Life is sweet: the cell biology of glycoconjugates

Alex C. Broussard and Michael Boyce*
Department of Biochemistry and Program in Cell and Molecular Biology, Duke University School of Medicine, Durham, NC 27710

ABSTRACT Cells are dazzling in their diversity, both within and across organisms. And yet, throughout this variety runs at least one common thread: sugars. All cells on Earth, in all domains of life, are literally covered in glycans, a term referring to the carbohydrate portion of glycoproteins and glycolipids. In spite of (or, perhaps, because of) their tremendous structural and functional complexity, glycans have historically been underexplored compared with other areas of cell biology. Recently, however, advances in experimental systems and analytical methods have ushered in a renaissance in glycobiology, the study of the biosynthesis, structures, interactions, functions, and evolution of glycans. Today, glycobiology is poised to make major new contributions to cell biology and become more fully integrated into our understanding of cell and organismal physiology.

THE BIOLOGY OF GLYCANS: A VERY BRIEF INTRODUCTION

Life on our planet is sweet indeed, as bacteria, archaea, and eukaryotes alike decorate their cells with species-specific and cell type–specific glycans (Abu-Qarn et al., 2008; Moremen et al., 2012; Kaminski et al., 2013; Varki, 2017). The fundamental building blocks of glycans are monosaccharides, or single sugars that cannot be hydrolyzed to simpler forms. Usually starting from protein and lipid scaffolds, cells create glycans through glycosidic linkages—the attachment of one monosaccharide to another, often through the hydroxyl of the anomeric carbon (Figure 1A). Several features of glycans contribute to their remarkable diversity. In animals, most glycoconjugates are assembled from a relatively small palette of monosaccharides (fewer than 20). In the prokaryotic world, however, hundreds of different sugars are used, contributing to an enormous range of microbial glycan structures that we have only begun to characterize. Equally important, glycosidic bonds can exist in α or β stereochemistry, differing in the position of the glycosidic oxygen relative to the sugar ring (Figure 1A). Stereochemistry matters: two glycans with identical chemical formulas and atom connectivities but different stereochemistries can have vastly different biological properties, as the comparison between cellulose and glycogen famously illustrates (Figure 1A). In addition, glycan-binding proteins, such as lectins, can discriminate—sometimes with exquisite specificity—between otherwise identical glycans linked via α versus β glycosidic bonds. Beyond stereochemistry, glycosidic bonds can exist at any of several positions on each monosaccharide ring, allowing for staggering variety in connectivity and branching (Figure 1B). Finally, glycans can be decorated postsynthetically with many covalent modifications, analogous to posttranslational modifications of proteins. For instance, the site-specific sulfation of glycosaminoglycans (Figure 1C) creates patterns of negative charges, greatly expanding the diversity of this glycan class and encoding biological information that is read out by receptors on other cells (Soares da Costa et al., 2017). These aspects of glycan structure produce a combinatorial complexity that dwarfs that of proteins or nucleic acids—while three different amino acids or nucleotides can be arranged into six distinct linear trimers, three different hexoses (six-carbon sugars) could be assembled into 27,648 distinct hypothetical trisaccharides (Varki and Kornfeld, 2015).

In eukaryotes, most classes of glycoproteins and glycolipids are biosynthesized in the endoplasmic reticulum (ER) and Golgi, often through the sequential action of a series of glycosyltransferases and glycosidases localized to specific subdomains of these organelles (Moremen et al., 2012; Varki, 2017). This spatial arrangement forms something like an assembly line, building glycans as lipid and protein substrates transit the secretory pathway. Beyond the ER and Golgi, some major extracellular glycan classes are biosynthesized and extruded directly at the plasmid membrane, including...
chitin in fungi, cellulose in plants, and hyaluronan in animals (Figure 1D). Glycosylation is also abundant within animal and plant cells, where thousands of soluble nuclear and cytoplasmic proteins are dynamically modified on serine and threonine residues by O-linked β-N-acetylglucosamine (O-GlcNAc), a transient, single-sugar modification that regulates protein localization and function (Figure 1D) (Hart et al., 2011; Bond and Hanover, 2013; Yang and Qian, 2017).

