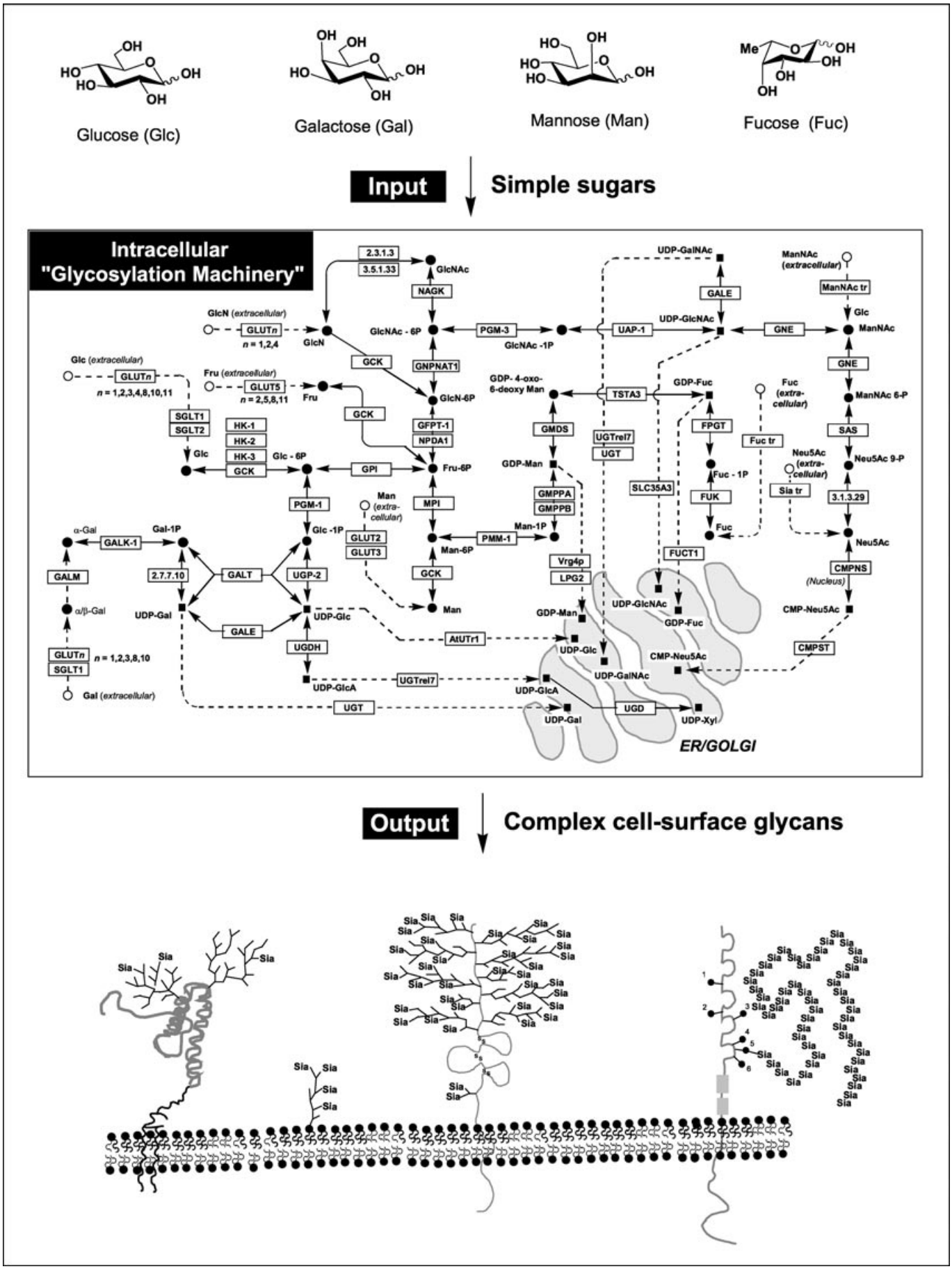


  DOI: 10.1002/cbic.200400143 ChemBioChem 2004, 5, 1334 – 1347

The Systems Biology of Glycosylation

Michael P. Murrell, Kevin J. Yarema,* and Andre Levchenko*[a]

Glycosylation can have a profound influence on the function of a variety of eukaryotic cells. In particular, it can affect signal transduction and cell–cell communication properties and thus shape critical cell decisions, including the regulation of differentiation and apoptosis. Regulation of glycosylation has multiple layers of complexity, both structural and functional, which make its experimental and theoretical analysis difficult to perform and interpret. Novel research methodologies provided by systems biology can help to address many outstanding issues and integrate glycosylation with other metabolic and cell regulation processes. Here we review the toolbox available for biochemical systems analysis of glycosylation.

Introduction

The surface of virtually every living cell is decorated with a layer of complex carbohydrates. In mammals, these sugars play important structural and signaling roles that are indispensable for the development and maintenance of the multicellular organism. The complexity of the glycosylation pathways that manufacture and constantly remodel surface sugars is truly astounding; this metabolic system utilizes many hundreds of components that connect metabolism, biosynthesis, and cell-signaling and -regulation events. We now increasingly accept that reductionist approaches, where components of complex systems such as glycosylation are studied alone, are often insufficient for understanding the inherent complexity in biological systems. More integrative methods have been termed *systems analysis* and emphasize global properties of robustness, regulation, and control. Glycosylation provides a perfect opportunity to employ a systems approach on an important biological system whose properties resist full definition when study remains limited to a reductionist approach.

The advent of high-throughput technology and the generation of annotated genome sequences have facilitated great strides in the reconstruction and quantitative treatment of biochemical metabolic networks of systems as complex as entire cells.[1–3] A tight integration of modeling and experimental studies enables the analysis of cells and organisms at the genome level by complementing the narrowly focused characterization of individual components. Metabolic systems, such as the glycosylation pathways responsible for the biosynthesis of cell-surface carbohydrates, are ideal for analysis by a systems approach as many computational and modeling tools have been developed for their study. These developments include the availability of the corresponding annotated genome sequences specifying glycosylation enzymes across different organisms, an increasing knowledge of the various regulatory mechanisms that these metabolic enzymes are subject to, and the characterization of the highly diverse metabolic products. The enzymes involved in glycosylation can be "promiscuous" in their specificity or redundant, thereby complicating the analysis of metabolic synthesis and conversion.[4] In terms of regulation, the activity of glycosylation enzymes can be controlled by gene activation, repression, and direct protein–protein interactions that are often "fine-tuned" by the spatial localization of these proteins. Furthermore, glycosylation is amenable to experimental validation, as defects in enzyme regulation often result in observable phenotypes. Thus, we propose that, by applying an integrative conceptual framework in studying global glycosylation properties, advantage can be taken of the considerable expertise that has been devoted to the systems study of metabolism. At the same time, however, glycosylation presents new challenges to the systems biologist/computational modeler; for instance, the incredible diversity of the metabolic products of the glycosylation pathways, namely, the collective complement of cell-surface carbohydrates, complicates analysis. Furthermore, glycosylation often lies at the interface between metabolism and cell signaling, thereby making analysis even more demanding.

Complexity in Glycosylation

Glycosylation is a complex process that requires a major commitment of cellular resources; it employs an estimated 2–3% of the genes in humans and many high-energy intermediates. The first step in glycosylation is the import of dietary sugars, such as glucose, into a cell (Scheme 1, Step A). This is followed by a series of phosphorylation, epimerization, and acetylation reactions that diversify these sugars and convert them into high-energy nucleotide sugar donors (Scheme 1, Step B, and Scheme 2). These compounds serve as the "building blocks" for the assembly of complex carbohydrates in the endoplasmic reticulum (ER) and Golgi apparatus (Scheme 1, Step C, and Scheme 3). The newly synthesized carbohydrates, which are generally attached to proteins or lipids, are then transported to the cell surface where they contribute to the interaction be-

[a] *M. P. Murrell, Prof. K. J. Yarema, Prof. A. Levchenko*
Department of Biomedical Engineering, Johns Hopkins University
3400 N. Charles Street, Baltimore, MD 21218 (USA)
Fax: (+1) 410-516-8152 (Yarema)
Fax: (+1) 410-516-6240 (Levchenko)
E-mail: kjyarema@bme.jhu.edu
alev@bme.jhu.edu

 DOI: 10.1002/cbic.200400143

tween a cell and its environment by playing important structural and signaling roles.[5,6]

