

Glycoscience finally comes of age

To take its place alongside genomics and proteomics, glycoscience needs recognition from scientists

Anthony H. Merry & Catherine L.R. Merry

Glycoscience—the study of the complex carbohydrates on the surface of proteins and lipids—has long been the neglected stepchild of molecular biology. Although genomics, proteomics and now systems biology have become hot topics of biomedical research, attracting both researchers and funding, glycoscience has played a rather minor role. Once of little interest to understanding biology on a larger scale, the field is increasingly being recognized as critically important for the next phase of biological and medical research. Following the genomic and proteomic revolutions, an increased knowledge of sugars, and their composition, synthesis and function in a wide variety of cellular processes, will be necessary to understand structural diversity and recognition, and the transfer of complex information among cells and organs. The perception among glycoscientists themselves is also changing—they now view their research in the larger context of biology rather than in isolation.

Glycoscience ... has long been the neglected stepchild of molecular biology

General interest in glycoscience continues to grow as academic and industrial researchers realize its importance to many biological processes. Recent advances in analytical technologies, carbohydrate chemistry, structural biology and transgenics—particularly knockout studies in mice—have made the discipline more accessible to the wider scientific community. Glycoscience is particularly relevant in the

post-genomic era, as glycosylation is one of the main post-translational events and is highly dynamic in nature. Its importance in cell signalling, the immune response and development is only now becoming fully appreciated. A major promise of glycoscience is that it could be applied to manipulate cell–cell communication in development or disease, which in turn could lead to greater precision and specificity in drug targeting. Potential applications in developing diagnostic or therapeutic products based on specific carbohydrates on the cell surface continue to attract both scientists and funds to this field.

The scant attention paid to glycoscience over the past few decades is surprising, considering that British scientist Walter Morgan discovered, almost 60 years ago, that carbohydrates are the cell markers that define the ABH blood group system (Morgan, 1950). His work indicated the importance of sugar moieties for immune reactions and cell recognition. Although further work inspired considerable research and commercial interest in glycoscience (Varki, 1993), the field has not yet joined the mainstream of scientific interest. This is despite observations and experimental demonstrations of the biological role of glycosylation in development, immune reactions, wound healing, disease and many other cellular processes (Varki, 1993; Sharon & Lis, 1993; Dwek, 1995, 1998; Haltiwanger & Lowe, 2004). These have generally been regarded by many scientists as noteworthy, but isolated, insights that are intrinsically random and too diverse to have any generalized function. One example of the lack of understanding of glycoscience is the fact that more than

30% of all glycoprotein entries in the Research Collaboratory for Structural Bioinformatics' Protein Data Bank are either incorrect or highly implausible (Lutteke & von der Lieth, 2004).

The low esteem for glycoscience is also reflected in the amount of funding it has received. The cost of sequencing the human genome was between US\$50 million and US\$150 million. By contrast, the total funding for glycoscience between 1993 and 2003 was probably less than US\$5 million (Schmidt, 2002).

A major promise of glycoscience is that it could be applied to manipulate cell–cell communication in development or disease...

This bias is also observed, apart from a few notable exceptions, in the investments made by the pharmaceutical and biotech industries. During the 1990s, commercial interest and funding were significantly and negatively influenced by the failure of various carbohydrate-based drugs, such as sialyl Lewis X. As this molecule has important roles in regulating inflammatory processes and in cancer metastasis, over-ambitious biotech companies hyped their products as the next wonder drugs in the war against cancer and inflammatory diseases. Their failure to fulfil their promises was most probably caused by insufficient basic research, but it further decreased commercial and medical interest in glycoscience.

