

Transforming Glycoscience: A Roadmap for the Future

A new focus on glycoscience, a field that explores the structures and functions of sugars, promises great advances in areas as diverse as medicine, energy generation, and materials science, this report finds. Glycans—also known as carbohydrates, saccharides, or simply as sugars—play central roles in many biological processes and have properties useful in an array of applications. However, glycans have received little attention from the research community due, in large part, to a lack of tools to probe their often complex structures and properties. This report presents a roadmap for transforming glycoscience from a field dominated by specialists to a widely studied and integrated discipline, which could lead to a more complete understanding of glycans and help solve key challenges in diverse fields.

Throughout the natural world, glycans are ubiquitous. Every living cell is covered with a glycan layer that provides a cell- and tissue-specific identity, and many proteins receive glycan tags that can act like a dimmer switch to modulate protein activity (see Figure 1). Glycans also play key roles in the reactions of photosynthesis and metabolism. In recent years, scientists have realized that understanding glycans' biological roles can yield advances in medicine; their polymer structures can produce renewable materials with useful properties; and the glycans found in plant cell walls can be processed to make biofuels. Yet despite these potential

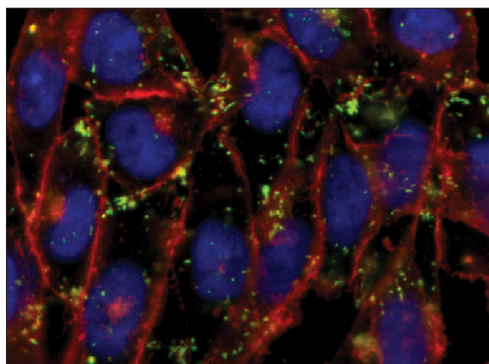


Figure 1. Cell surfaces are covered with a layer of glycans. Here, glycans on Chinese hamster ovary cells are imaged in red, and cell nuclei and Golgi apparatus are blue and green, respectively. Source: Carolyn Bertozzi, Scott Laughlin, and Jeremy Baskin.

applications, glycans remain less studied than other molecules of life, such as DNA and proteins (see Box 1).

Glycoscience in Health, Energy, and Materials Science

The committee investigated three key areas in which glycoscience can make significant contributions:

Health

Glycans play roles in almost every biological process and are involved in every major disease. They are integral to cell adhesion and movement, such as when white blood cells migrate to a site of injury or infection. Glycans inside cells help influence the expression of genes and proteins, forming part of a cell's response to biological signals. The human immune system has evolved the ability to recognize pathogen-specific glycan sequences on bacteria, viruses, and fungi, triggering an immune response. These properties mean glycans are useful as components of therapeutic drugs to help treat chronic and infectious diseases and as biomarkers to detect diseases like cancer.

Glycoscience can also contribute to personalized medicine. In recent years, there have been substantial investments in studying the human genome to develop disease treatments

Box 1. What are Glycans?

Glycans—also known as carbohydrates, saccharides, or simply as sugars—are molecules made up of combinations of different sugar units, such as glucose. These units can link together to form complex, three dimensional structures. Unlike DNA and proteins, glycans are not created by following a template. Instead, the reactions that link individual sugar units together are influenced by factors including cellular metabolism, cell type, developmental stage, and nutrient availability. These factors provide substantial diversity and allow for glycans with a wide array of properties, but also make glycans more difficult to study and manipulate in the laboratory.

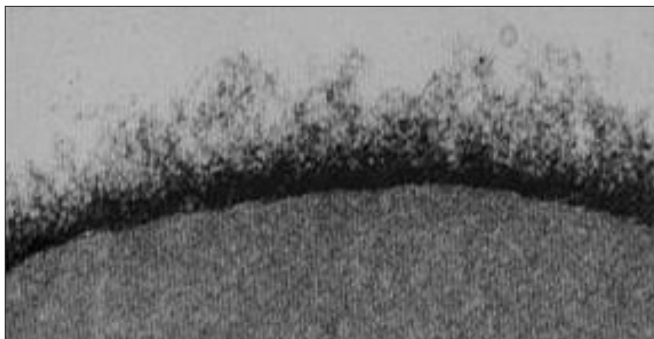


Figure 2. The glycan layer on the surface of a red blood cell. Source: Voet and Voet, *Biochemistry*, John Wiley and Sons, Inc.

tailored to the individual. Because an individual's glycome—the full complement of glycans in cells and tissues—helps mediate the transfer of information from genotype to phenotype, glycoscience could help develop personalized medical care.

Energy

The reactions of photosynthesis capture the sun's energy and convert it to sugars that fuel the growth of plants and the animals that feed on them. Some of these sugars are used to make plant cell walls—structures composed of glycans, structural proteins, and lignin. Large-scale generation of biofuels from biomass requires an efficient process to convert the energy stored in plant cell walls into liquid fuels. However, a phenomenon called recalcitrance—the resistance of plant cell walls to the degradation that's necessary to convert biomass into fuel—makes current methods for producing biofuel expensive.