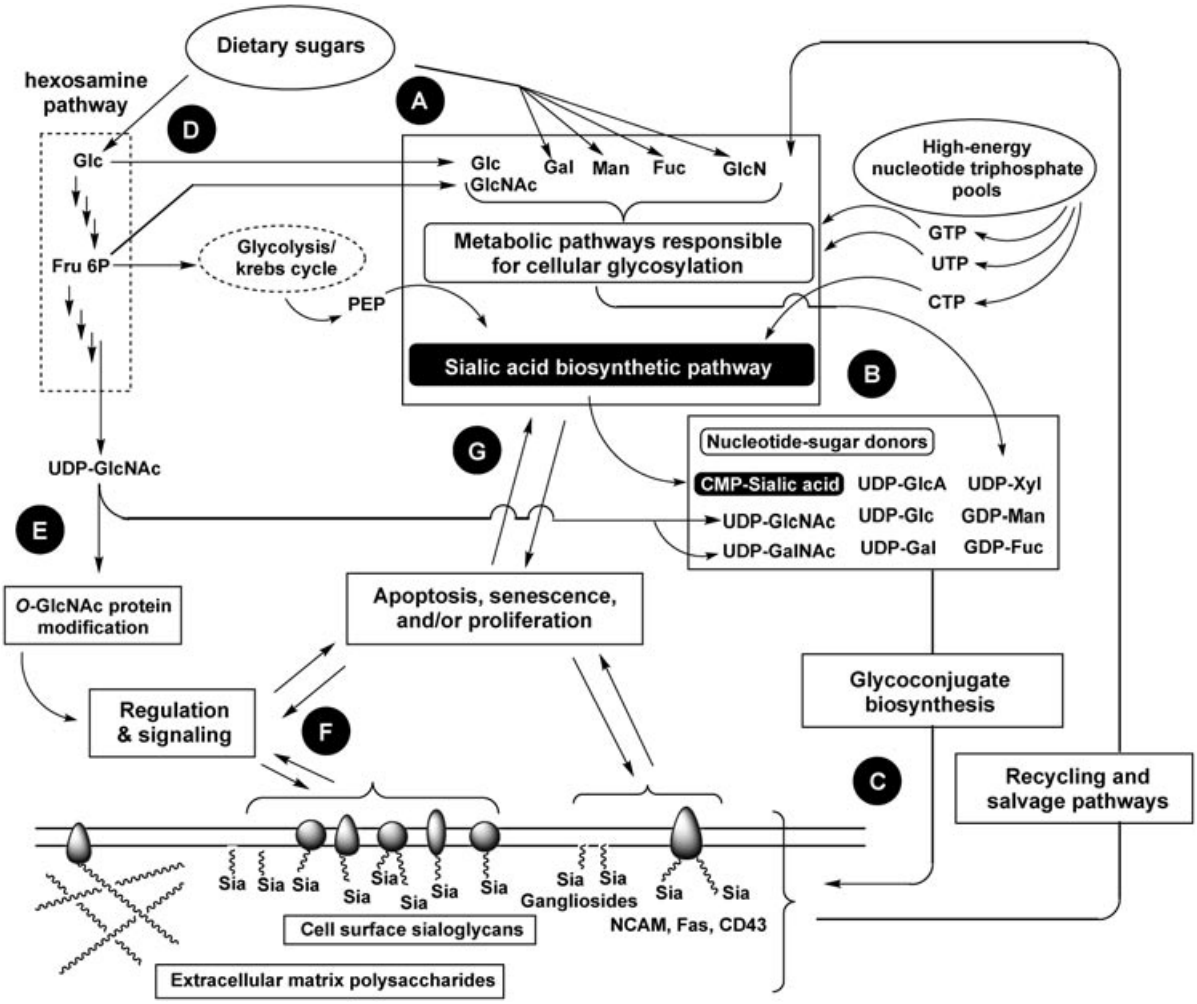
A signature of glycosylation reactions is the immense multiplicity of chemical structures they produce. In combination with the cellular machinery involved in the synthesis and modification of these complex carbohydrates, glycosylation processes are very rich in products, even on the metabolic-system scale. Carbohydrates exhibit greater structural diversity than other cellular macromolecules, such as nucleic acids and proteins, because each sugar residue has three or four different hydroxy groups that can provide a point of attachment to another monosaccharide unit, thereby allowing oligosaccharide structures to form branches, as opposed to other macromolecules which typically form linear chains. As a result, in theory, the nine common monosaccharides found in humans could be assembled into more than 15 million possible tetrasaccharides, all of which would be considered relatively simple glycans.[7] The complexity of glycosylated molecules is increased still further because many glycoproteins have multiple (sometimes dozens) sites for the attachment of the sugar and each site can host many different oligosaccharides. The prion protein (Prp), for example, can be decorated with over 50 different sugar structures at either of two different attachment sites.[8] Interestingly, glycosylation variants of the diseased form of the prion protein (PrPsc) influence the development of spongiform encephalopathies,[9] thereby providing just one of many examples where diversity in cell-surface carbohydrate architecture is coupled with the development of disease-state phenotypes. Uncovering the underlying connections between structure and function is a daunting task considering that the heterogeneity of the cell-surface carbohydrates makes characterizing even a single glycoprotein a formidable challenge.

The complexity of carbohydrate structure is matched by the intricacy of the intracellular mechanisms that *produce* glycans.

Michael Murrell was born in Milwaukee, WI, USA, in 1981. He received his BS degree in biomedical engineering from the Johns Hopkins University, where his primary work focused on glycosylation and carbohydrate engineering in the Laboratory of Cell Engineering under K.J.Y. As a visiting researcher at the laboratory of Professor Lodish of the Whitehead Institute for Biomedical Research, he has also contributed to studies into the effects of the ACRP30 protein on fatty acid oxidation. Under the sponsorship of a Howard Hughes Fellowship, Michael Murrell investigated methods of network reconstruction of biological systems with A.L. in the Laboratory of Cell Signaling and Cell–Cell communication. He also worked in the X-Ray Global Components division of GE Medical Systems. He is now a graduate student in the Division of Biological Engineering at the Massachusetts Institute of Technology, in Cambridge, MA.

Kevin Yarema was born in Williams Lake, BC, Canada, in 1965. After completing an undergraduate degree in bioengineering at Walla Walla College, he went on to receive a PhD degree in biological chemistry for research done under the direction of Professor Essigmann at the Massachusetts Institute of Technology. He then completed a postdoctoral fellowship in the laboratory of Professor Bertozzi at the Berkeley campus of the University of California where he specialized in chemical biology approaches to manipulate glycosylation in living cells. He is now an assistant professor in the Biomedical Engineering department of the Johns Hopkins University, where he is pursuing systems biology approaches for determining the role of glycosylation in health and disease and developing metabolic-substrate-based approaches for cell-surface carbohydrate engineering applications.

Andre Levchenko was born in Kishinev, USSR, in 1970 and grew up in Siberia. He received his BS and MS degrees in biological physics from the Moscow Institute for Physics and Technology in 1992 and his PhD degree in mechanical and biomedical engineering from Columbia University in 1998. While at the graduate school, he also worked at the Memorial Sloan–Kettering Cancer Center on drug resistance in cancer cells. In 1998 he started his postdoctoral studies at the California Institute of Technology (Caltech) in the Divisions of Electrical Engineering and Biology, working with Professor Sternberg and Professor Bruck. At Caltech, Andre Levchenko was awarded the Burroughs–Wellcome Fellowship award for his studies of signaling pathways in yeast. In 2001 he accepted a position as Assistant Professor of Biomedical Engineering at the Whitaker Institute for Biomedical Engineering of the Johns Hopkins University. He has been on the faculty at Johns Hopkins ever since, working on experimental and theoretical research into biological signal transduction and cell–cell communication. Andre Levchenko is the author of multiple publications focusing on the mechanisms of cell regulation. He is also on the Editorial Board of IEE Systems Biology. He lives with his family in Columbia, MD, USA. For more information on Levchenko's group, visit www.bme.jhu.edu/labs/levchenko.

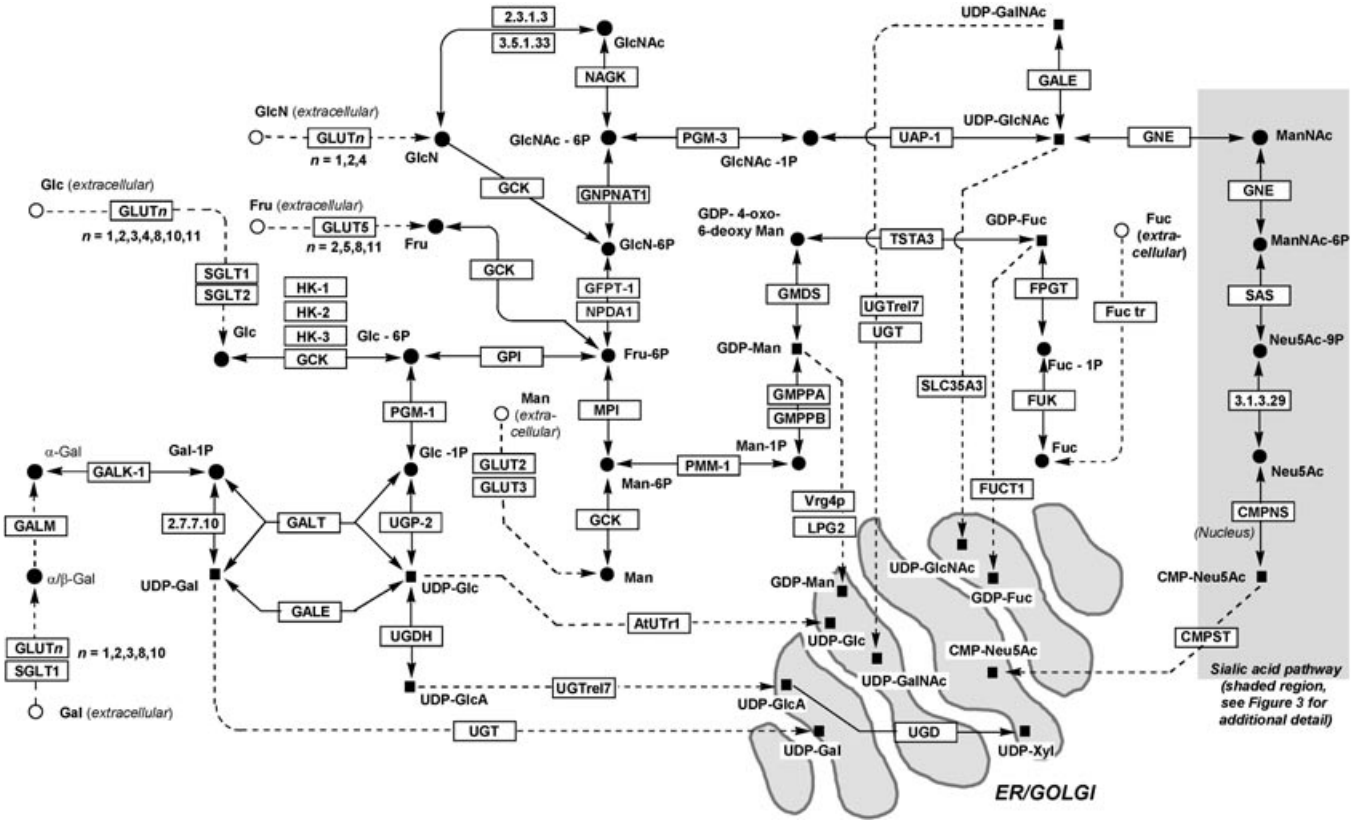


Scheme 1. *An overview of mammalian glycosylation (Steps A–C) and a sampling of the metabolic and signaling connections formed by the "glycosylation machinery" of a cell (Steps D–G). Please see the text for a detailed discussion and the tables for the full names of the monosaccharide intermediates, transporters, and enzymes involved in these processes (as well as those in Schemes 2–4.*

Recent evidence contradicts the previously held central dogma of glycobiology, that is, the "one-enzyme, one-linkage" rule that was thought to govern glycosyltransferase activity.[4] More specifically, as progress continues in fully elucidating the constituent components of the "glycosylation machinery", more than 250 enzymes devoted to oligosaccharide biosynthesis (including glycosyltransferases) have been described and it has been discovered that multiple enzymes are capable of catalyzing exactly the same reaction. In other cases, a single enzyme can catalyze the synthesis of more than one linkage. These findings complicate attempts to connect a carbohydrate structure with a unique sequence of enzyme actions. Further complicating the biosynthetic process is the fact that, unlike proteins whose amino acid sequence is specified by a DNA template, carbohydrate synthesis is not template based. Instead, oligosaccharide biosynthesis results from a series of step-by-step enzymatic conversions occurring in the ER and Golgi apparatus and the final carbohydrate structures produced are thought to depend largely on factors such as the transport of nucleotide sugar donor "building blocks" into these organelles and the spatial orientation of the biosynthetic enzymes.

