*
H New algorithmic approaches for complex multi-source materials data M Mathematical R&D to drive development of informatics for materials design M Modeling and simulation tools	T Computer-based decision-making methods T Modular combinatorial tool kit, which includes an infrastructure for many different applications M Apply optimization and analysis to informatics, to enable robustness to 6 sigma quality	

INFRASTRUCTURE

Near-Term (0-3 yrs)	*Mid-Term (>3 -10 yrs)*	*Long-Term (>10 yrs)*
T Curriculum for teaching informatics in universities	T Standard interface for instrumentation in general T Standard experimental tools H Enduring access to data generated today, through national organizations, digital libraries H Legally binding electronic notebooks M Reliable experimental databases (e.g., calibration, x-ray assumes flat sample) M New standards and better management of standard-setting process	

T = Top priority; H = High priority; M = Medium priority

Enabling Models and Theory—Supporting models and theory are needed to build an effective, usable informatics system. A top priority in modeling is the development and implementation of computer-based decision-making techniques to facilitate the selection process for "hits". In the near term, a high priority is the development of new algorithmic approaches for handling materials data from a variety of sources. This can be a particularly difficult challenge, considering the complexity and quantity of data that will be generated. To enable the widespread use of informatics, an overarching top priority is to create a modular combinatorial tool kit which is flexible enough to be used for a multitude of applications, and by a diverse set of end-users. This will require a variety of modeling and simulation tools, along with developments in mathematical theory that can be applied specifically to materials design.

Infrastructure—There are a number of critical areas where research or other efforts must be pursued to ensure that an adequate infrastructure exists to support a viable informatics approach. First, a curriculum for teaching informatics must be implemented in universities, to ensure that scientists with an appropriate skill set will be entering the work force over the next ten years. Second, a standard interface must be established for instrumentation in general, and in particular for instruments used in library production and characterization. Supporting this would be better management and design of standard-setting processes. Third, consistency in experimentation must be implemented through the development of a set of standardized experimental tools. To support this, techniques should be implemented to create reliable experimental databases. Last, a system must be established to ensure enduring access to the data that is generated today and in the future. This could potentially be done through national organizations and digital libraries. Other important issues include addressing intellectual property concerns, possibly through legally binding electronic notebooks.

Priority Non-research Activities/Strategies—In addition to research, there are a number of activities and strategies that should be pursued to enable the successful development and use of informatics (see Exhibit 7-4). In the near-term, a top priority is to increase the presence of informatics in the scientific community through articles in journals, papers at conferences and symposium, and similar venues. Along with increased awareness should come ways to articulate the cost and benefits of informatics to the business community, particularly investment decision-makers (both technical and non-technical). To do this, economic metrics will need to be established to measure the results and impacts of informatics. An initial focus should be on current process technologies and applications that can more easily illustrate the bottom line benefits. As informatics becomes integrated into research, publically-accessible demonstration experiments should be made available.

One top priority is to encourage the use of more complex systems and approaches into combinatorial informatics. A complementary approach is to implement a Toyota-type production system which incorporates elements such as make-on-demand and waste minimization. Over the mid-longer term, a priority is to solve problems at the highest level, then at the more practical or realistic level, and to use the data obtained to increase knowledge of how to refine the informatics system.

Exhibit 7-4. Priority Non-Research Activities/Strategies for Informatics

NEAR-TERM (0 - 3 Years)

- **Cultivate informatics community (journals, conferences)**
- **Foster introduction of complex systems and approaches into combinatorial informatics**
- **Establish economic metric to measure results and impacts of informatics**
- Implement Toyota production system in combinatorial methods (make-on-demand, and one-at-a-time, minimize waste, lowest possible cost)
- Focus more on current process technologies and applications to illustrate bottom line benefits

MID-TERM (>3 - 10 Years)

- **Go from high-level to realistic-level to solve specific problems and learn**
- Develop ways to articulate cost/benefit of informatics
- Get business schools to make a business case for informatics
- Develop publically accessible demonstration experiments

Note: **Bold face** = top priority; Regular type = high priority.

8 The Road Forward

Common Themes

Combinatorial methods can facilitate the discovery of new materials, and help to optimize the processes used to manufacture them. They can also be used for increasing productivity and profit margins, improving environmental performance, and meeting the challenges of globalization. The common themes and goals for the future of combinatorial methods are shown in Exhibit 8-1. These illustrate the integral success factors that will impact all aspects of the combinatorial system—library production, library characterization, informatics, experiments and models.

