Chapter 1

Sugars And Proteins: Getting It Together

The discovery of major roles for carbohydrates in biology is far from a recent development. Indeed, our knowledge of the existence of the simple sugars found in natural foodstuffs, such as honey, and their use in the fermenting process is as old as civilised man himself. Moreover, beer- and wine-making is by far the earliest practical exploitation of any biological transformation. However, following the advent of modern science and the triumphant elucidation of the pathways of carbohydrate metabolism during the first half of the twentieth century (and of the defects responsible for many of the associated human disorders, such as galactosaemia, shortly afterwards), sugars were prematurely consigned to the margins of scientific research. By the spring of 1953, amid a flurry of activity, Francis Crick and James Watson had unveiled their proposed structure for DNA,[1] heralding an unprecedented sea-change in scientific thinking and providing scientists with a new focus. Unfortunately, in contrast to the 'life molecule,' carbohydrates continued to hold little appeal for the vast majority of scientists: after all, their structures and functions were at the time considered well-defined, and their biology, at best, unexciting. As Nathan Sharon, a pioneer in the field of glycoproteins, notably remarked in 1993, carbohydrates were long regarded as 'second-class citizens' of the cell.[2] This is not to say that researchers failed to appreciate the fundamental importance of carbohydrates in living systems: organisms as diverse as bacteria and humans had been found to metabolise dietary sugars (be they simple sugars such as glucose, sucrose and lactose, or much larger carbohydrates like starch and cellulose) to provide energy, or, in the higher species, to appropriate them for the production of polysaccharides, a class of high molecular-weight carbohydrate polymers. The latter appeared to act solely as structural and protective materials, such as cellulose in plants and chitin in the exoskeletons of insects, or alternatively as space-efficient storehouses

R. M. Bill et al., *Protein Glycosylation*

designed as food reserves; starch and glycogen being classic examples. Yet these obvious, global functions for sugars were disingenuous, in that they misled investigators to believe that the biology of carbohydrates was 'cut and dried' and that such molecules were lacking in any sort of biological specificity or 'intelligence.' Even the presence of the sugar, ribose, as a building block in the newly-characterised nucleic acids saw carbohydrates merely serving a familar structural role; they simply provided the molecular framework within which the four information-carrying bases, G, C, A, and T, were accommodated. Furthermore, although carbohydrates were known to be the essential products of photosynthesis in green plants, arguably the singular most important process of life on Earth, the key to such biological puzzles, and hence the attentions of the biochemists, lay not with the sugars themselves but with their manufacture, as facilitated by the molecular light-trap, chlorophyll, and as orchestrated by a battery of enzymes within the cell.

The widely-held, but mistaken belief that biological macromolecules neatly fall into the four mutually-exclusive groups of nucleic acids, proteins, carbohydrates and lipids (largely founded on the structural analytical work of Nobel Laureate, Albrecht Kossel), led researchers to long disregard a key observation: namely the often-noted presence in highly-purified protein samples of small quantities of sugar. Conversely, traces of protein were routinely detected in otherwise pure samples of the so-called 'mucopolysaccharides' (later to be called 'proteoglycans'). Scientists of the time quite reasonably attributed the appearance of such 'contaminants' to the limitations of the techniques then available, and assumed that only time was needed before improved purification methods would inevitably solve the problem. Yet it was not so, and speculation began to grow that there might exist a new class of hybrid biological molecules, consisting of proteins and carbohydrates in molecular partnership. Nonetheless, it was not until 1958 that there emerged conclusive proof of a specific covalent linkage between the two groups of compounds.[3]

The onset of the 1960s marked the beginning of a new era of research into the biology of carbohydrates, although it would be some time before a suitable neologism, 'glycobiology,' was coined to describe this emergent field.[4] Over the next two decades, it rapidly transpired that sugars were in fact an integral component of a broad range of molecules, collectively termed by French biochemist, Jean Montreuil, 'the glycoconjugates.'[5] These comprise compounds in which sugar chains are covalently attached to either polypeptide or lipid chains to form, variously, glycoproteins, proteoglycans and glycolipids, or in which carbohydrate acts as a molecular bridge, linking protein on the one hand to lipid chains on the other, as in the glycosylphosphatidylinositol (GPI) membrane anchors (see Fig. 1, right). It also became abundantly clear that carbohydrate moieties were performing a

host of important, theretofore unsuspected, biological functions. These range widely from conferring certain immediate environmental advantages on the molecule as a whole, such as greater stability and improved resistance to degradation, to acting as the key molecular arbiters for much of the social communication that goes on between the cells of complex organisms. In particular, the discovery, in 1968,[6] that the sugar chains of serum glycoproteins were responsible for signalling the timely removal of these molecules from the blood stream of higher animals sparked a sudden surge of interest amongst scientists; in retrospect this finding has major implications for the design of glycoprotein drugs. Sugars, it emerged, were laying claim to important physiological and pharmaceutical properties, and could no longer be ignored. In the more recent past, immunologists and physicians have begun to realise that the absence of the normal saccharide portions from recombinant therapeutics and adjuvants provokes an immune responsefrom some patients during trials.[7] Still more exciting is the overwhelming body of evidence indicating that the carbohydrates attached to