Glycosylation Is Tied to Other Cellular Processes

The complex glycosylation pathways of a cell do not exist in isolation; instead they are intimately connected and intertwined with other critical metabolic and regulatory networks of a cell (Scheme 1). For instance, the major dietary sugar glucose is not used exclusively for biosynthesis of larger carbohydrates, but rather the majority of it enters the "hexosamine pathway" (Scheme 1, Step D)[10] where it is converted into Fru-6P, which is in turn used for energy production through glycolysis. Interestingly, Fru-6P can also be converted into GlcNAc, which is used as a building block for glycosylation. Also, phosphoenol pyruvate (PEP), a product obtained from Fru-6P through glycolysis and the Krebs cycle, is a required cosubstrate in the sialic acid biosynthetic pathway. Diversion of metabolic flux from the hexosamine pathway for these purposes is expected to alter the formation of UDP-GlcNAc (itself used for glycosylation), a compound whose concentration largely determines the extent of *O*-GlcNAc protein modification within a cell (Scheme 1, Step E). *O*-GlcNAc protein modification is a



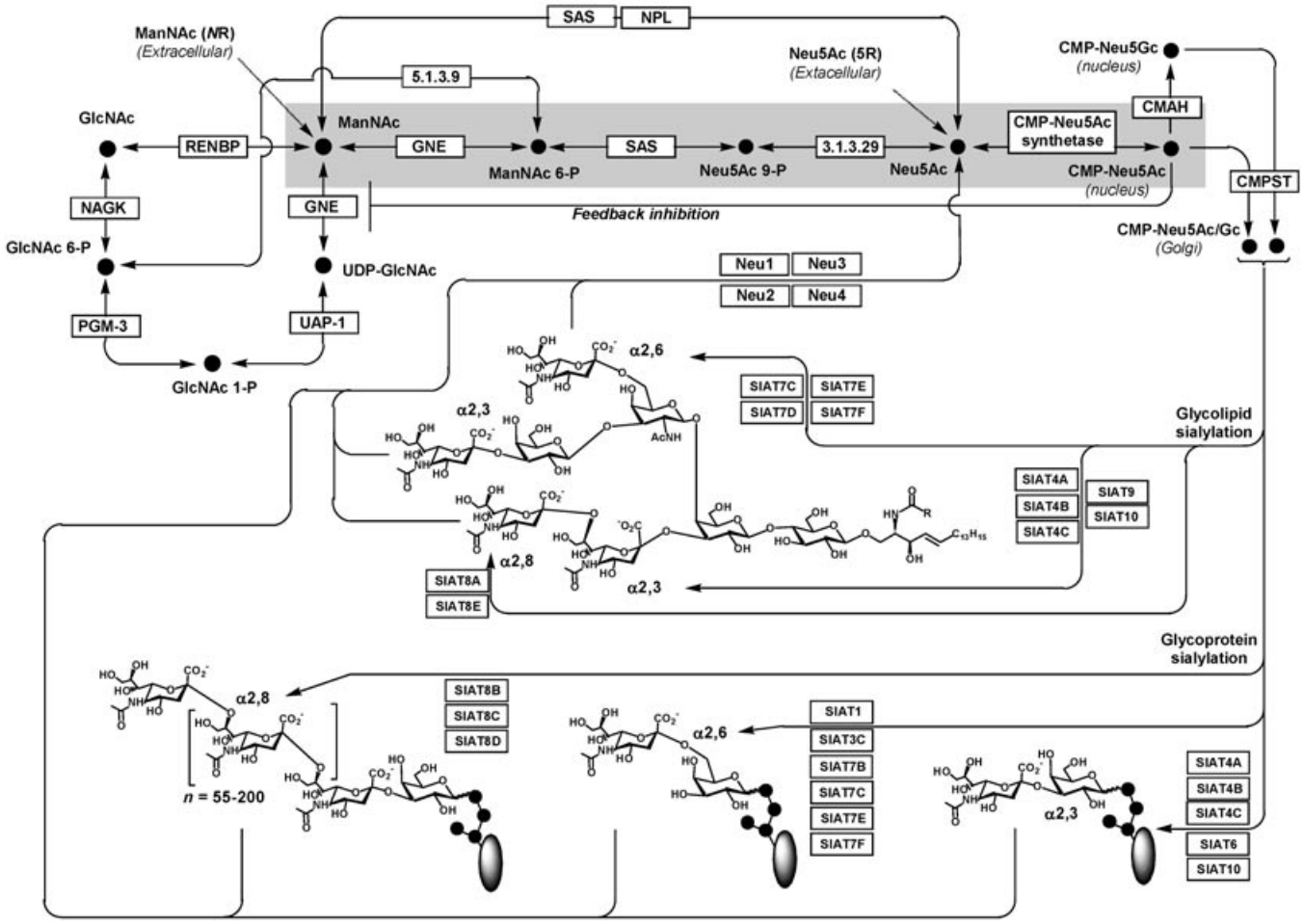
Scheme 2. *Monosaccharide transport and metabolic processing reactions. Extracellular sugars are transported into cells (dashed lines), converted into additional varieties of monosaccharides by a series of phosphorylation, epimerization, and acetylation reactions, then condensed with CTP, GTP, or UTP to form nucleotide sugar donors (solid lines). The nucleotide sugar donors function as the "building blocks" for the production of complex carbohydrates; their biosynthesis generally occurs after they are transported into the ER or Golgi compartments (dashed lines).*

posttranslational addition of a single GlcNAc to hundreds of different cytosolic proteins; this modification influences the activity of the host protein and is emerging as a major intracellular regulatory/signaling mechanism that affects many cellular behaviors (Scheme 1, Step F).[10,11]

Apoptosis provides an illustrative example of how *intracellular* glycosylation processes and *cell-surface* carbohydrates engage in a complex interplay to modulate important biological events. First, in intracellular events, the above-mentioned *O*-GlcNAc protein modification reaction is connected to the regulation of apoptosis. Overall levels of *O*-GlcNAc proteins are up-regulated in cells undergoing stress; this increase functions as a prosurvival mechanism for the affected cell. Upon exceeding a certain stress threshold, however, *O*-GlcNAc protein levels rapidly decline and the apoptotic program is allowed to proceed.[12] Another intracellular connection between apoptosis and glycosylation is emerging wherein the rate of metabolic flux through the sialic acid pathway appears to play a role in modulating apoptosis (Scheme 1, Step G).[13] The surface-displayed products of the intracellular pathways are also involved in apoptosis. For example, a *reduction* in sialic acid presented on specific cell-surface markers, such as Fas, sensitizes a cell to apoptosis;[14,15] similarly, overall cellular levels of sialic acid decline during apoptosis. By contrast, an *increase* in sialic acid on another cell-surface protein, CD43, occurs during early onset of apoptosis.[16] The challenge is to define how different individual forms of surface sialic acid are simultaneously up- (CD43) or down- (Fas) regulated during apoptotic cell death when the only known regulator of sialic acid biosynthesis, the GNE gene,[17] lies far upstream of these two metabolic products and should affect the production of each sialoglycoprotein equally, according to current understanding.

Considerations in Modeling Glycosylation

The process outlined above where cellular energy utilization, *O*-GlcNAc protein modification, and surface carbohydrates engage in a complex interplay to govern a crucial biological response such as apoptosis provides only one of many possible examples that could be used to illustrate how glycosylation involves a level of complexity that can ultimately only be understood by employing an integrated systems approach. To fully understand how this complex system works in healthy cells, as well as to discover how glycosylation-modulated behaviors contribute to the progression of disease states, we propose that it is necessary to identify the systems properties of glycosylation, to determine the mechanisms of control and regulation, and to elucidate the critical elements that determine the fragility and robustness of the underlying biochemical processes to various perturbations. Applying a systems-level analysis



Scheme 3. *Metabolic details of sialoglycoconjugate biosynthesis. The basic "sialic acid biosynthetic pathway" (as shaded in Scheme 2) has many additional potential and realized metabolic connections that need to be considered when building a rigorous model. For instance, both exogenously supplied ManNAc and Neu5Ac can gain access to the pathway although no specific transporters have been identified. Also, the renin-binding protein (RENBP), although not expressed in all cells, provides an alternative means of ManNAc being supplied into the pathway (or flux being diverted away from the pathway). Upon production of the CMP-Neu5Ac "building blocks", these compounds contribute a sialic acid residue during the assembly of sialoglycoconjugates. The final biosynthetic step is catalyzed by a suite of sialyltransferases (SIAT*n*x)that work in parallel to provide α2,3-, α2,6-, or α2,8-linked sialoglycoconjugates, as discussed in more detail in the text.*

to glycosylation, however, presents unique challenges to the researcher and efforts to model this system have lagged behind other well-studied networks such as the MAP kinase[18] or NF-kappaB[19] pathways. Briefly stated, a systems study of glycosylation requires the simultaneous incorporation of three lines of investigation, up to now typically addressed separately but with limited success, into one coherent modeling effort. These aspects are a) integrating the catalytic and spatial aspects of carbohydrate biosynthesis, b) studying the role of glycosylation at the intersection of metabolism, cell adhesion, and signaling, and c) accounting for how metabolic products can be "fine-tuned" in myriad ways. Each of these factors will be discussed briefly, with the intention not to provide a definitive guide to how glycosylation should (or must) be modeled but rather to motivate the investigation of the underlying systems properties in glycosylation and spark interest in the development of new approaches to its study.