Unlike DNA replication, RNA transcription, or protein translation, glycan biosynthesis is not directed by a preexisting template molecule. Instead, the final glycoforms of proteins and lipids depend on the intricate interplay among the glycan biosynthetic machinery, the available nucleotide sugars (small metabolites that serve as monosaccharide donors), and signals from the intracellular and extracellular environments. Owing in part to this biosynthetic complexity, a

FIGURE 1: Diversity of glycans. (A) Stereochemistry can greatly affect the physical and biological properties of glycans, as exemplified by the differences between glycogen (left) and cellulose (right), both polymers of 1,4-linked glucose. The glycosidic bonds are highlighted in yellow. (B) Glycan diversity is increased by branching. For example, in a mammalian N-glycan, a mannose residue (shown) might be linked to another mannose or an N-acetylglucosamine (GlcNAc) at its 1, 2, 3, 4, or 6 carbon position. (C) Postsynthetic modifications further diversify glycans. Here, a heparan sulfate disaccharide is pictured with arrows indicating known sites of possible sulfation. (D) Major classes of animal glycans are depicted. Most glycan types reside on the cell surface or in the extracellular space (here shown above the membrane). O-GlcNAc is intracellular, modifying nuclear and cytoplasmic proteins (below the membrane). (E) A given protein can exist in multiple glycoforms within a single cell or tissue, a phenomenon referred to as microheterogeneity. (F) A given protein can exist in distinct glycoforms across different tissues or organs, a phenomenon referred to as macroheterogeneity. (G) A single glycosylation site at Ser-49 is required for the assembly of the cytoskeletal protein vimentin into intermediate filaments in human cells. Interestingly, the same residue is dispensable for filament assembly in vitro, indicating an in vivo regulatory role for glycosylation at this site. Here, wild-type or unglycosylatable Ser-49-Ala mutant vimentin-GFP was expressed in vimentin−/− HeLa cells and imaged by fluorescence microscopy, as described in Tarbet et al. (2018).
given protein can exist in multiple glycoforms within a single cell or tissue, a phenomenon referred to as “microheterogeneity” (Figure 1E). Moreover, a single protein may be glycosylated differently in one tissue or organ versus another, producing “macroheterogeneity” as well (Figure 1F). Seemingly small changes in glycoforms can have profound consequences for protein function. To give one example: we recently showed that a single glycosylation site in the cytoskeletal protein vimentin is required for its assembly into intermediate filaments in human cells (Figure 1G) (Tarbet et al., 2018). The biological imperatives driving the wide assortment of glycoconjugates are poorly understood, but many organisms may use multiple glycoforms to generate greater functional diversity from a limited set of protein or lipid scaffolds, or to confound viruses and other parasites that might more easily exploit simple, predictable cell surface patterns for attachment and invasion.

Whatever the reason, this prodigious variety of structures supports a broad range of functions (Varki, 2017). For example, glycans provide structural integrity to cells and tissues, including the well-known examples of cellulose, chitin, and glycosaminoglycans like hyaluronan, mentioned earlier. Structural glycans often also perform passive signaling functions, such as sculpting gradients of morphogens by modulating their diffusion rate through the extracellular space or even by relieving their autoinhibition upon contact with receptor glycans (Metzler et al., 1994; Inatani et al., 2003; Sarris et al., 2012; Kiermaier et al., 2016). Glycosylation heavily influences the protein quality-control systems of the secretory pathway as well, where dedicated machinery in the ER lumen interacts directly with glycans on newly translated polypeptides, ensuring that only properly folded proteins are secreted (Breitling and Aebi, 2013). On the cell surface and beyond, many glycans also mediate essential receptor–ligand interactions. In one canonical example, secreted human chorionic gonadotropin changes its glycoforms over time during pregnancy, resulting in evolving signaling effects by the same polypeptide hormone (Fournier, 2016). Other well-documented functions of glycans include the regulation of endocytosis and protein trafficking, membrane organization and signaling, nutrient storage, immune recognition and evasion, and cell–cell interactions during fertilization. Finally, glycans play a key role in human health and disease, both because major illnesses are characterized by aberrant glycosylation (e.g., cancer, neurodegeneration, diabetes, and congenital disorders) and because a number of critical therapeutics are glycosylated (e.g., heparin, aminoglycoside antibiotics, erythropoietin, monoclonal antibodies, and many vaccines) (Freeze and Aebi, 2005; Ohtsubo and Marth, 2006; Hart et al., 2011; Bond and Hanover, 2013; Hennet and Cabralzar, 2015; Yilmaz, 2017).