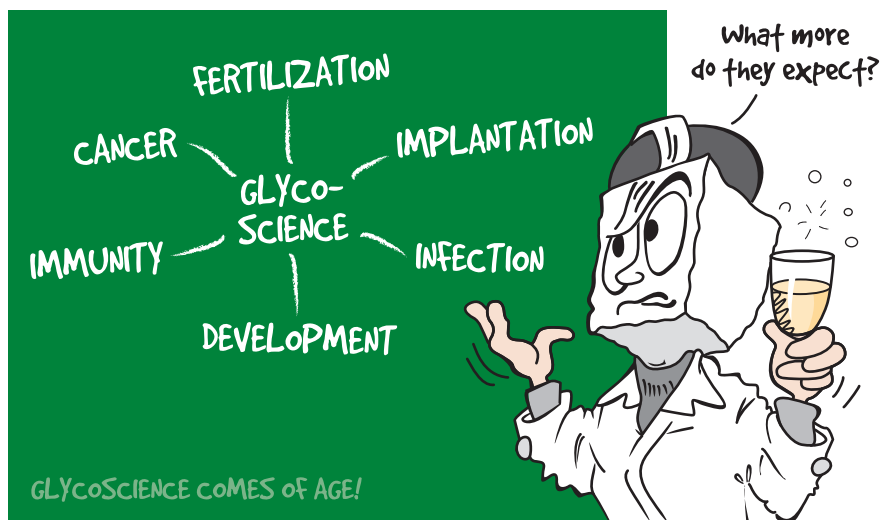
The true nature and importance of glycoscience in the larger context of molecular and cellular biology has not been fully

understood. Nor has it lived up to its original promises of novel therapeutic approaches. This may be because many glycoscientists approach problems in an isolated way rather than looking at the bigger picture—similar to a botanist studying the function of a leaf by concentrating on the differences between individual leaves, rather than asking about the purpose and function of the leaf for the plant as a whole.

Now, after a troublesome beginning, glycoscience can proudly join the rest of molecular biology. In the post-genomic era of the twenty-first century, attitudes are changing. It is becoming clear that, although knowledge of the genome is vital, this alone cannot provide a comprehensive answer as to how biological systems operate. Glycoscience is thus pertinent to the growing research field—often termed ‘systems biology’ (Liu, 2005)—that looks at how biological systems function as a whole, rather than investigating their individual components. Indeed, much of glycoscience only makes sense when viewed in such a holistic way.

Specific glycosylation patterns on the surface of proteins and lipids are one of the most characteristic features of certain cell types, such as embryonic stem cells (Muramatsu & Muramatsu, 2004); tissues, such as the kidney (Palcic *et al*, 1990); or even a whole organism, such as the ABH blood group antigens (Morgan & Watkins, 2000). The sugars in the supramembrane surface layer of cells—the glycocalyx—are optimally positioned to help the cell communicate with its neighbours and the environment. Furthermore, the combinatorial possibilities inherent in glycan synthesis and structure exceed peptide-based structural diversity by several orders of magnitude (Gabius, 2000). This creates an enormous number of options to interact with molecules in the environment or on neighbouring cells, and to modulate the binding affinities of their specific receptors on the cell surface.

Glycosylation is also a prime example of a cellular process—the post-transcriptional modification of proteins and lipids—that is not under the direct control of the genome (Paulson & Colley, 1989). It is affected by a multitude of factors, such as the general cell metabolism and the rate of cell growth. Although the enzymes that synthesize individual monosaccharides and the glycosyltransferases that add these



monosaccharides to glycan chains are defined by the genome, their activity in any given cell at any given time cannot be predicted, even when the transcriptome of that cell is available (Varki, 1998). The enormous variability within the glyco-repertoire is a result of the dynamics of sugar metabolism, as a few biosynthetic components—enzymes and monosaccharides—are able to generate multiple sugar structures depending on factors such as their distribution within the Golgi apparatus and their relative abundance (de Graffenried & Bertozzi, 2004).

The true nature and importance of glycoscience in the larger context of molecular and cellular biology has not been fully understood

The ability to characterize cells or tissues by their glycosylation patterns shows that this diversity is controlled, but that this control—unlike the synthesis of other macromolecules—is not hardwired into DNA. This dynamic regulation of the overall structure and shape of sugar molecules enables the cell to respond quickly to change; a small change in the activity of one glycosyltransferase, for example, is able to redefine an entire sugar epitope because of the effect on subsequent glycosynthesis (Mare & Trinchera, 2004). These dynamics are a critical part of development (Haltiwanger & Lowe, 2004) and of our immune system's ability to respond to disease, injury and infection. There are

interesting therapeutic possibilities for modulating these systems, which have made glycoscience research a topic of growing importance. The field is also aided considerably by recent cloning and knock-out technologies that allow scientists to study the biological role and function of specific sugar-synthesizing enzymes and glycosyltransferases (Furukawa *et al*, 2001; Forsberg & Kjellén, 2001). Nevertheless, much work on the basic mechanisms of glycosylation remains to be done. Although some synthetic pathways and their regulatory mechanisms are relatively well understood, most notably for *N*-linked glycans, many others remain puzzling, such as those for glycosaminoglycans.