Carbohydrate Biosynthesis Involves Both Catalytic and Spatial Considerations

As mentioned above, the application of a computational approach to study the systems properties of glycosylation involves unique and novel challenges. Consequently, it is beneficial to briefly consider the basic subsystems and elements that need to be included in initial efforts to model glycosylation. In the past, attempts to apply computational approaches to glycosylation were performed with the assumption that consideration of the metabolic processes that occur in the ER and Golgi, where the glycans are actually assembled and diversified, was sufficient. For example, Monica and co-workers assumed that CMP-Neu5Ac, the building block for sialic acid biosynthesis in human cells, was maintained at saturating concentrations in the Golgi and competition between the various sialyltransferases that produce α2,3-linked, α2,6-linked, or α2,8-linked sialosides was the primary factor to consider in model production (Scheme 3).[20] It is now known that this approach

was over-simplified and at least two additional factors must be considered: the availability of nucleotide sugar donors in the ER or Golgi and the expression and spatial localization of the biosynthetic enzymes. Each issue will be discussed briefly in the following paragraphs.

In-depth analysis of the mechanisms responsible for glycosylation indicates that the supply of nucleotide sugar donors into the Golgi or ER can play a defining role in determining the finer details of cell-surface architecture. In one study, certain galactose-containing products are made normally (heparin sulfate, chondroitin 4-sulfate) while others are severely reduced (glycoproteins, glycolipids, keratin sulfate) when UDP-Gal is limited.[21] An altered supply of nucleotide sugar donors in the ER or Golgi can result from two different metabolic abnormalities. The first type of abnormality is illustrated by the congenital disease leukocyte adhesion deficiency II (LADII), which occurs when defective GDP-Fuc transport into the Golgi results in a decrease in the availability of this nucleotide sugar donor to the glycosyltransferases that collectively assemble oligosaccharides displayed on the cell surface.[22–26] This transport defect results in decreased *N*-linked glycosylation but little change in *O*-linked glycosylation.[27] Thus, while the bulk levels of fucosylation have not changed significantly, the altered *pattern* of glycosylation, where certain crucial epitopes such as blood group antigens or selectin ligands are defective, leads to clinical manifestation of the disease. The second type of metabolic abnormality is the up- or down-regulation of nucleotide sugar production in the cytosol. The sialic acid biosynthetic pathway (Scheme 3), which produces CMP-Neu5Ac for transport into the Golgi, provides an example of both regulation abnormalities. When the key regulatory enzyme, GNE, suffers one type of single amino acid mutation, feedback inhibition is lost and sialic acid production increases by several hundred times, thereby resulting in the disease sialuria.[28–30] A second set of single amino acid mutations reduce the activity of this enzyme and result in a different human disease, hereditary inclusion body myopathy (HIBM).[31–33]

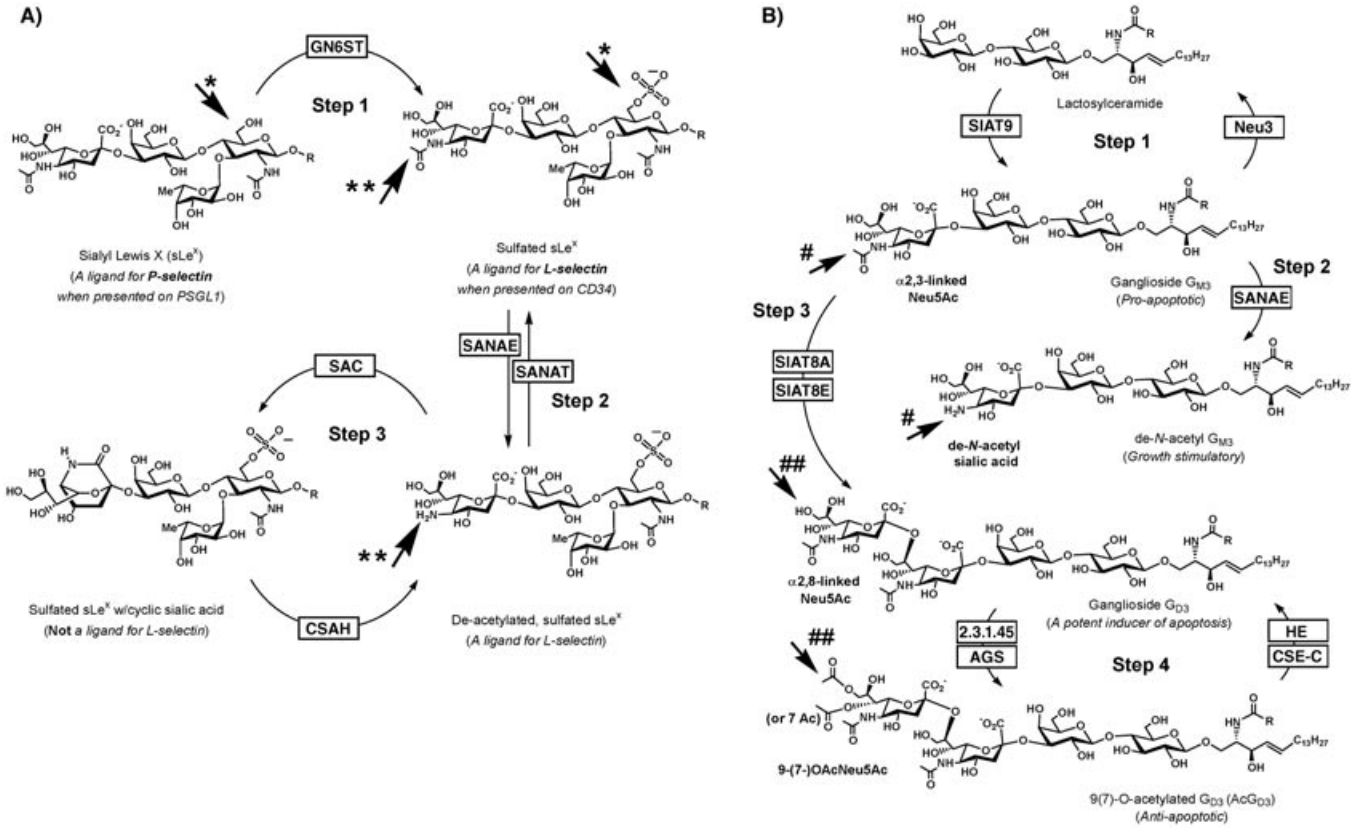
Sialic acid metabolism provides an intriguing system to begin a detailed computational modeling study of glycosylation. This subsystem of glycosylation is considerably more complex than shown in Scheme 2 or in many reviews of sialic acid metabolism.[34,35] Indeed, the core "pathway", designated in gray shading, is amplified in Scheme 3 to show that the metabolic machinery involved in the positioning of just one of the nine monosaccharides incorporated in human glycans is highly complex. Moreover, two important considerations remain for the elements shown in Scheme 3. The first is the regulation of the expression of both the Golgi-localized sialyltransferase and the pre-Golgi "pathway" enzymes. This consideration is important because, in the past, attempts to model metabolic flux without considering the regulation of the enzymes involved have led to the construction of models with poor predictive abilities in general[36–38] and for sialic acid in particular.[20] Accordingly, another factor to consider when producing a computational model of glycosylation is the effect of metabolic flux on the constituent biosynthetic enzymes. The second important consideration is the spatial localization of the sialyltransferases that install sialic acid residues onto the nascent oligosaccharide under synthesis. As indicated in Scheme 3, overlapping sets of these enzymes produce α2,3-, α2,6-, and α2,8-glycosidic linkages on either protein- or lipid-hosted carbohydrate structures. Current understanding is that the exact placement of each type of linkage is largely determined by the exact spatial localization of each glycosyltransferase in the ER or Golgi;[39,40] these results suggest that emerging efforts to incorporate spatial features into modeling studies will be crucial to successful computational approaches to understanding glycosylation.

Glycosylation Lies at the Intersection of Metabolism, Adhesion, and Cell Signaling

Typically, systems level approaches to modeling in metabolism have focused on the output of secreted molecules or on the optimization of factors such as energy production, cell growth, or increase in bulk levels of biomass but not on persistent changes such as those seen in cell-surface carbohydrate architecture and the intracellular processes that produce them. The example of apoptosis discussed above provides one situation where surface-displayed sugars remain intricately connected to, and continue to regulate, both the metabolic pathways that synthesize them and other aspects of cellular function; many other equally important examples are known, including the interaction of growth-factor receptors with intrinsic tyrosine kinases,[41,42] modulation of integrin function,[43–45] and activation of cytoplasmic signal transducers, such as Src family kinases and small G proteins.[5,6,46,47] The complexity of events mediated at the cell surface can be seen by considering the "glycosynapses", which are microdomains characterized by distinctive clustering of glycolipids or glycoproteins. Three types of glycosynapses are now known, each having the ability to simultaneously interact with key sets of transmembrane receptors or signal transducers and modulate cell adhesion and motility.[5] Tying all these complex systems together based on their common connections to glycosylation will require an intense level of investigation for the foreseeable future.