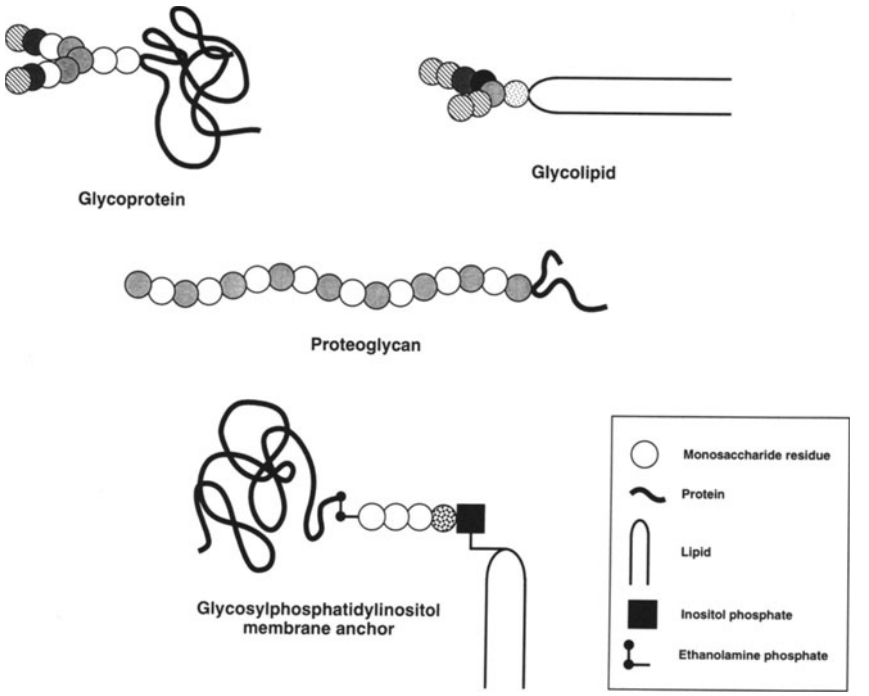


Figure 1. A schematic representation of the four major types of glycosylated macromolecule found in Nature. Shading of circles indicates different types of monosaccharide components.

glycoconjugates at the surface of cells play pivotal roles in the early stages of infection by pathogenic invaders, such as bacteria and viruses.[8] They also provide a gauge of malignancy, and have been highlighted as structures of primary importance in metastasis and the progression of cancer.[9-11] This realisation that sugars are not just dull raw materials, as previously thought, but in fact serve as sophisticated information-carriers with subtle biological specificities has attracted a great deal of attention from researchers in recent years, and has promoted a universal re-evaluation of the importance of carbohydrates in biology.

As with many other life sciences, glycobiology has been revolutionised during the last decade by the rapid and powerful techniques of molecular biology. As a result, great strides have now been made, bridging our knowledge of the way in which glycoconjugates are biosynthesised, and of the tasks they perform in vivo. In the former case, an extensive repertoire of biosynthetic enzymes, the glycosyltransferases and glycosidases, have been isolated and studied; in the latter, new genetic and immunological methods have emerged, such as cloning and the raising of monoclonal antibodies, which allow us to search the molecular haystack for the proverbial needle: carbohydrate chains with distinct biological functions. Today, glycobiology is a well-established branch of biological research, complete with its own scientific journals, international meetings and specialist commercial businesses, and the tremendous expansion it has enjoyed is reflected by the emergence of a growing spectrum of sub-genres such as glycotechnology,[12] glycoimmunology[13] and glycopathology,[14] to name a few. For convenience, these related fields are commonly referred to under the umbrella of 'glycosciences.'

1. THE CHEMICAL STRUCTURE OF SIMPLE SUGARS IN NATURE: BUILDING BLOCKS

Before casting our eyes back over some of the historical highlights of carbohydrate research, it is useful first to summarise the salient features of sugar chemistry, and in particular to define the chemical structures of those simple components that, when linked together, form the more complex carbohydrate moieties commonly found attached to proteins and lipids. The family of molecules collectively known as carbohydrates are so-called because they were determined by early chemical methods to be just that: hydrates of carbon, sharing the general chemical formula $C_n(H_2O)_x$. Carbohydrates are frequently termed saccharides (from the Latin, *saccharum*, for sugar), allowing for further, convenient classification and nomenclature using the well-established chemical prefixes—mono-, di-, tri-,

oligo-, poly-, etc.—to define the number of minimal carbohydrate units—one, two, three, several or many, respectively—that comprise the chosen compound. Logically, then, monosaccharides are the simplest sugar molecules because hydrolysing them no longer releases smaller carbohydrates, and larger saccharides are merely the result of linking various monosaccharides together in a particular configuration (although, as we shall see, many such oligo- and polysaccharides are further elaborated by the addition of inorganic groups such as phosphate and sulphate). The distinction between oligosaccharides, on the one hand, as carbohydrates consisting of several monosaccharide components and polysaccharides, on the other comprising many is somewhat arbitrary. Generally speaking, polysaccharides are characterised by large numbers of monosaccharides (typically, 30 or more) linked together in a fairly uniform, repeating fashion to form long chains in much the same way as are other organic polymers; oligosaccharides, meanwhile, possess singular structures that are often characterised by complex branching patterns.