Given the ubiquity and significance of glycans, we will need deep insights from glycobiology to develop a truly comprehensive and integrated understanding of the cell. However, despite this importance, glycobiology has been understudied in the last several decades compared with other aspects of cell biology. One reason for this lag is that the chemical complexity and analytical challenges of glycans have presented formidable barriers to research (National Research Council, 2012; Agre et al., 2016; Varki, 2017). Recently, however, advances in carbohydrate chemistry, chemical biology, mass spectrometry, structural biology, and biological model systems have empowered new avenues of exploration in glycobiology, setting the stage for great progress in the coming years. To illustrate a few of these recent trends, we provide three very brief and highly selective snapshots of exciting work on the roles of glycans outside, inside, and beyond the cell. For a more thorough introduction to these and other topics in the field, we refer the reader to the excellent Essentials of Glycobiology, a graduate-level textbook available free online (www.ncbi.nlm.nih.gov/books/NBK310274), and to outstanding recent review articles (Moremen et al., 2012; Varki and Kornfeld, 2015; Varki, 2017).

Outside: cell surface receptor–ligand interactions

In animals, many cell surface proteins require specific glycans to perform their jobs. While this has long been appreciated, current research has revealed beautiful new aspects of this glycan function. For example, Notch is festooned with unusual O-linked glycans that are essential for its ligand binding. Notch glycosylation has been studied for years, but improved imaging methods and developmental model systems have yielded recent fresh insights in the biosynthesis and function of Notch glycans during vertebrate embryogenesis (Boskovski et al., 2013; Sawaguchi et al., 2017). Equally interesting, groundbreaking structural studies revealed that individual sugars within Notch glycans make direct contacts with ligands, triggering conformational changes and a “catch bond” behavior that bolsters the receptor–ligand interaction across the range of tensions that Notch requires for its activation (Luca et al., 2015, 2017). Mechanistic knowledge of Notch signaling, coupled with novel chemical approaches, has empowered new efforts to manipulate Notch signaling with sugar analogues (Schneider et al., 2018), which may hold promise for future cancer therapies. In another current example, a flurry of discoveries has shed new light on the unusual O-glycans of dystroglycan, a protein that forms a transmembrane link between the extracellular matrix and the cytoskeleton and is dysregulated in muscular dystrophy (Goddeers et al., 2013; Yoshida-Moriguchi et al., 2013; Briggs et al., 2016; Kanagawa et al., 2016; Praissman et al., 2016; Zhu et al., 2016). Thanks in part to new semisynthetic and structural approaches combined with established cellular and animal models, the composition and effects of normal and aberrant dystroglycan glycans have come into sharper focus. Finally, recent work has elucidated unexpected bulk functional properties of the glycoalyx, or the shell of glycans that surrounds all cells. Sophisticated imaging and chemical methods have revealed that a tall and dense glycoalyx promotes integrin activation, not through direct action on integrin glycans themselves, but rather by trapping integrins near each other when they are squeezed between larger glycoproteins, such as mucins (Paszek et al., 2014; Woods et al., 2017; Barnes et al., 2018). These bulk effects may promote cell survival and proliferation in tumor cells, which often exhibit abnormalities in their glycoalyx. It will be interesting to learn in future studies whether bulk glycoalyx properties are harnessed in normal tissue to tune the functions of integrins and other cell surface receptors.

Inside: intracellular signaling

As noted earlier, large numbers of intracellular proteins are reversibly modified by O-GlcNAc in animals and plants (Figure 1D). O-GlcNAc was discovered nearly 35 years ago, but many aspects of its biology remain incompletely understood, including how it impacts signaling pathways and downstream phenotypes. O-GlcNAc is added to substrates by the glycosyltransferase OGT, using UDP-GlcNAc as its nucleotide-sugar donor. A prevailing notion is that O-GlcNAc often serves a nutrient-sensing function, particularly because UDP-GlcNAc itself is biosynthesized from several essential metabolites, including glucose, glutamine, ATP, uridine, and acetyl-coenzyme A (Hart et al., 2011; Bond and Hanover, 2013; Yang and Qian, 2017). Indeed, global O-GlcNAc levels have long been known to fluctuate in response to nutrient changes, but the specific glycoprotein substrates and functional effects involved are largely unclear. Recent work has revealed important new aspects of this...
signaling, as O-GlcNAc has been linked to such metabolically important pathways and processes as cardiac calcium signaling through calcium/calmodulin-dependent kinase II (Erickson et al., 2013), circadian entrainment (Li et al., 2013), mitochondrial motility in neurons (Wang and Schwarz, 2009; Pekkumaz et al., 2014), epigenetic transcriptional control of metabolic genes through the TET family of 5-methylcytosine oxidases (Chen et al., 2013; Deplus et al., 2013; Vella et al., 2013; Hrit et al., 2018), and the Hippo-YAP organ size control pathway (Peng et al., 2017). Interestingly, other reports have suggested nutrient-sensing functions for O-GlcNAc at the organ level as well, as mammalian OGT is required in distinct subpopulations of hypothalamic neurons to govern peripheral white fat browning, satiety, and feeding (Ruan et al., 2014; Lagerlof et al., 2016). Importantly, however, O-GlcNAc influences many processes beyond nutrient sensing. For example, our lab has recently described roles for O-GlcNAc in the cytoskeleton (Tarbet et al., 2018), protein trafficking (Cox et al., 2018a,b), redox stress signaling (Chen et al., 2017), and mediation of protein–protein interactions (Toleman et al., 2018). Going forward, improved quantitative methods for pinpointing dynamic, site-specific occupancy on key substrates will be needed to fully dissect the role of O-GlcNAc in cell and tissue physiology.