Carbohydrate Biosynthesis Produces "Fine-Tuned" Products that Remain Subject to Dynamic Modification

Perhaps paradoxically, the many large-scale influences of glycosylation in cellular processes are often determined by very slight or subtle chemical changes to carbohydrate structure. Furthermore, such molecular-level "fine-tuning" is often transient and reversible, thereby allowing dynamic modulation of signal transduction and cell adhesion. Once again, sialic acid can be used to provide an illustrative example. As shown in Scheme 4 A, multiple posttranslational modifications control the activity of the sialyl Lewis X (sLe^X) tetrasaccharide in the leukocyte homing process. sLe^X has been shown to function as a ligand for all three (E-, L-, and P-) selectins when studied by in vitro assays but in vivo specificity is more highly stringent and is determined by several factors. First, the identity of the



Scheme 4. *Postsynthetic modifications fine-tune the activity of cell surface carbohydrates. Subtle postsynthetic chemical modifications to cell-surface carbohydrate epitopes influence important cellular behaviors including leukocyte homing (panel **A**) and function as a switch between proliferation and apoptosis (panel **B**). Each of these examples is discussed in more detail in the text. The * and # symbols are visual aids to direct attention to the chemical changes taking place.*

host molecule plays a role in specificity; sLe^X is a ligand for P-selectin when attached to PSGL-1, but it is a ligand for L-selectin when presented on CD34 or Glycam-1 and a ligand for E-selectin when it is hosted by a ganglioside.[48] Further discrimination is achieved by sulfation of position 6 of the GlcNAc and/or galactose residues (Step 1); this modification is required for binding to L-selectin under physiological conditions.[49,50] Further modification of the sialic acid residue of sLe^X by stepwise de-*N*-acetylation of position 5 (Step 2) followed by cyclization of the sialic acid residue (Step 3) inhibits binding to L-selectin. It has been proposed that these two posttranslational modifications of sLe^X allow the immune system to control the recruitment of leukocytes at different rates during routine homing and inflammation. The negative feedback sialic acid cyclase system allows routine homing of leukocytes at a "slow and steady" rate, whereas massive accumulation of lymphocytes to sites of inflammatory lesions is mediated mainly by the interaction of nonsulfated sLe^X with P-selectin.[51,52]

A second example of the role of sialic acid in modulating key cellular events is given in Scheme 4B where the subtle changes in sialic acid structure dramatically affect cellular decision processes by acting as a switch for proliferation and apoptosis modes. More specifically, when the glycolipid lactosylceramide (LacCer) is converted into the G_{M3} ganglioside by addition of a single α2,3-linked sialic acid residue (Step 1), this metabolic product is weakly or moderately proapoptotic in many cell types[6,53] but stimulates cell proliferation when it is further converted into the de-*N*-acetyl form (Step 2).[41,54] In an alternative modification of G_{M3}, the addition of an α2,8-linked sialic acid residue produces the G_{D3} ganglioside (Step 3), which is a potent inducer of apoptosis; interestingly, 9-*O*- or 7-*O*-acetylation of the terminal sialic acid residue of G_{D3} (Step 4) produces AcG_{D3}, a molecule with antiapoptotic properties.[55,56] The ability of cells to rapidly modify their cell-surface expression of oligosaccharides in response to internal signals or extracellular stress in these subtle manners illustrates how minor local changes can have global, whole-cell repercussions that potentially involve numerous cell-regulation pathways. The resultant implications for cell-wide behavior are that command of cell-wide fragility can occur on the molecular scale and the effects of local changes to carbohydrate structure resonate to *various* global processes. Careful attention to detail in modeling efforts will be required to ensure that minor local glycosylation changes are not overlooked amid much larger system-wide changes.

Methods of Systems Analysis

As discussed above, a systems-level analysis of glycosylation must relate the large-scale interactions that contribute to carbohydrate synthesis and cell-surface presentation with single molecule-level processes that fine-tune product conversion in

individual reactions. One of the important aspects of systems analysis is the identification of key regulation mechanisms that occur across both levels, from the molecular feedback inhibition of allosteric enzyme binding, such as the GNE enzyme in the sialic acid network, that affects an entire biosynthetic pathway (Scheme 3) to the impact of an event such as conversion of G_{M3} into G_{D3} (Scheme 4B). This latter event, while only involving one of over a dozen or more sialyltransferases found in a cell and only one of hundreds of different metabolic products bearing sialic acid, has the potential to trigger apoptosis and completely shut down sialic acid biosynthesis.[13] These examples show that, while systems properties have been identified in various extensively studied metabolic pathways, glycosylation provides a rich conduit from which we can explore the complexity of these intracellular mechanisms by observing the resultant cell-surface characteristics, that is, the diverse structure of complex carbohydrates can act as "signatures" of unique sequences of intracellular events. Up to now, there have been few attempts at modeling glycosylation; however, we predict that constraint-based methods, such as metabolic flux analysis (MFA), and more quantitative approaches, such as metabolic control analysis (MCA), can identify these mechanisms to benefit the study of carbohydrate synthesis and surface presentation. Owing to the relative infancy of modeling in glycosylation, we will merely provide the biochemist with an overview of analytical and computational methods currently used to study many metabolic systems and thereby motivate extension of the analysis to glycosylation. In particular, emphasis will be placed on the connections between the intracellular metabolic processes that produce glycans and the surface aspects that present them, as these are often complex and involve many large-scale regulatory motifs, including both synthesis in the forward direction and recycling of surface sugars, as well as the influence of signaling.

Constraint-Based Modeling in Metabolic Systems

In one approach to modeling metabolic systems, entire biochemical reaction networks can be expressed as systems of linear equations, with parameters representing the stoichiometric coefficients. These equations can be further represented by an (m,n) stoichiometric matrix S of m metabolites and n reactions. More technically, the stoichiometric matrix is a linear transformation of an n-length vector of the system variables, the metabolic fluxes (v), which describe the rates of metabolite conversion. The result is a vector of time derivatives of metabolite concentrations, which when expressed at a steady-state yield steady-state fluxes. It can also be viewed as mass-balance constraints on the system, conferred by the stoichiometric reaction coefficients. This is known as the *nullspace* of the linear system [Eq. (1)].[57]

$$Sv = 0 \quad (1)$$

The solutions of Equation (1) are often not unique and additional assumptions, such as reaction capacities, are used to further constrain the possible solution space (an approach common to flux balance analysis[58,59] (FBA)). This establishes a limit on the enzyme turnover rate and allows us to study the flux distributions in more detail. Specifically, from some experimental considerations, one can assume that certain fluxes must be within the capacities given in Equation (2), where lb_i and ub_i are the lower and upper bounds on the reaction flux v_i (where i indicates the index for the fluxes). They would also be subject to thermodynamic constraints expressing the generally valid assumption that each *irreversible* reaction flux must proceed in the forward direction and would therefore need be nonnegative, so Equation (3) would apply.

$$lb_i > v_i > ub_i \quad (2)$$

$$v_i \geq 0 \quad (3)$$

This is obviously true for the fluxes that act as the input and output of the pathway. Furthermore, the elimination of thermodynamically infeasible solutions can further constrain the solution.[60] Given sufficient constraints, whether they originate from thermodynamics, mass balance, or enzyme capacities, we can fully characterize the system, thereby yielding precise regions of allowable metabolic flux.[58]

This stoichiometric method of representation is useful from a systems perspective, as the stoichiometric matrix contains information about the topology or connectivity of reaction species, which is very important in determining the properties of the network.[61,62] Modeling in metabolism typically aims at determining a unique (or precise) set of metabolic flux distributions in a pathway or identifying the elements that can control or yield an optimal flux. In this methodology, one is not required to specify the actual functional form of kinetic equations governing the chemical reactions. Rather, one is satisfied with the simple notion of flux values. These are the central premises behind classic modeling methods, such as MFA, and are extended for more sophisticated methods of pathway reconstruction, such as FBA, elementary node analysis,[63] and extreme pathway analysis.[64] It also provides a more conceptual framework for more precise methods, such as MCA.

Metabolic Flux Analysis

Metabolic flux analysis is a constraint-based method for modeling biochemical reaction networks that relies on reaction stoichiometries and mass balances to determine a precise distribution of fluxes.[65,66] It takes metabolic fluxes as its system variables and applies constraints to reduce their allowable ranges (the S matrix above). Each metabolite is considered a "pool", where the rates of transfer into the pool must be balanced with what exits the pool. The way to analyze this transfer is determined by the difference between the number of unknown fluxes and the metabolites, known as the degrees of freedom. The number of measured fluxes can therefore be equal to, less than, or greater than the degrees of freedom. In the case where there is an equal number of measured fluxes and de-

grees of freedom, the system is uniquely determined and the linear relation in Equation (1) has one solution, which would then satisfy MFA. A unique solution however is rare and it is more common that the degrees of freedom and the measured fluxes are unequal. In cases where there are more measured fluxes, the system is overdetermined and there is redundancy in the system. This redundancy can be exploited to get more precise estimations on both the measured and nonmeasured fluxes. This is however a somewhat ideal case and does not occur frequently. The most common case is one where the system is underdetermined and the degrees of freedom exceed the number of measured fluxes.[67] In this case, there are an infinite number of solutions to satisfy the system and methods of linear programming must be used to reduce the total set of solutions. To do so requires that further assumptions are made concerning the metabolic system. We often presume that the metabolic system in question has an objective, that its behavior is optimized towards achieving one or more goals, which can help us to determine this precise set of flux distributions. In methods like FBA, this goal takes the form of an objective function, a mathematical relationship that relates the fluxes to some extreme value. Often, we are looking to maximize or minimize a particular function and obtain a set of fluxes that correspond to that unique optimized metabolic state. In a landmark paper by Palsson and co-workers, a quantitative relationship between carbon uptake (acetate or succinate), oxygen uptake, and maximal growth was determined in silico for *Escherichia coli* MG1655 and was experimentally verified by making the assumption that metabolism was optimized for growth.[68] By using the reaction stoichiometry to constrain the total solution space into what is termed a "flux cone",[1,69] it was possible to identify the set of flux distributions that corresponded to the maximum growth rates under different conditions. The resultant predicted and experimentally observed fluxes agreed and thus confirmed the proposed objective, the rate of cell growth. Consequently, it was also shown that constraint-based methods of reaction flux determination can be valuable in the absence of precise kinetic information.

The concept of flux cones, in particular, is conceptually accessible for illustrating the importance of a constraint-based approach. The flux cone corresponds to a set of allowable fluxes in the metabolic network. If all fluxes were allowable, then the boundaries of the flux cone would be just the set of axes. Based on just the stoichiometry, however, not all fluxes are allowable and thus a restricted set can be identified, even if the precise values cannot. Moreover, often certain fluxes are known to be not possible, from the knowledge that certain reactions may not occur in the presence of other metabolites, etc. The edges of the flux cones are the extreme pathways of the system fluxes that describe the absolute range of feasibility for the metabolic network. The magnitude of these boundaries can illustrate important systems properties of a metabolic network, such as robustness and fragility. Further analysis can be done to adjust the boundaries or the extreme pathways of the system from experimental invalidation. Additionally, there are various avenues for reducing a set of flux distributions to within acceptable ranges that can characterize a metabolic system uniquely. Such a case could be a paradigm for modeling in glycosylation. Objectives in glycosylation will center on the production of cell-surface carbohydrates but could take the position that the goal is overall diversity or perhaps the maximization of a specific glycoform. The use of FBA alone, however, cannot give quantitative analysis of how the flux is *controlled*. For such information, a more precise, quantitative method should be used.