Exhibit 8-1. Common Vision Themes and Goals for 2020

- Cost of combi must be reduced
- Combi should be used to explore complex systems
- High throughput experiments should be possible throughout the organization, with understanding of output
- Combi should be explored for process-based applications
- Students should be combinatorially-educated
- All groups (industry, academia, government) should be engaged in combi efforts
- Combi should be web-enabled
- Molecular modeling should be coupled with combi
- Materials supply should be seamlessly integrated with end-user needs
- Part of materials funding should be extended to combi
- Faster scale-up must be achieved
- Combi should have the capability to solve pilot-plant/scale problems
- There must be completely new thinking about how discovery is done
- External factors must be considered in combi development (e.g., population growth, China as world power, human genome)
- The number of products that need to be screened to get to market should be quantified
- 5-way combinations should be attained
- Computers should be loaded with algorithms, models, data, etc., and used routinely for screening/predictions

The Role of Government

While much activity has been undertaken in the commercial sector for combi drug discovery, it is clear that precompetitive research and development may be needed to advance the state-of-the-art of combinatorial methods for materials use. There are, in fact, significant opportunities for the government to play an important role in accelerating advances in combi techniques (see Exhibit 8-2). For example, Federal funds can be used to support national test facilities and equipment for developing new characterization methods as well as equipment for handling and processing samples. Work at these facilities can facilitate the development of standardized methods and tests to help ensure consistency in library production and information transfer. Federal funds can also be used to help industry access the unique capabilities of the national laboratories and various universities involved in combinatorial research. Federal programs like Small Business Innovations Research (SBIR), which provide seed

Exhibit 8-2. Library Production: Government Role

High Priority Research	Government Role	Impact of No Government Funding
Fast, nondestructive methods for morphology characterization **Equipment for processing and handling library samples**	• Provide national test facilities and equipment • Fund specialized expertise in national labs and universities • Develop standard methods • Fund smaller equipment companies through SBIRs	• Lose standardization power • Intellectual property issues and problems • More spotty coverage of materials • Loss of U.S. competitive edge in new technology area • Significantly slower growth in emerging field
Automated, parallel synthesis methods for raw materials	• Fund research proposals that solicit paradigm change	• As above
Virtual library	• Fund specialized expertise in national labs, universities • Develop standard methods	• Resource-intensive—company will not do • As above
Improved educational curricula or consortia	• Fund curriculum development • Fund Sematech-like consortia	• Significantly slower growth in emerging field • Loss of U.S. competitive edge in new technology area
Strategic/intellectual property issues	• Provide leadership at the patent office	• Current situation which impedes patents

money to small companies for promising inventions, can also be a source of funding for smaller equipment companies interested in developing specialized combinatorial instruments. Federal funds are essential for resource-intensive concepts like a virtual library, which few individual companies have the resources to fully develop. In academia, Federal funds can be used to fund development of curricula that supports the multi-disciplinary combinatorial concept.

Without Federal funding, new technology will be developed much more slowly, and the competitive edge of the U.S. in this arena could suffer. Lack of Federal funds will also make it more difficult to achieve standardization, as individual companies will continue to work in isolation developing unique systems. Federal funding for generic, pre-competitive research will help to alleviate intellectual property issues, and permit more data to be available to a broader range of industry customers.

Bibliography

J.D. Hewes, “High Throughput Methodologies for Chemicals & Materials Research, Development and Engineering,” U.S. Department of Commerce, National Institute of Standards and Technology, March 2000.

D.S. King, “Combinatorial Methodologies: New Tools for the Chemicals and Materials Industries,” U.S. Department of Commerce, National Institute of Standards and Technology, April 2000.

John K. Borchardt, “Combinatorial Chemistry: Not Just for Pharmaceuticals,” *Today’s Chemist at Work*, November 1998, 7 (10), 34–41, American Chemical Society.

Stu Borman, “Combinatorial Chemistry,” *Chemical & Engineering News*, April 6, 1998, American Chemical Society.

Stu Borman, “Combinatorial Chemistry—Researchers continue to refine techniques for identifying potential drugs in ‘libraries’ of small organic molecules,” *Chemical & Engineering News*, February 24, 1997, American Chemical Society.

“Combinatorial Encoding Technique: A Lucky Charm,” Chemical and Engineering News, July 24, 2000 (pps 40–41).