Beyond: interspecies relations
Because they coat the surface of all cells, glycans are often frontline mediators of interactions between species. Examples of this phenomenon abound, but perhaps the best studied is the human gut microbiome, which is influenced by, and influences, the host’s metabolism and glycome. Contemporary work has highlighted interesting new aspects of this complex relationship. For instance, deletion of a single glycosyltransferase gene in mice resulted in significant remodeling of the microbiome and its response to dietary carbohydrates, with particularly pronounced changes in the gene expression program of Bacteroides thetaiotaomicron, a common gut bacterium (Kashyap et al., 2013). Other recent studies have documented striking relationships among dietary carbohydrates and the intestinal flora, including the effects of the probiotic arabinoxylan (a hemicellulose glycan) on Bacteroides species (Wu et al., 2015) and a microbiome-dependent augmentation of host anabolic pathways, lean weight gain, and bone morphogenesis in mice and piglets fed oligosaccharides containing siaic acid (a nine-carbon monosaccharide), which are normal components of human breast milk (Charbonneau et al., 2016). These and other studies indicate that it may be possible to improve human health through rational manipulation of the gut microbiome using probiotic glycans. Beyond Homo sapiens, glycans are the common language among other kingdoms as well. For instance, in legumes, a receptor kinase that it may be possible to improve human health through rational manipulation of the gut microbiome using probiotic glycans. Beyond Homo sapiens, glycans are the common language among other kingdoms as well. For instance, in legumes, a receptor kinase

CONCLUSIONS AND FUTURE DIRECTIONS
As these and countless other examples illustrate, glycobiology is an integral—and fascinating—aspect of cell biology. Though not emphasized here, glycobiology also has profound implications for medicine, bioenergy, and materials science, underscoring its exceptionally broad significance (National Research Council, 2012). Thanks in part to recent advances in chemical, analytical, and structural techniques, we are now poised to address some of the exciting challenges remaining in the field. For example, how is glycan diversity (e.g., micro- and macroheterogeneity) regulated, and what is its purpose? How is the information encoded in complex glycans “read” by proteins and other biomolecules? What is the structural basis of glycan function? Do the bulk properties of glycans (e.g., the glycoalkali or heavily O-GlcNacylated nuclear pore complex) serve specific biological functions, including the regulation of phase separations or similar biological phenomena? How and when can glycosylation be rationally manipulated for therapeutic benefit in human disease? Addressing these questions will require not only technological advances, but also generations of young scientists who are well-trained in glycobiology as a fundamental aspect of cell biology (Agre et al., 2016). Given the tremendous scientific discoveries in the field lying on and beyond the horizon, life for these future glycobiologists will be sweet indeed.

ACKNOWLEDGMENTS
Research in our lab has been supported by the National Institutes of Health (R01GM118847 and R01GM117473) and a Scholar Award from the Rita Allen Foundation to M.B.