Sensitivity Analysis in Metabolic Systems

Metabolic control analysis

MCA is a quantitative method for studying how homeostasis is maintained at the molecular level.[70,71] It focuses on the role of enzymes and enzyme activities in controlling how metabolic flux is distributed and how flux distributions are affected by varying external conditions. Specifically, it studies the deviation from steady-state flux levels given infinitesimal alterations to enzyme activities; this is conceptually similar to sensitivity analysis, a methodology used in many different fields. One important aspect of MCA is that it requires detailed knowledge of the reaction kinetics for all the reactions involved. MCA is particularly useful from a systems perspective, as many global properties can be affected by alterations in a single molecule or by a change in the details of a single metabolic reaction. Feedback inhibition, cooperativity, and allosteric effects all contribute to the regulation and control of metabolite synthesis, modification, and degradation. Regulation of enzyme activity through gene expression, proteolytic enzyme activity, and covalent modification all add an additional layer of complexity that can be exploited from MCA.

For such high-resolution analysis, MCA requires nearly complete enzyme characterization and knowledge of the kinetics and constants of the enzyme reactions. Thus, the level of detail required is much higher than for MFA and other constraint-based methods. Additionally, MCA hinges on a number of assumptions. The most critical assumption is that the system under study can be observed under steady state or constant conditions. This means that the rates of "entry" and "exit" for the aforementioned metabolic pools must be balanced. Due to its high reliance on enzyme characterization and kinetics, MCA is a much less scalable method than MFA and it is most appropriate for relatively small metabolic pathways and networks whose molecular and kinetic details are well characterized.

Each enzyme confers a particular sensitivity and elicits a certain change in metabolic flux, which is related by flux control coefficients (FCCs).[70,72] Mathematically, the FCCs ($C^{v_i}_j$) represent a fractional change in metabolite flux (v_i) in response to a fractional change in enzyme activity (E_j), as described by Equation (4).

$$C^{v_i}_j = \frac{\mathrm{d}v_i}{\mathrm{d}E_j}\frac{E_j}{v_i} \text{ for } i,j \in \{1,\ldots,n\} \quad (4)$$

Here, the full derivative is used to imply that both direct and indirect influences of the change in enzyme activity on any

flux v are taken into account. (dE may influence other system variables as well.) Each control coefficient in itself has little value, however, as they cannot be compared across separate systems. Rather, all control coefficients are normalized and are most informative in relation to other control coefficients characterizing enzymes of a given pathway or network. Additionally, control coefficients are not constant, as they reflect the sensitivity to enzymatic change at a certain steady state. Furthermore, if the enzymatic activity is changed significantly beyond equilibrium levels, the control coefficients become less accurate, as the system can behave nonlinearly when far from equilibrium. The concept of metabolic control coefficients may be viewed as a part of a greater theory in metabolic modeling called distributed control, which states that all enzymes in a pathway can share the control of flux (and therefore no individual enzyme has control over flux through the pathway).[37] The proportion of control is related by the *summation theorem* [Eq (5)], which asserts that all of the coefficients must add up to unity (for *each* flux).

$$\sum_{j=1}^{n} C_j^{v_i} = 1 \text{ for i} \in \{1,\ldots,n\} \quad (5)$$

This theorem does, however, leave open the possibility that any individual coefficient can take negative values or values greater than one, as can occur in pathways that contain branches and cycles. Also, by this theorem, the magnitude of any coefficient is most closely related with the *location* of the enzyme in the overall pathway. In contrast, elasticity coefficients (EC) are used to reflect more local properties of the enzymes in the pathway. Elasticity coefficients are similar to flux control coefficients but relate the sensitivity of fluxes (v_i) with metabolite concentrations (X_j), as described by Equation (6).

$$\varepsilon_{X_j}^{i} = \frac{\delta v_i}{\delta X_j}\frac{X_j}{v_i} \text{ for } i \in \{1,\ldots,n\} \text{ and } j \in \{1,\ldots,k\} \quad (6)$$

The difference is that with elasticity coefficients we want to measure the precise change in flux with respect to a particular metabolite. Generally speaking, we only want the direct effects. Furthermore, ECs relate the activity of an enzyme to the substrate concentration and thus capture more of the molecular detail than control coefficients alone. Like control coefficients, however, elasticity coefficients are normalized; the partial differential operators are shown to illustrate that all other variables must be held constant.

For each metabolite X_j, we can also use the *connectivity theorem*, which states that the flux coefficients are related to the elasticity coefficients according to Equation (7), where X_j can actually represent any compound that influences the reaction rate, although it is more commonly applied to metabolite concentrations.

$$\sum_{i=1}^{n} C_i^{v_L}\varepsilon_{X_j}^{i} = 0 \text{ for } i \in \{1,\ldots,n\} \text{ and } j \in \{1,\ldots,k\} \quad (7)$$

Equation (7) is considered to be the most important of the MCA theorems, because it allows us to understand how enzyme kinetics affect flux control.[72]

MCA in glycosylation

Glycosyltransferases are the primary gene products in the "glycosylation machinery", and are critical to the promotion of carbohydrate structural diversity. They could be the focus of a control approach to dissect the diversity of surface carbohydrates. Glycosyltransferases are a general family of carbohydrate-binding enzymes located in the ER and Golgi and can

Table 1. *Monosaccharides, nucleotides, derivatives, and nucleotide sugar donors.*

Abbreviation	Name/description
ATP	adenosine 5′-triphosphate
CMP	cytidine 5′-monophosphate
CMP-Neu5Ac	CMP–*N*-acetylneuraminic acid
CMP-Neu5Gc	CMP–*N*-glycolylneuraminic acid
CMP-Sia	CMP–sialic acid
CTP	cytidine 5′-triphosphate
Fru	D-fructose
Fru-6P	D-fructose-6-phosphate
Fuc	L-fucose
Fuc-1P	L-fucose-1-phosphate
Gal	D-galactose
Gal-1P	D-galactose-1-phosphate
GalNAc	*N*-acetyl-D-galactosamine
GDP	guanidine 5′-diphosphate
GDP-Fuc	GDP–L-fructose
GDP-Man	GDP–D-mannose
GDP-4-oxo-6-deoxy-Man	GDP–4-oxo-6-deoxy-D-mannose
Glc	D-glucose
Glc-1P	D-glucose-1-phosphate
Glc-6P	D-glucose-6-phosphate
GlcA	D-glucuronic acid (or glucuronate)
GlcN	D-glucosamine
GlcN-6P	D-glucosamine-6-phosphate
GlcNAc	*N*-acetyl-D-glucosamine
GlcNAc-1P	*N*-acetyl-D-glucosamine-1-phosphate
GlcNAc-6P	*N*-acetyl-D-glucosamine-6-phosphate
Man	D-mannose
Man-1P	D-mannose-1-phosphate
Man-6P	D-mannose-6-phosphate
ManNAc	*N*-acetyl-D-mannosamine
ManNAc-6P	*N*-acetyl-D-mannosamine-6-phosphate
Neu5Ac	*N*-acetylneuraminic acid (or *N*-acetylneuraminate)
Neu5Ac-9P	*N*-acetylneuraminic acid 9-phosphate
Neu5Gc	*N*-glycolylneuraminic acid (or *N*-glycolylneuraminate)
PEP	phosphoenol pyruvate
Sia	sialic acid (generic term for over 50 compounds)
UDP	uridine 5′-diphosphate
UDP-Gal	UDP–D-galactose
UDP-GalNAc	UDP–*N*-acetyl-D-galactosamine
UDP-Glc	UDP–D-glucose
UDP-GlcA	UDP–D-glucuronic acid
UDP-GlcNAc	UDP-*N*-acetyl-D-glucosamine
UDP-Xyl	UDP–D-xylose
UTP	uridine 5′-triphosphate
Xyl	D-Xylose

Table 2. *Hexose (monosaccharide) transporters.*

Type	Tissue	Sugar transported (Michaelis constant, K_m [mM])	Function/comments
SGLT family			SLC5A family, member no. as shown in left column
SGLT-1 (A1)	small intestine, kidney	glucose (0.1–0.8), galactose	major apical glucose transporter
SGLT-2 (A2)	kidney	glucose (1.6)	glucose uptake
GLUT family			
GLUT-1 (subfamily (sf) I)	erythrocytes, blood/brain barrier	glucose (3–5), galactose, mannose, glucosamine	primary glucose transporter in fetal cells
GLUT-2 (sf I)	liver, small intestine, brain	glucose (17), galactose (92), mannose (125), fructose (76), glucosamine (0.8)	high capacity and low affinity
GLUT-3 (sf I)	neurons, placenta	glucose (1–2), galactose, mannose, maltose, xylose, dehydroascorbate	primary glucose transporter of neurons
GLUT-4 (sf I)	adipose tissue, skeletal muscle,	glucose (5), glucosamine (3.9), dehydroascorbate	insulin-stimulated glucose uptake
GLUT-5 (sf II)	small intestine	fructose (6), glucose (in rat)	primarily fructose absorption
GLUT-8 (sf III)	testis, blastocysts, brain, muscle, adipocytes	glucose (2), galactose, fructose	retained in an intracellular compartment and not responsive to insulin; formerly designated as GLUTX1
GLUT-10 (sf III)	heart, lung, brain, liver, skeletal muscle, pancreas, placenta, kidney	2-deoxy-glucose (0.3), glucose, galactose	
GLUT-11 (sf II)	heart, skeletal muscle	glucose, fructose	mRNA for this protein is detected in many tissues, GLUT-11 occurs in three different splice forms (a–c).

degrade, modify, or create glycosidic linkages, thereby synthesizing oligosaccharides, polysaccharides, and glycoconjugates. Each glycosyltransferase confers high specificity, through their ability to generate one type of glycosidic bond. However, it is the combination of redundancy and "promiscuity" of glycosyltransferases, the phenomenon of multiple enzymes generating the same glycosidic bond and, conversely, the ability of one glycosyltransferase to generate multiple bonds, that accounts for the wide heterogeneity seen at the surface. Moreover, while glycosyltransferases mediate carbohydrate generation for the entire cell surface, they maintain molecular precision and resolution.