Appendix A. Workshop Participants

Jay Akhave, Avery Dennison
Joe Akkara, NSF
Eric Amis, NIST
John Babiak, Pharmacopeia, Inc.
Joel S. Bader, CuraGen Corp
Peter Brooks, Symyx
David Casebier, ArQule
Richard Cavanagh, NIST
James Cawse, GE
Ray Chrisman, Dow Chemical
Michael Cima, MIT
Rich Colton, Naval Research Lab
A. Gonis, Lawrence Livermore NL
Brian Goodall, Albemarle
Laurel Harmon, Nonlinear Dynamics Inc
Jim Harness, Mettler-Toledo Bohdan, Inc
Dennis Healy, DARPA
Steve Heilmann, 3M
John Hewes, NIST
Joe Hogan, Alveus
Jennifer Holmgren, UPO LLC
Andrew Kaldor, ExxonMobil
Alamgir Karim, NIST
David King, NIST
Doug Kiserow, ARO
Sundaram (Kris) Krishnamurthy, Kodak
John Lau, PQ Corp
Don Lauffer, Phillips Petroleum
Terry Leib, GE CRD
Steve M. Levine, MSI-Pharmacopeia
Gregory Lewis, UOP LLC
Nate Lewis, California Institute of Technology
Erik Lifshin, GE CRD
Andy Lovenger, NSF
Suresh Mahajan, Catalytica
Wilhelm F. Maier, Max-Planck-Institut für Kohlenforschung
Costas Maranas, Penn State
Ian Maxwell, Avantium
David Myers, WR Grace
Rajesh Nair, GSE Systems
J. Narayan, NC State
John Newsam, hte
Robert Pinschmidt, Air Products & Chemicals
Geoffrey Prentice, NSF
Krishna Rajan, RPI
Roberto J. Rioja, Alcoa Technical Center
John Rodgers, Toth Information Systems
Dave Rothman, Dow Chemical
Sab Sabesan, DuPont
Ray Salemme, 3DP
Krishnan Sankaran, Boeing
Lida Sethi, Chemical Abstracts Services
Raveen Sharma, Spotfire
Jack Solomon, Praxair
Brian Southern, GSE Systems
D.W. Smith, Jr., Clemson University
Al Sylwester, Sandia NL
Ichiro Takeuchi, U. Maryland
Liong Tee, BP Amoco
Windy Thompson, Glaxo Wellcome, Inc.
David B. Wallace, MicroFab Tech, Inc.
Ted Williams, Symyx
Bob Wilson, SRI
X.D. Xiang, Lawrence Berkeley NL
Eric Zeigel, BP Amoco

Appendix B. Detailed R&D Tables

Exhibit B-1. Research Needs to Overcome the Barriers to Library Production

✪ = Top Priority; ● = High Priority; ❍ = Medium Priority

Time Frame	Education	Equipment for Library Production/ Sample Handling	Raw Materials/ Chemical Synthesis	Scalability (Performance and Process)	Strategic Issues
NEAR (0-3 Years)	• Increase exposure of combi chem in materials/chemical trade associations – MRS, ACS, AIChE, SAMPE, APS, ACerS ✪✪✪✪●● • Lobby government for combi chem funding and support (both state and Federal) ●●❍❍❍❍ • Introduce formal programs in universities (NSF funding) ●●❍❍ • Funds for joint industry/ academia cooperation ●●❍ • More inexpen- sive workshops on combi chem that academia can attend ● • Establish industry/ academic R&D consortia analogous to Sematech (e.g., combitech) ❍❍ • Educate chemical engineers ❍	• Develop technology for precise handling and mixing of powders and highly viscous liquids at small scales, handle wide temperature ranges, pH ranges, and pressure (near to long) ✪●●●●❍❍ • Develop a series of simple to complex, mult-reactor automated devices ●●❍❍ • Develop micro-arrayed calorimetry ●❍❍❍ • Develop equipment for processing, handling and storage of library samples (near to mid) ●❍❍❍ • Develop fiber optic spectroscopy for in-situ measurement ❍❍❍❍ • Develop improved procedures for making and evaluating air-sensitive compounds when making libraries ❍❍ • SBIR/STTR/ ATP grants for equipment companies ❍		• Develop fast, nondestructive methods for morphological characterization – Substrate and solvent effects – Correlations between thin film and bulk properties ✪✪✪✪●●❍❍	• Education of patent office (limit wildly broad claims) ●●●●❍❍ • Application-oriented patents only ●●❍❍❍ • Use consortium approach to help with info sharing ❍❍
MID (> 3-10 Years)		• Develop robust, cheap sensors (mid to long) ❍❍❍ • Apply split and pool technique to materials research (mid to long) ❍❍	• Develop automated and parallel synthesis methods to address need for raw materials ✪●●❍❍❍❍❍❍	• Develop scientific understanding to map miniaturization to lab scale performance – Understand limits – Translatability from lab to pilot ●●●●●❍❍	
LONG (> 10 Years)					**Virtual Library** • Develop predictive tools ●●❍❍❍❍