REFERENCES
Abu-Qarn M, Eichler J, Sharon N (2008). Not just for Eukarya anymore: protein glycosylation in bacteria and Archaea. Curr Opin Struct Biol 18, 544–550.
Agre P, Bertozi C, Bissell M, Campbell KP, Cummings RD, Desai UR, Estes M, Flotte T, Fogelman G, Gage F, et al. (2016). Training the next generation of biomedical investigators in glycosciences. J Clin Invest 126, 405–408.
Barnes JM, Kaushik S, Bainer RO, Sa JK, Przybyla L, Lee M, Lee HW, Tung JC, et al. (2018). A tension-mediated glycoalkalyx-integran feedback loop promotes mesenchymal–glialblasta. Nat Cell Biol 20, 1203–1214.
Bond MR, Hanover JA (2013). O-GlcNAc cycling: a link between metabolism and chronic disease. Annu Rev Nutr 33, 205–229.
Boskovski MT, Yuan S, Pedersen NB, Goth CK, Makova S, Clausen H, Brueckner M, Khokha MK (2013). The heterotaxy gene GALNT11 glycosylates Notch to orchestrate cilia type and lateriality. Nature 504, 456–459.
Breitling J, Aebl M (2013). N-linked protein glycosylation in the endophasmic reticulum. Cold Spring Harb Perspect Biol 5, a013359.
Briggs DC, Yoshida-Moriguchi T, Zheng T, Venzke D, Anderson ME, Strazzulli A, Moracci M, Yu L, Hohenester E, Campbell KP (2016). Structural basis of laminin binding to the LARGE glycans on dystroglycan. Nat Chem Biol 12, 810–814.
Charbonneau MR, O’Donnell D, Blanton LV, Totten SM, Davis JC, Barratt MJ, Cheng J, Guruge J, Talcott M, Bain JR, et al. (2016). Sialylated milk oligosaccharides promote microbiota-dependent growth in models of infant undernutrition. Cell 164, 859–871.
Chen PH, Smith TJ, Wu J, Siesser PF, Bissert BJ, Khan F, Hogue M, Soderblom E, Tang F, Marks JR, et al. (2017). Glycosylation of KEAP1 links nutrient sensing to redox stress signaling. EMBO J 36, 2233–2250.
Chen Q, Chen Y, BIAN F, Fujiki R, Yu X (2013). TET2 promotes histone O-GlcNAcylation during gene transcription. Nature 493, 561–564.
Cox NJ, Luo PM, Smith TJ, Bissert BJ, Soderblom EJ, Boyce M (2013). A tension-mediated glycoalkalyx-integran feedback loop promotes mesenchymal–glialblasta. Nat Cell Biol 20, 1203–1214.
Cox NJ, Unlu G, Bissert BJ, Meister TR, Condon BM, Luo PM, Smith TJ, Hanna M, Chhetri A, Soderblom EJ, et al. (2018b). Dynamic glycosylation governs the vertebrate COPII protein trafficking pathway. Biochemistry 57, 91–107.
Deplus R, Delatte B, Schwinn MK, Defrance M, Mendez J, Murphy N, Dawson MA, Vollmar M, Putmans P, Calonne E, et al. (2013). TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. EMBO J 32, 645–655.

Dinh C, Farnholt T, Hirose S, Zhuchenko O, Kupsa A (2018). Lectins modulate the microbiota of social amoebae. Science 361, 402–406.

Erckson JR, Pereira L, Wang L, Han G, Ferguson A, Dao K, Copeland RJ, Despa F, Hart GW, Ripplinger CM, Bers DM (2013). Diabetic hyperglycaemia activates CaMII and arrhythmias by O-linked glycosylation. Nature 502, 372–376.

Fournier T (2016). Human chondronic gonadotropin: different glycoforms and biological activity depending on its source of production. Ann Endocrinol (Paris) 77, 75–85.

Freyre HH, Aebi M (2005). Altered glycans: the molecular basis of congenital disorders of glycosylation. Curr Opin Struct Biol 15, 490–498.

Goddeeris MM, Wub B, Venzke D, Yoshida-Moriguchi T, Saito F, Matsuura K, Moore SA, Campbell KP (2013). LARGE glycans on dystroglycan function as a tunable matrix scaffold to prevent dystrophy. Nature 503, 136–140.

Hart GW, Slawson C, Ramirez-Correa G, Delatte B, Schwinn MK, Defrance M, Mendez J, Murphy N, Dawson MA, Vollmar M, Putmans P, Calonne E, et al. (2013). TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. EMBO J 32, 645–655.

Inatani M, Irie F, Plump AS, Tessier-Lavigne M, Yamaguchi Y (2003). Mamma Hrit J, Goodrich L, Li C, Wang BA, Nie J, Cui X, Martin EA, Simental E, Hennet T, Cabalzar J (2015). Congenital disorders of glycosylation: a con}

Kiermaier E, Moussion C, Veldkamp CT, Gerardy-Schahn R, de Vries I, Kanagawa M, Kobayashi K, Tajiri M, Manya H, Kuga A, Yamaguchi Y, Kaminski L, Lurie-Weinberger MN, Allers T, Gophna U, Eichler J (2013). Inatani M, Irie F, Plump AS, Tessier-Lavigne M, Yamaguchi Y (2003). Mamma Hrit J, Goodrich L, Li C, Wang BA, Nie J, Cui X, Martin EA, Simental E, Hennet T, Cabalzar J (2015). Congenital disorders of glycosylation: a con