Obstacles to metabolic control analysis in glycosylation

The primary obstacle to using these methods for the determination of metabolic fluxes is incomplete characterization of pathway enzymes. Study of the sensitivity of fluxes to small perturbations in enzyme activities implies that the enzyme activities are already known. Accordingly, MCA requires almost complete characterization of all of the enzymes in a metabolic pathway as well as their modes of binding. Incomplete characterization increases the "degrees of freedom" of the system. Typically, with a small number of uncharacterized enzymes, the enzyme parameters are allowed to vary within bounds established by allowable fluxes and the knowledge of the sensitivity of the other parameters in the model. The more "free" parameters there are in the system, however, the less reliable MCA becomes and the more suitable a constraint-based approach becomes. Generally speaking, glycosylation pathways should be amenable to an MCA approach, given that these pathways tend to have a smaller number of enzymes than other biochemical. Nevertheless, we continue to struggle with quantitative modeling of glycosylation due to incomplete characterization.

Table 3. *Nucleotide transport in the Golgi and ER.*

Nucleotide sugar donor	Human gene	Transporter name
ATP		rat Golgi membrane ATP transporter
CMP-Sia	*SLC35A1*	CMPST; CMP–sialic acid Golgi transporter
GDP-Fuc	*SLC35C1*	FUCT1; GDP–fucose transporter 1
GDP-Man		Vrg4p; Yeast GDP–mannose Golgi transporter
3′-phosphoadenosine 5′-phosphosulfate (PAPS)	*PAPST1*	PAPS transporter
UDP-Gal	*SLC35A2*	UGT; UDP–galactose transporter
UDP-GalNAc	*SLC35A2*	UGT (see UDP-Gal)
	SLC35D1	UGTrel7 (see UDP-GlcA)
UDP-Glc		AtUTr1; *Arabidopsis thaliana* UDP–-galactose/UDP–glucose transporter
UDP-GlcA	*SLC35D1*	UGTrel7; UDP–glucuronic acid/UDP–*N*-acetylgalactosamine transporter
UDP-GlcNAc	*SLC35A3*	UDP–*N*-acetylglucosamine transporter

Conclusion

Cell-surface carbohydrates are critical in numerous cellular processes, including cell–cell communication and cell signaling. Glycosylation pathways are known to mediate these processes and to influence cellular decisions, such as those about proliferation and apoptosis. Contrastingly, glycosylation is also able to retain the sensitivity for fine-tuning carbohydrates with molecular detail and resolution. Moreover, the intracellular mechanisms that synthesize these glycans exhibit complexity that is complemented by the extreme diversity and variation of complex carbohydrates seen at the surface. Yet, unlike many other

Table 4. *Enzymes used for monosaccharide processing and carbohydrate assembly.*

Abbreviation	Name/description
2.3.1.3	glucosamine *N*-acetyltransferase
2.3.1.45	*N*-acetylneuraminate 7-*O*-(or 9-*O*-)acetyltransferase
2.7.7.10	galactose-1-phosphate uridylyltransferase
3.1.3.29	*N*-acylneuraminic acid 9-phosphate phosphatase
3.5.1.33	*N*-acetylglucosamine deacetylase
5.1.3.9	*N*-acylglucosamine-6-phosphate 2-epimerase
AGS	*O*-acetyl ganglioside synthetase
CMAH	CMP–Neu5Ac hydroxylase
CMPNS	CMP–Neu5Ac synthetase
CSAH	cyclic sialic acid hydrolase
CSE-C	cytosolic sialic acid 9-*O*-acetylesterase homologue
FPGT	fucose-1-phosphate guanylyltransferase
FUC TR	FUCT1
FUK	L-fucose kinase
GALE	UDP–galactose-4-epimerase
GALK1	galactokinase 1
GALM	galactose mutarotase (aldose 1-epimerase)
GALT	galactose-1-phosphate uridylyltransferase
GCK	glucokinase (hexokinase 4)
GCNT1	glucosaminyl (*N*-acetyl) transferase 1, core 2 (β-1,6-*N*-acetylglucosaminyltransferase)
GCS1	glucosidase 1
GFPT-1	glutamine-fructose-6-phosphate transaminase 1
Ggta1	β-D-galactosyl-1,4-*N*-acetyl-D-glucosaminide α-1,3-galactosyltransferase (mouse)
GMDS	GDP–mannose 4,6-dehydratase
GMPPA	GDP–mannose pyrophosphorylase A
GMPPB	GDP–mannose pyrophosphorylase B
GN6ST	*N*-acetylglucosamine-6-*O*-sulfotransferase
GNE	UDP–GlcNAc 2-epimerase/ManNAc 6-kinase
Gnpnat-1	glucosamine-phosphate N-acetyltransferase 1(mouse)
GNPNAT-1	glucosamine-phosphate N-acetyltransferase 1(human)
GPI	glucose phosphate isomerase
HE	influenza C virus glycoprotein
HK-1	hexokinase 1 isoform HKI-td (brain form hexokinase)
HK-2	hexokinase 2 (muscle form hexokinase)
HK-3	hexokinase 3 ATP:D-hexose 6-phosphotransferase
LPG2	Leishmania GDP–mannose transporter
MPI	mannose phosphate isomerase
NAGK	*N*-acetylglucosamine kinase
Neu1	sialidase 1 (lysosomal sialidase)
Neu2	sialidase 2 (cytoplasmic sialidase)
Neu3	sialidase 3 (membrane sialidase)
Neu4	sialidase 4
NPL	*N*-acetylneuraminate pyruvate lyase
PGM1	phosphoglucomutase 1
PGM3	phosphoglucomutase 3
PMM1	phosphomannomutase 1
RENBP	renin-binding protein (a GlcNAc 2-epimerase)
SAC	sialic acid cyclase
SANAE	sialic acid *N*-acetylesterase
SANAT	sialic acid *N*-acetyltransferase
SAS	N-acetylneuraminic acid phosphate synthase
SIAT1	sialyltransferase 1 (β-galactoside-α2,6-sialyltransferase)
SIAT3C	sialyltransferase 3C
SIAT4A	sialyltransferase 4A
SIAT4B	sialyltransferase 4B
SIAT4C	sialyltransferase 4C
SIAT6	sialyltransferase 6
SIAT7A	sialyltransferase 7A
SIAT7B	sialyltransferase 7B
SIAT7C	sialyltransferase 7C
SIAT7D	sialyltransferase 7D
SIAT7E	sialyltransferase 7E
SIAT7F	sialyltransferase 7F
SIAT8A	sialyltransferase 8A (α2,8-linkages in gangliosides)
SIAT8B	sialyltransferase 8B (α2,8-linkages in polysialic acid)
SIAT8C	sialyltransferase 8C (α2,8-linkages in polysialic acid)
SIAT8D	sialyltransferase 8D (α2,8-linkages in polysialic acid)
SIAT8E	sialyltransferase 8E (α2,8-linkages in gangliosides)
SIAT9	sialyltransferase 9
SIAT10	sialyltransferase 10
TSTA3	tissue-specific transplantation antigen P35B
UAP1	UDP–*N*-GlcNAc pyrophosphorylase 1
UGD	UDP–glucuronate decarboxylase 1 (UXS-1)
UGDH	UDP–glucose dehydrogenase
UGP-2	UDP–glucose pyrophosphorylase 2

metabolic systems and cell-signaling pathways, they have yet to be studied extensively by a systems approach. The demand, however, has been established, as we are increasingly realizing the importance of connecting metabolic product synthesis with other large-scale cellular functions. Thus, modeling in glycosylation will present new challenges for the systems biologist but will also inevitably contribute to our knowledge of the regulation and control of many cellular processes.

Tables 1–4 give definitions of the abbreviations found in the schemes and comments on enzyme functions.

Acknowledgements

K.J.Y. was supported by funding from a Culpeper Biomedical Engineering Pilot Award, the Arnold and Mabel Beckman Foundation, and the Foundation for Advancement in Research in Myopathies (ARM).

Keywords: biochemical pathways · carbohydrates · glycosylation · metabolic flux · systems biology

[1] J. S. Edwards, B. O. Palsson, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5528–5533.
[2] J. S. Edwards, R. U. Ibarra, B. O. Palsson, *Nat. Biotechnol.* **2001**, *19*, 125–130.
[3] R. U. Ibarra, J. S. Edwards, B. O. Palsson, *Nature* **2002**, *420*, 186–189.
[4] S. A. Brooks, M. V. Dwek, U. Schumacher, *Functional and molecular glycobiology*, BIOS Scientific, Oxford, **2002**.
[5] S.-I. Hakomori, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 225–232.
[6] M. Sorice, A. Longo, T. Garofalo, V. Mattei, R. Misasi, A. Pavan, *Glycoconjugate J.* **2004**, *20*, 63–70.
[7] A. Dove, *Nat. Biotechnol.* **2001**, *19*, 913–917.
[8] T. Endo, D. Groth, S. B. Prusiner, A. Kobata, *Biochemistry* **1989**, *28*, 8380–8388.
[9] P. M. Rudd, T. Endo, C. Colominas, D. Groth, S. F. Wheeler, D. J. Harvey, M. R. Wormald, H. Serban, S. B. Prusiner, A. Kobata, R. A. Dwek, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13044–13049.
[10] J. A. Hanover, *FASEB J.* **2001**, *15*, 1865–1876.
[11] N. E. Zachara, G. W. Hart, *Chem. Rev.* **2002**, *102*, 431–438.
[12] N. E. Zachara, C. Butkinaree, G. W. Hart, *Glycobiology* **2003**, *13*, 833.
[13] E. J. Kim, S. G. Sampathkumar, M. B. Jones, J. K. Rhee, G. Baskaran, S. Goon, K. J. Yarema, *J. Biol. Chem.* **2004**, *279*, 18342–18352.
[14] O. Suzuki, Y. Nozawa, M. Abe, *Int. J. Oncol.* **2003**, *23*, 769–774.