Glycoscience may be able to help develop enzymes and other catalysts to improve this conversion. Because cell wall structures can vary between plant species and even between the different tissues of a single plant, a better understanding of plant cell wall production could also help scientists select or manipulate crops to maximize biomass production, while minimizing inputs of fertilizer and water.

Materials

Many of the most widely used plastics are made from the polymers found in petroleum. But with concerns about petroleum cost and supply, there is increasing pressure to find alternative sources of polymer materials.

Scientists are turning to the glycan-based polymers found in nature as alternative sources of material feedstocks. Glycan-based polymers can be purified, altered chemically to provide a range of new properties, or even broken down to their constituent sugar units to make chemical precursors for use in the chemical and engineering industries. The properties of

glycan-based polymers could even outperform petroleum-based products at lower cost for some applications. For example, the glycans that make up the exoskeletons of crustaceans and insects and the woody parts of plants are strong, yet lightweight and flexible.

A Roadmap to Advance Glycoscience

The report's authoring committee identified needs for a glycoscience toolkit and outlined a roadmap to fill these gaps and ensure that interested researchers can effectively incorporate glycoscience in their work, along with a timeline of specific roadmap goals for the next 5 to 15 years.

1. Glycan Synthesis

While significant advances have been made in glycan synthesis, these technologies are only available in specialized laboratories and produce only small quantities of a specific glycan.

For glycoscience to advance, widely applicable methods to generate both large and small quantities of glycans are needed. This capability would allow researchers to characterize the glycomes of various organisms, study glycans isolated at the intermediate steps of reactions, and produce large quantities of glycans for use as therapeutic drugs or in new polymer materials.

The development of transformative methods for the facile synthesis of carbohydrates and glycoconjugates should be a high priority for the National Institutes of Health, the National Science Foundation, the Department of Energy, and other relevant stakeholders.

Roadmap Goals:

- Within 7 years, have synthetic tools to be able to synthesize all known carbohydrates of up to eight carbohydrate units, including substituents (e.g., acetyl, sulfate groups). This goal encompasses all human glycoconjugates, which are currently estimated to be 10,000-20,000 structures, along with plant and microbial glycans and polymers.
- Within 10 years, have synthetic tools to be able to synthesize uniform batches, in milligram quantities, of all linear and branched glycans that will enable glycan arrays for identifying protein binding epitopes, provide standards for analytical methods

Box 2. The Glycoscience Toolkit

Although much can be accomplished with the tools already available, gaps remain in the glycoscience toolkit in areas such as the synthesis of glycans and analytical techniques to determine glycan structure and function.

development, and enable improved polysaccharide materials engineering and systematic studies to be conducted. This includes methods for synthesis of structures with isotopic enrichment of specific desired atoms that may be needed for a wide variety of studies.

- Within 15 years, be able to synthesize any glycoconjugate or carbohydrate in milligram to gram quantities using routine procedures. Community access should be available through a web ordering system with rapid delivery.

2. *Glycan Analysis*

A suite of tools analogous to those available for studying nucleic acids and proteins is needed to detect, describe, and purify glycans from natural sources, and characterize their chemical composition and structure.

The development of transformative tools for detection, imaging, separation, and high-resolution structure determination of carbohydrate structures and complex mixtures should be a high priority for the National Institutes of Health, the National Science Foundation, the Department of Energy, and the Food and Drug Administration, and other relevant stakeholders.

Roadmap Goals:

- Over the next 5-10 years, develop the technology to purify, identify, and determine the structures of all the important glycoproteins, glycolipids and polysaccharides in any biological sample. For glycoproteins, determine the significant glycans present at each glycosylation site. Develop agreed upon criteria for what constitutes the acceptable level of structural detail and purity.
- Within 10 years, have the ability to undertake high-throughput sequencing of all nitrogen- and oxygen-linked glycans from a single type of cell in a single week.
- Within 10 years, have the ability to routinely determine the complete carbohydrate structure of any glycan or polymer repeat sequence including branching, anomeric linkages between glycans, and substituents.
- Within 15 years, have the ability to determine glycoforms (a complete description of molecular species within a population that have the same polypeptide sequence) of any glycoprotein in a biological sample.

3. *Molecular Modeling*

Continued advances in molecular modeling can generate insights for understanding glycan structures and properties.

Robust, validated informatics tools should be developed in order to enable accurate carbohydrate and glycoconjugate structural prediction, computational modeling, and data mining. This capability will broaden access of glycoscience data to the entire scientific community.

Roadmap Goals:

- Within 5 years, develop an open-source software package that can automatically annotate an entire glycan profile (such as from a mass spectrometry experiment) with minimal user interaction.
- Within 5 years, develop the technology to perform computer simulations of carbohydrate interactions with other entities such as proteins and nucleic acids.
- Within 10 years, develop the software to simulate a cellular system to predict the effects of perturbations in glycosylation of particular glycoconjugates and polysaccharides.

4. *Glycoenzymes*

An expanded toolbox of enzymes and enzyme inhibitors would help scientists produce, degrade, and study the function of glycans of interest, driving progress in many areas of science.