It is now widely accepted that it is not only the type of carbohydrate that is important for biological function, but also the nature and presentation of glycan chains on the protein or lipid molecule. Generally, carbohydrates interact only weakly, but the combined presentation of many epitopes on a common core can enhance binding affinity; indeed, strong carbohydrate-based interactions with proteins occur frequently on the cell surface (Evans & MacKenzie, 1999). This is exemplified in the studies of carbohydrate-binding proteins known as selectins, transmembrane molecules on the surface of leukocytes and activated endothelial cells and many other cell types. During inflammation, they regulate the initial binding of leukocytes to, and their subsequent slow movement along, the endothelium by reversible adhesive interactions with glycans on the endothelial

cells. The ligands for this family of proteins are now being identified, and the results show that the presentation of the same glycan in different ways modifies its binding affinity to different members of the selectin family (Crocker & Varki, 2001; Daniels *et al*, 2002).

...glycoscientists—like the sugars they are studying—must be highly interactive to function well

In addition, a single glycan chain can contain several regions with different and specific functions. A well-investigated example for this intrinsic diversity is heparan sulphate, which is involved in the binding and regulation of various growth factors to their specific receptors on the cell surface (Gallagher, 2001). Although heparan sulphate molecules show a high diversity among different chains, the linear sequence of individual monosaccharides in heparan sulphate consists of well-defined domains with highly sulphated regions of variable length, separated by extended non-sulphated sugar repeats. Most ligand-binding motifs characterized so far are within these sulphated regions. The pattern of these charged groups dictates the strength of growth-factor binding and regulates cell growth and differentiation by modulating the binding affinity to various growth factors. However, the spacing between these domains also governs interactions with other—particularly multimeric—proteins such as the anti-angiogenic molecules endostatin (Blackhall *et al*, 2003) and platelet factor 4 (Stringer & Gallagher, 1997). Finally, recent evidence suggests that heparan-sulphate chains may fulfil another function when shed from the cell surface. Experiments in mice indicate that an excess of soluble heparan sulphate delays wound repair by inhibiting cell proliferation at wound edges (Elenius *et al*, 2004). This specialized complex carbohydrate is thus able to catalyse a wide range of diverse functions, with the primary pattern of saccharides and the secondary domain structure within the chain both conferring specificity.

Two more examples from completely different areas illustrate the significance of glycosylation and its implications for

biomedical research. Recently, research revealed that the product of *fringe*, a universal and essential developmental gene, is a glycosyltransferase, that points to the importance of glycosylation in development (Haltiwanger & Lowe, 2004). Similarly, neural biology increasingly acknowledges the importance of glycosylation in forming basic neural junctions during the development of the nervous system (Martin, 2003). This is not surprising given that congenital disorders of glycosylation are known to impair neural development (Miossec-Chauvet *et al*, 2003).

...most of us know our own blood group, but how many of us will ever know our complete genome sequence?

The second example concerns the recent discovery that human embryonic stem cells grown under certain conditions express glycoconjugates on their surface that contain a form of sialic acid, N-glycolyl neuraminic acid, which is not found in humans (Martin *et al*, 2005). As this modification usually produces an antigenic response against the alien carbohydrate when the cells are implanted, it would render these stem-cell lines useless for therapeutic purposes. The possibility that this could invalidate further stem-cell research with these specific cell lines received worldwide attention (Vogel, 2005). The finding provoked a rapid response from stem-cell researchers, and several groups are now working to avoid contamination by developing alternative culture conditions for these stem cells, which are notoriously difficult to grow (Sjogren-Jansson *et al*, 2005). Far from being a novel observation in stem-cell biology, one of the most interesting aspects of this discovery is that glycans provide effective markers for characterizing the state of pluripotency of stem cells, both from mouse and human (Muramatsu & Muramatsu, 2004). One of the main problems in culturing human stem cells so far is a lack of suitable surface markers to characterize their state of pluripotency and differentiation. More research into these cells' specific carbohydrates could eventually provide such markers, with important implications for both research and future therapeutic applications.