[15] O. T. Keppler, M. E. Peter, S. Hinderlich, G. Moldenhauer, P. Stehling, I. Schmitz, R. Schwartz-Albiez, W. Reutter, M. Pawlita, *Glycobiology* **1999**, *9*, 557 – 569.
[16] S. Eda, M. Yamanaka, M. Beppu, *J. Biol. Chem.* **2004**, *279*, 5967 – 5974.
[17] O. T. Keppler, S. Hinderlich, J. Langner, R. Schawartz-Albiez, W. Reutter, M. Pawlita, *Science* **1999**, *284*, 1372 – 1376.
[18] U. S. Bhalla, P. T. Ram, R. Iyengar, *Science* **2002**, *297*, 1018 – 1023.
[19] A. Hoffmann, A. Levchenko, M. L. Scott, D. Baltimore, *Science* **2002**, *298*, 1241 – 1245.
[20] T. J. Monica, D. C. Andersen, C. F. Goochee, *Glycobiology* **1997**, *7*, 515 – 521.
[21] L. Toma, M. A. S. Pinhal, C. P. Dietrich, H. B. Nader, C. B. Hirschberg, *J. Biol. Chem.* **1996**, *271*, 3897 – 3901.
[22] K. Luhn, M. K. Wild, M. Eckhardt, R. Gerardy-Schahn, D. Vestweber, *Nat. Genet.* **2001**, *28*, 69 – 72.
[23] T. Marquardt, T. Brune, K. Luhn, K. P. Zimmer, C. Korner, L. Fabritz, N. van der Werft, J. Vormoor, H. H. Freeze, F. Louwen, B. Biermann, E. Harms, K. von Figura, D. Vestweber, H.G. Koch, *J. Pediatr.* **1999**, *134*, 681 – 688.
[24] C. B. Hirschberg, *J. Clin. Invest.* **2001**, *108*, 3 – 6.
[25] T. Lübke, T. Marquardt, K. von Figura, C. Korner, *J. Biol. Chem.* **1999**, *274*, 25 986 – 26 989.
[26] M. K. Wild, K. Luhn, T. Marquardt, D. Vestweber, *Cells Tissues Organs* **2002**, *172*, 161 – 173.
[27] L. Sturla, R. Rampal, R. S. Haltiwanger, F. Fruscione, A. Etzioni, M. Tonetti, *J. Biol. Chem.* **2003**, *278*, 26 727 – 26 733.
[28] K. J. Yarema, S. Goon, C. R. Bertozzi, *Nat. Biotechnol.* **2001**, *19*, 553 – 558.
[29] R. Seppala, V. P. Lehto, W. A. Gahl, *Am. J. Hum. Genet.* **1999**, *64*, 1563 – 1569.
[30] R. Seppala, F. Tietze, D. Krasnewich, P. Weiss, G. Ashwell, G. Barsh, G. H. Thomas, S. Packman, W. A. Gahl, *J. Biol. Chem.* **1991**, *266*, 7456 – 7461.
[31] I. Eisenberg, N. Avidan, T. Potikha, H. Hochner, M. Chen, T. Olender, M. Barash, M. Shemesh, M. Sadeh, G. Grabov-Nardin, I. Shmilevich, A. Friedmann, G. Karpati, W. G. Bradley, L. Baumbach, D. Lancet, E. Ben Asher, J. S. Beckmann, Z. Argov, S. Mitrani-Rosenbaum, *Nat. Genet.* **2001**, *29*, 83 – 87.
[32] S. Hinderlich, I. Salama, I. Eisenberg, T. Potikha, L. R. Mantey, K. J. Yarema, R. Horstkorte, Z. Argov, M. Sadeh, W. Reutter, S. Mitrani-Rosenbaum, *FEBS Lett.* **2004**, *566*, 105 – 109.
[33] S. Noguchi, Y. Keira, K. Murayama, M. Ogawa, M. Fujita, G. Kawahara, Y. Oya, M. Imazawa, Y. Goto, Y. K. Hayashi, I. Nonaka, I. Nishino, *J. Biol. Chem.* **2004**, *279*, 11 402 – 11 407.
[34] K. J. Yarema, *BioTechniques* **2001**, *31*, 384 – 393.
[35] O. T. Keppler, R. Horstkorte, M. Pawlita, C. Schmidt, W. Reutter, *Glycobiology* **2001**, *11*, 11R – 18R.
[36] W. Wiechert, *J. Biotechnol.* **2002**, *94*, 37 – 63.
[37] J. A. Morgan, D. Rhodes, *Metab. Eng.* **2002**, *4*, 80 – 89.
[38] M. W. Covert, C. H. Schilling, B. Palsson, *J. Theor. Biol.* **2001**, *213*, 73 – 88.
[39] E. G. Berger, *Glycobiology* **2002**, *12*, 29R – 36R.
[40] A. S. Opat, C. van Vliet, P. A. Gleeson, *Biochimie* **2001**, *83*, 763 – 773.
[41] E. G. Bremer, J. Schlessinger, S.-I. Hakomori, *J. Biol. Chem.* **1986**, *261*, 2434 – 2440.
[42] A. J. Yates, J. Van Brocklyn, H. E. Saqr, Z. Guan, B. T. Stokes, M. S. O'Dorisio, *Exp. Cell Res.* **1993**, *204*, 38 – 45.
[43] E. C. Seales, G. A. Jurado, B. A. Brunson, S. L. Bellis, *Glycobiology* **2003**, *13*, 860 – 861.
[44] A. C. Semel, E. C. Seales, A. Singhal, E. A. Eklund, K. J. Colley, S. L. Bellis, *J. Biol. Chem.* **2002**, *277*, 32 830 – 32 836.
[45] E. C. Seales, G. A. Jurado, A. Singhal, S. L. Bellis, *Oncogene* **2003**, *22*, 7137 – 7145.
[46] M. Sorice, I. Parolini, T. Sansolini, T. Garofalo, V. Dolo, M. Sargiacomo, T. Tai, C. Peschle, M. R. Torrisi, A. Pavan, *J. Lipid Res.* **1997**, *38*, 969 – 980.
[47] K. Iwabuchi, K. Handa, S. Hakomori, *J. Biol. Chem.* **1998**, *273*, 33 766 – 33 773.
[48] M. M. Burdick, B. S. Bochner, B. E. Collins, R. L. Schnaar, K. Konstantopoulos, *Biochem. Biophys. Res. Commun.* **2001**, *284*, 42 – 49.
[49] A. Bistrup, S. Bhakta, J. K. Lee, Y. Y. Belov, M. D. Gunn, F.-R. Zuo, H. Chiao-Chain, R. Kannagi, S. D. Rosen, S. Hemmerich, *J. Cell Biol.* **1999**, *145*, 899 – 910.
[50] K. Uchimura, H. Muramatsu, T. Kaname, H. Ogawa, T. Yamakawa, Q. Fan, C. Mitsuoka, R. Kannagi, O. Habuchi, I. Yokoyama, K. Yamamura, T. Ozaki, A. Nakagawara, K. Kadomatsu, T. Muramatsu, *J. Biochem.* **1998**, *124*, 670 – 678.
[51] C. Mitsuoka, K. Ohmori, N. Kimura, A. Kanamori, S. Komba, H. Ishida, M. Kiso, R. Kannagi, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1597 – 1602.
[52] R. Kannagi, *Curr. Opin. Struct. Biol.* **2002**, *12*, 599 – 608.
[53] Y. Nakatsuji, R. H. Miller, *Exp. Neurol.* **2001**, *168*, 290 – 299.
[54] N. Hanai, T. Dohi, G. A. Nores, S. Hakomori, *J. Biol. Chem.* **1988**, *263*, 6296 – 6301.
[55] H. Y. Chen, A. Varki, *J. Exper. Med.* **2002**, *196*, 1529 – 1533.
[56] F. Malisan, L. Franchi, B. Tomassini, N. Ventura, I. Condo, M. R. Rippo, A. Rufini, L. Liberati, C. Nachtigall, B. Kniep, R. Testi, *J. Exper. Med.* **2002**, *196*, 1535 – 1541.
[57] H. Anton, C. Rorres, *Elementary Linear Algebra: Applications Version*, 8th ed., Wiley, New York, **2000**.
[58] M. W. Covert, I. Famili, B. O. Palsson, *Biotechnol. Bioeng.* **2003**, *84*, 763 – 772.
[59] B. Palsson, *Nat. Biotechnol.* **2000**, *18*, 1147 – 1150.
[60] D. A. Beard, S. D. Liang, H. Qian, *Biophys. J.* **2002**, *83*, 79 – 86.
[61] J. A. Papin, N. D. Price, S. J. Wiback, D. A. Fell, B. O. Palsson, *Trends Biochem. Sci.* **2003**, *28*, 250 – 258.
[62] C. H. Schilling, D. Letscher, B. O. Palsson, *J. Theor. Biol.* **2000**, *203*, 229 – 248.
[63] S. Schuster, D. A. Fell, T. Dandekar, *Nat. Biotechnol.* **2000**, *18*, 326 – 332.
[64] C. H. Schilling, S. Schuster, B. O. Palsson, R. Heinrich, *Biotechnol. Prog.* **1999**, *15*, 296 – 303.
[65] J. Nielsen, *Biotechnol. Bioeng.* **1998**, *58*, 125 – 132.
[66] W. Wiechert, *J. Biotechnol.* **2002**, *94*, 37 – 63.
[67] K. J. Kauffman, P. Prakash, J. S. Edwards, *Curr. Opin. Biotechnol.* **2003**, *14*, 491 – 496.
[68] J. S. Edwards, R. U. Ibarra, B. O. Palsson, *Nat. Biotechnol.* **2001**, *19*, 125 – 130.
[69] M. W. Covert, C. H. Schilling, B. Palsson, *J. Theor. Biol.* **2001**, *213*, 73 – 88.
[70] D. A. Fell, *Biochem. J.* **1992**, *286*, 313 – 330.
[71] D. Fell, *Understanding the Control of Metabolism (Frontiers in Metabolism 2)*, Portland Press, London, **1997**.
[72] G. Stephanopoulos, A. A. Aristidou, J. H. I. Nielsen, *Metabolic Engineering: Principles and Methodologies*, Academic Press, San Diego, **1998**.

Received: May 7, 2004