The development of transformative capabilities for perturbing carbohydrate and glycoconjugate structure, recognition, metabolism, and biosynthesis should be a high priority for the National Science Foundation, the National Institutes of Health, the Department of Energy, and other relevant stakeholders.

Roadmap Goals:

- Within 5 years, identify the genes involved in glycan and glycoconjugate metabolism in any organism whose genome has been sequenced, and identify the activities of at least 1,000 enzymes that may have utility as synthetic and research tools.
- Within 10 years, be able to use all glyco-metabolic enzymes (e.g. glycosyltransferases, glycosidases, others) as well as other state of the art tools for perturbing and modifying glyco-metabolic pathways (knockouts, siRNAs, etc) of utility to the bio-medical and plant research communities.
- Within 10 years, develop methods for creating specific inhibitors to any human, plant or microbial

Glycoconjugates are glycans linked to other molecules, for example lipids and proteins.

Glycosylation is a type of reaction in which a glycan is attached to another molecule.

glycosyltransferase suitable for in vitro and in vivo conducted in order to perturb the biology mediated by these enzymes.

- Within 15 years, develop imaging methods for studying glycan structure, localization, and metabolism in both living and non-living systems.

5. Informatics and Databases

The development of a centralized, accessible glycan database, linked to other molecular databases, is needed to make full use of the knowledge generated by an expanded effort in glycoscience.

A long-term-funded, stable, integrated, centralized database, including mammalian, plant and microbial carbohydrates and glycoconjugates, should be established as a collaborative effort by all stakeholders. The carbohydrate structural database needs to be fully cross-referenced with databases that provide complementary biological information, for example, Protein Data Bank and GenBank. Furthermore, there should be a requirement for deposition of new structures into the database using a reporting standard for minimal information.

Roadmap Goals:

- Within 5 years, develop a long-term-funded, centralized glycan structure database with each entry highly annotated using standards adopted by the community and all the world's repositories of glycan structures. The database should be cross-referenced and open source to allow the community to develop database resources that draw on this resource and improve its utility to investigators that wish to incorporate glycoscience in their work.

- Within 5 years, employ an active curation system to automatically validate glycan structures deposited into a database so that journals can provide authors with an easily accessible interface for submitting new glycan structures to the database.

6. Education

Although there is increasing recognition of glycans' diverse roles and uses, glycoscience is still seen as a niche research area. Building glycoscience into education would broaden the community of glycoscientists and enable the mainstream scientific community to integrate glycoscience into their research in order to realize the full potential of biological and chemical sciences.

Integrating glycoscience into relevant disciplines in high school, undergraduate, and graduate education, and developing curricula and standardized testing for science competency would increase public as well as professional awareness.

Roadmap Goals:

- Within 5 years, integration of glycoscience as a significant part of the science curriculum would include glycoscience as both lecture materials and hands-on experiments or activities.
- Within 10 years, integrate and teach glycoscience at every level for which it is relevant to understand the scientific content. Competency in glycoscience could also be included in all standardized testing wherever relevant (for example, as part of the SAT and GRE Subject Tests, MCAT, and Medical Board Exams).

Committee on Assessing the Importance and Impact of Glycomics and Glycosciences: David Walt (*Chair*), Tufts University; **Kiyoko F. Aoki-Kinoshita**, Soka University, Japan; **Brad Bendiak**, University of Colorado, Denver; **Carolyn R. Bertozzi**, University of California, Berkeley; **Geert-Jan Boons**, Complex Carbohydrate Research Center, University of Georgia; **Alan Darvill**, Complex Carbohydrate Research Center, University of Georgia; **Gerald Hart**, Johns Hopkins University School of Medicine; **Laura L. Kiessling**, University of Wisconsin, Madison; **John Lowe**, Genentech, Inc.; **Robert Moon**, Forest Products Laboratory, U.S. Forest Service; **James C. Paulson**, The Scripps Research Institute; **Ram Sasisekharan** Massachusetts Institute of Technology; **Ajit P. Varki**, University of California, San Diego; **Chi-Huey Wong**, Academia Sinica, Taiwan and The Scripps Research Institute; **Katherine Bowman**, (*Co-Study Director*), **Douglas Friedman**, (*Co-Study Director*), **Sheena Siddiqui** (*Senior Program Associate*), **Rachel Yancey** (*Senior Program Assistant*), National Research Council.



The National Academies appointed the above committee of experts to address the specific task requested by the National Institutes of Health, the National Science Foundation, the Department of Energy, the Food and Drug Administration, and the Howard Hughes Medical Institute. The members volunteered their time for this activity; their report is peer-reviewed and the final product signed off by both the committee members and the National Academies. This report brief was prepared by the National Research Council based on the committee's report. For more information, contact the Board on Chemical Sciences and Technology at (202) 334-2156 or visit <http://glyco.nas.edu>. Copies of *Transforming Glycoscience: A Roadmap for the Future* are available from the National Academies Press, 500 Fifth Street, NW, Washington, D.C. 20001; (800) 624-6242; www.nap.edu.

*Permission granted to reproduce this brief in its entirety with no additions or alterations.
Permission for images/figures must be obtained from their original source.*