These and numerous other findings all contribute to a greater recognition of the importance of glycoscience in human development and disease. As a result, the field is slowly moving back into the focus of mainstream biology. It also seems that the tide is now turning with respect to funding, with the establishment of the Consortium for Functional Glycomics (CFG; www.functionalglycomics.org), an initiative funded by the US National Institute of General Medical Sciences (Bethesda, MD, USA) to understand the role of carbohydrate-protein interactions at the cell surface and in cell-cell communication. Another promising sign is the establishment of a glycomics programme as part of the Human Genome Project in Japan and the creation of organizations such as the multidisciplinary Glycoscience Forum (www.glycosciences.org.uk) in the UK. Bioinformatics will have a vital role in all of these projects, and it is therefore promising that the European Union is funding, under its Sixth Framework Programme, the creation of EuroCarbDB (www.eurocarbdb.org), a series of integrated carbohydrate databases and informatics tools to pool information on glycoscience and glycomics. The long-term goal is to expand this into a depository for carbohydrate-related data, similar to the extensively used data collections in genomics and proteomics. Similarly, the CFG databases of carbohydrate structures, carbohydrate-binding proteins and glycosyltransferases are also becoming a useful resource for researchers in glycoscience. These developments indicate a growing interest in this area from academia, industry and governments.

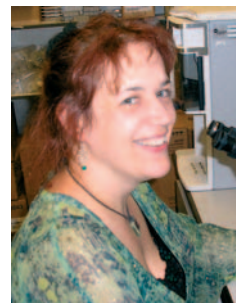
For glycoscience to further gain attention and funding in the future, researchers in the field will need to change their habits. Instead of working in isolation and looking only at their molecule of interest, glycoscientists—like the sugars they are studying—must be highly interactive to function well. Rather than only attending specialized meetings that can, at times, be introspective and as concerned with semantics as with scientific novelty, it is time for researchers to make the most of this renewed interest. The time has come to infiltrate the wider scientific community and to welcome researchers from other fields. It is also important that those outside the field who have an interest in

glycoscience are invited to meetings so that they can show what they expect from the community.

To truly come of age and take its place alongside genomics and proteomics, glycoscience needs proper recognition as a routine part of biological research, not only from funding agencies, but also from scientists themselves. With growing interest and newly available technologies, glycomic analysis should become as common as PCR, and sugar sequencing as easy as DNA sequencing. After all, most of us know our own blood group, but how many of us will ever know our complete genome sequence?