Exhibit B-2. Research Needs in Library Characterization and Properties Analysis
✪ = Top Priority; ● = High Priority; ❍ = Medium Priority

Time Frame	System Integration, from Sample-to-Enduse	Mechanical Properties	Sample Prep	Structure	Broad Tools	Other
NEAR (0-3 Years)	• Integrate analytical equipment with reactors ●●●●●❍❍❍	• Develop tool for measurement of stress/strain curves for 100 micro samples per day ●❍❍❍❍	• Experiments/ samples that are compatible with testing ❍ • Miniaturized sample prep/mixing of non-liquid samples ❍	• Structural characterization tools specifically designed for high throughput screening (robust, sensitive, accurate) • X-ray system for structure and composition analysis ✪✪✪❍	• Tools to measure catalyst performance ●●❍❍ • Low cost modular units (for formulation/ mixing, multiple property testing (gloss, hardness, etc.) ● • Quantitative assessment of bio-compatibility	
MID (>3-10 Years)	• Microfluidic reactors coupled to detectors ✪● • Real time tools for measure- ments of equilibration ●❍❍ • Increase fundamental understanding of continuous solids ❍ • Develop a series of “nose” type sensors to assist transfer functions for improved performance properties (use continuous vs batch experiments) ●●●❍❍	• Develop tool for high throughput tensile testing of 100 micro samples per day ✪●❍❍	• High throughput compatible methods for measuring surface energy on uneven surfaces (e.g., powders) ❍❍❍	• Defect and grain structure analysis, micro to macro property prediction ❍	• Tool for high throughput porosity and pore size screening (100 samples/day) ●●	

Exhibit B-2. Research Needs in Library Characterization and Properties Analysis
✪ = Top Priority; ● = High Priority; ❍ = Medium Priority

Time Frame	System Integration, from Sample-to-Enduse	Mechanical Properties	Sample Prep	Structure	Broad Tools	Other
LONG (> 10 Years)	• Combine micro-electronics and MEMS for on-board measurement in reactor cells ✪ • Work station approach to integrate sample prep, library production, reaction analysis, data management ●●●●●●❍❍ • Transfer function to correlate between properties and performance ❍	• Develop tool for accelerated aging tests capable of 100 micro samples per day ●		• Measurement tools for anisotropic structure and property ❍ • Rapid x-ray differentiation and fluorescence techniques for small samples—must be fast and give nonambiguous results ❍		
ONGOING		• Rhealogical measurements on small samples ●❍❍❍❍		• Structure property mapping of ternary systems ❍	• Tools to measure material/ product performance at multiple scales/ performance based assays ✪✪ ✪✪❍❍ • Multiple scanning tips techniques for parallel determination of physical properties (speed) ❍❍❍❍	

Exhibit B-3. R&D Needs for Informatics
✪ = Top Priority; ● = High Priority; ❍ = Medium Priority