REFERENCES

- Blackhall FH, Merry CL, Lyon M, Jayson GC, Folkman J, Javaherian K, Gallagher JT (2003) Binding of endostatin to endothelial heparan sulphate shows a differential requirement for specific sulphates. *Biochem J* **375**: 131–139
- Crocker PR, Varki A (2001) Siglecs, sialic acids and innate immunity. *Trends Immunol* **22**: 337–342
- Daniels MA, Hogquist KA, Jameson SC (2002) Sweet 'n' sour: the impact of differential glycosylation on T cell responses. *Nat Immunol* **3**: 903–910
- de Graffenried CL, Bertozzi CR (2004) The roles of enzyme localisation and complex formation in glycan assembly within the Golgi apparatus. *Curr Opin Cell Biol* **16**: 356–363
- Dwek RA (1995) Glycobiology: more functions for oligosaccharides. *Science* **269**: 1234–1235
- Dwek RA (1998) Biological importance of glycosylation. *Dev Biol Stand* **96**: 43–47
- Elenius V, Gotte M, Reizes O, Elenius K, Bernfield M (2004) Inhibition by the soluble syndecan-1 ectodomains delays wound repair in mice overexpressing syndecan-1. *J Biol Chem* **279**: 41928–41935
- Evans SV, MacKenzie CR (1999) Characterization of protein–glycolipid recognition at the membrane bilayer. *J Mol Recognit* **12**: 155–168
- Forsberg E, Kjellén L (2001) Heparan sulfate: lessons from knockout mice. *J Clin Invest* **108**: 175–180
- Furukawa K, Takamiya K, Okada M, Inoue M, Fukumoto S, Furukawa K (2001) Novel functions of complex carbohydrates elucidated by the mutant mice of glycosyltransferase genes. *Biochem Biophys Acta* **1525**: 1–12
- Gabijs HJ (2000) Biological information transfer beyond the genetic code: the sugar code. *Naturwissenschaften* **87**: 108–121
- Gallagher JT (2001) Heparan sulfate: growth control with a restricted sequence menu. *J Clin Invest* **108**: 357–361
- Haltiwanger RS, Lowe JB (2004) Role of glycosylation in development. *Annu Rev Biochem* **73**: 491–537
- Liu ET (2005) Systems biology, integrative biology, predictive biology. *Cell* **121**: 505–506
- Lutheke T, von der Lieth CW (2004) pdb-care (PDB carbohydrate residue check): a program to support annotation of complex carbohydrate structures in PDB files. *BMC Bioinformatics* **5**: 69
- Mare L, Trincherà M (2004) Suppression of β 1,3galactosyltransferase β 3Gal-T5 in cancer cells reduces sialyl-Lewis x and enhances poly N-acetylglucosamines and sialyl-Lewis x on O-glycans. *Eur J Biochem* **271**: 186–194
- Martin PT (2003) Glycobiology of the neuromuscular junction. *J Neurocytol* **32**: 915–929
- Martin MJ, Muotri A, Gage F, Varki A (2005) Human embryonic stem cells express an immunogenic nonhuman sialic acid. *Nat Med* **11**: 228–232
- Miossec-Chauvet E, Mikaeloff Y, Heron D, Merzoug V, Cormier-Daire V, de Lonlay P, Matthijs G, Van Hulle C, Ponsot G, Seta N (2003) Neurological presentation in pediatric patients with congenital disorders of glycosylation type Ia. *Neuropediatrics* **34**: 1–6
- Morgan WT (1950) Nature and relationships of the specific products of the human blood group and secretor genes. *Nature* **166**: 300–302
- Morgan WT, Watkins WM (2000) Unravelling the biochemical basis of blood group ABO and Lewis antigenic specificity. *Glycoconj J* **17**: 501–530
- Muramatsu T, Muramatsu H (2004) Carbohydrate antigens expressed on stem cells and early embryonic cells. *Glycoconj J* **21**: 41–45
- Palcic MM, Ripka J, Kaur KJ, Shoreibah M, Hindsgaul O, Pierce M (1990) Regulation of N-acetylglucosaminyltransferase V activity. Kinetic comparisons of parental, Rous sarcoma virus-transformed BHK, and L-phytohemagglutinin-resistant BHK cells using synthetic substrates and an inhibitory substrate analog. *J Biol Chem* **265**: 6759–6769
- Paulson JC, Colley KJ (1989) Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J Biol Chem* **264**: 17615–17618
- Schmidt K (2002) Sugar rush. *New Sci* **176**: 34–37
- Sharon N, Lis H (1993) Carbohydrates in cell recognition. *Sci Am* **268**: 82–89
- Sjogren-Jansson E, Zetterstrom M, Moya K, Lindqvist J, Strehl R, Eriksson PS (2005) Large-scale propagation of four undifferentiated human embryonic stem cell lines in a feeder-free culture system. *Dev Dyn* **223**: 1304–1314
- Stringer SE, Gallagher JT (1997) Specific binding of the chemokine platelet factor 4 to heparan sulfate. *J Biol Chem* **272**: 20508–20514
- Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**: 97–130
- Varki A (1998) Factors controlling the glycosylation potential of the Golgi apparatus. *Trends Cell Biol* **8**: 34–40
- Vogel G (2005) Ready or not? Human ES cells head toward the clinic. *Science* **308**: 1534–1538



Anthony H. Merry is at Glycoscience Consultancy in Charlbury, UK. Catherine L.R. Merry is at Cancer Research UK and the Department of Medical Oncology at the University of Manchester, UK. E-mail: tony@merrystone.freeserve.co.uk

doi:10.1038/sj.embor.7400547