Time Frame	Cultural	Enabling Models and Theory	Infrastructural	Database Design and Interpretation
NEAR (0-3 Years)	• Establish measurable economic metric to measure result/impact of informatics ✪✪✪● • Foster introduction of complex systems and approaches into combi informatics✪●❍❍❍❍❍ • Cultivate informatics community (journal, conferences)●●❍❍❍❍ • Implement Toyota production system in combi ● – Make on demand – Make one-at-a-time – Waste minimization – Lowest possible cost • Focus on current technologies and applications to make bottom line benefits more obvious ❍❍	• Develop new algorithmic approaches for complex multi-source materials data ●●❍ • Pursue mathematical R&D that is needed to drive development ❍❍❍ • Modeling and simulation tools ❍	• Develop curriculum for teaching informatics in colleges ●●●	• Develop alternates to the data visualization that exists today ●●●❍❍ • Tools to integrate and analyze structured and unstructured data ●●❍❍ • Textual data mining and compression ●❍❍❍ • Better enablement in data mining in relational databases – Better tools in database—must get in with database vendors to direct mining ●❍❍ • Develop new methods for database aggregation and federation, instead of integration for various types of data ●❍❍
MID (> 3-10 Years)	• Go from high-level to realistic-level to solve specific problems and learn ✪ • Develop ways to articulate cost/benefit of informatics; educate business community ● • Include economics/ financial scientists in the actual science of informatics/combi ❍❍ • Get business schools to make a business case for informatics ❍❍ • Develop publically accessible demonstration experiments ❍	• Computer-based decision-making methods ✪✪✪●●●❍ • Modular tool kit—infrastructure is there for many different applications ●●●❍ • Optimization and analysis applied to informatics— robustness to 6 sigma quality ❍	• Develop standard interface for instrumentation in general ●●❍❍❍ • Develop standard experimental tools ●●❍❍ • Develop enduring access to data generated today – National organizations – Digital libraries ●❍❍ • Develop electronic notebooks that are legally binding ❍❍❍❍ • Develop reliable experimental databases (e.g., calibration, x-ray assumes flat sample) ● • Develop standards and enhance/manage process by which they are set ❍ • Feasibility studies by regulatory/government to set funding priority ❍	• Better methods to compare/ contrast model libraries versus experimental libraries ❍

Exhibit B-3. R&D Needs for Informatics
✪ = Top Priority; ● = High Priority; ❍ = Medium Priority

Time Frame	Cultural	Enabling Models and Theory	Infrastructural	Database Design and Interpretation
LONG (> 10 Years)				• Develop alternative to current relationals database technology – Reducing work in database design, implementation, and modifications – Dynamically updatable architectures ✪✪✪●●●❍❍❍❍

Exhibit B-4. Research Needs for Design of Experiment
✪ = Top Priority; ● = High Priority; ❍ = Medium Priority

Time Frame	Modeling Tools	Hardware	Experimental Methods	Economic/Cultural Shifts	Advanced Computer Capabilities
NEAR (0-3 Years)	• Develop algorithms for rapid search capabilities ✪✪●❍❍❍ • Develop technology to query data bases ❍❍❍❍ • Integration of content ✪ • Develop CFD tools – Needs to be accessible ❍❍		• Develop experimental methodologies that are repeatable ✪●❍❍ • Develop tracking mechanism for samples ❍ • Set up protocols/ standards for data ❍	• Develop multi-disciplinary needs – Establish collaborative efforts – Consortiums to collect/share data and results ●●●❍❍❍❍ • Find analogous applications for “paradigm shift” (e.g., parallel computers) – Find commonalities in other fields (geology, etc.) – Find experts willing to tackle these problems – Universities – Raise the profile of combi-chem – Introduce serendipity ✪●●❍ • Develop financial modeling tools ❍	• Problem taxonomy ✪❍❍❍❍ • Capability to handle larger models ●●● • Pattern recognition ● • Scalable model ❍❍
MID (> 3-10 Years)	• Experimentally enable theory development for new models ●●●●●❍❍ – Create libraries for data and knowledge – Manage the knowledge—access it, get it quickly • Models that consider time dependence – Accelerates studies – Life-cycle stability – Accelerated testing applied to chemistry (GE) – ASTM-type testing ●❍❍❍❍ • Integrate tools ●❍❍❍❍ • Develop process modeling tools ❍ • Build better models for continuum mechanics	• New hardware – Micro-reactors – New analytical methods—faster – Micro-channels – Develop “expert-system” type of product to guide new users, order of process ●●●●●●❍❍ ❍❍❍❍ • Develop a central nervous system—system architecture that is flexible ✪●●❍❍❍	• Experiment design based on reaction pathways ✪✪✪✪✪●●❍❍ • High throughput analysis of assay data ●❍❍	• Develop more of an “on-line” process – Use of computers – Visualization software and hardware ❍❍	• New efficient descriptor generation ❍❍ • Properties that experiments measure ❍

Exhibit B-4. Research Needs for Design of Experiment
✪ = Top Priority; ● = High Priority; ❍ = Medium Priority

Time Frame	Modeling Tools	Hardware	Experimental Methods	Economic/Cultural Shifts	Advanced Computer Capabilities
LONG (> 10 Years)	• Develop models that span multiple scales – Predict one level to another scale-ups ✪●●●●●❍❍ • Develop "fuzzier" search capabilities—learning engines ●❍				• New/improved computational algorithms ●❍❍