Whilst the deduced formula of simple sugars is usually $C_n(H_2O)_x$ the structural reality is, of course, that such compounds do not contain molecular water per se, but rather consist of multiple, chiral hydroxymethylene (CHOH) units sequentially linked to form a carbon chain that is commonly terminated at one end by a hydroxymethyl group and, at the other, by either an aldehyde or an α-hydroxyketone group (hence the terminology aldose and ketose; the chemical notations are $CH_2(OH)(CHOH)_{n-2}CHO$ and $CH_2(OH)(CHOH)_{n-3}COCH_2OH$, respectively). As such, most monosaccharides are reducing sugars since they exhibit aldehyde-like chemistry and are readily oxidised, for example by Tollen's and by Benedict's reagents to produce characteristic positive tests. The end of the molecule bearing the carbonyl moiety is termed the reducing terminus—the carbon chain is always numbered from this point—and the other end is termed non-reducing. However, in an aqueous environment the common reducing sugars are in dynamic structural equilibrium, existing only fleetingly as open-chain molecules, and instead spending a much larger proportion of their time as cyclic hemiacetals (see Fig. 2A, overleaf). Indeed, such cyclisation explains why there are two separable, crystalline stereoisomers of a given monosaccharide—the α and β anomers—because this process of ring-formation introduces an additional stereocentre at the site of the carbonyl carbon (known as the anomeric centre; C-1 for aldoses, C-2 for ketoses). These anomers are discrete from one another in the solid phase, but as soon as either one is dissolved in a nucleophilic solvent such as water, they may undergo interconversion via the open-chain form, a process that leads to an inevitable loss of anomeric purity. This equilibration is evidenced by the phenomenon of mutarotation, first observed by Dubrunfaut in 1846, whereby

the specific optical rotation of the solution is observed to change over time once an anomerically-pure reducing sugar is initially dissolved (see Fig. 2C). From the point of view of the chemist, simple monosaccharides are therefore classifiable by three criteria: (i) the number of carbon atoms present in the molecule, excluding side-groups, (ii) whether the carbonyl group in the open-chain form is an aldehyde or ketone, and (iii) whether the cyclic hemiacetal form is five- or six-membered (furanose and pyranose forms, respectively). The vast majority of sugars encountered in Nature, as we shall see, are either five-carbon pentoses or six-carbon hexoses that differ from one another only in the stereoconfiguration at each of the chiral centres. In stereochemical terms, saccharides differing at only a single chiral centre, other than the anomeric centre, are epimers, whilst those differing at all stereocentres are related to the configuration of glyceraldehyde, and to each other, by the rather less useful, though historically entrenched, prefixes D and L.[16] Typically, a saccharide is defined as belonging to the D family when the highest-numbered stereocentre, or 'reference carbon atom,' (C-4 for pentoses, C-5 for hexoses) has the same configuration as D-glyceraldehyde, whereas compounds with the opposite configuration at this position (i.e. that of L-glyceraldehyde) belong to the L family (see Fig. 2B). In addition, the term anomer is reserved to describe saccharides that vary only in the configuration of their anomeric centres.

During the first decade of the twentieth century, C. S. Hudson proposed a relationship between optical activity and the anomeric configuration of sugars such that 'in the D-series, the more dextrarotatory member of an α,β-pair of anomers is to be named α-D-, the other being named β-D-' whereas 'in the L-series, the more laevorotatory member of such a pair is given the name α-L- and the other β-L-.'[15] Furthermore, with respect to Fischer's representation of simple sugars and their methyl glycosides, 'the —OH or —OCH_3 group on C-1 is on the right in an α-D-anomer and on the left in a β-D-anomer.'[15] Hudson's rule, as this came to be known, is remarkably accurate, although it is violated by some anomeric pairs, and so anomers are now more strictly defined according to absolute structural relationships. Thus, the α anomer is always the compound in which the substituents attached to the anomeric and reference carbon atoms have a formal *cis* relationship whereas the β anomer has the same groups *trans* to one another.

The presence in carbohydrates of several nucleophilic hydroxyl groups and an electrophilic carbonyl centre renders them ideal molecules for polymerisation, in much the same way that amino acids are able to link together via peptide bonds to form proteins, and nucleotides through phosphodiester bonds to produce RNA and DNA. Coupling between two monosaccharides occurs by chemical condensation of the anomeric functional group of the glycosyl donor, on the one hand, and a hydroxyl

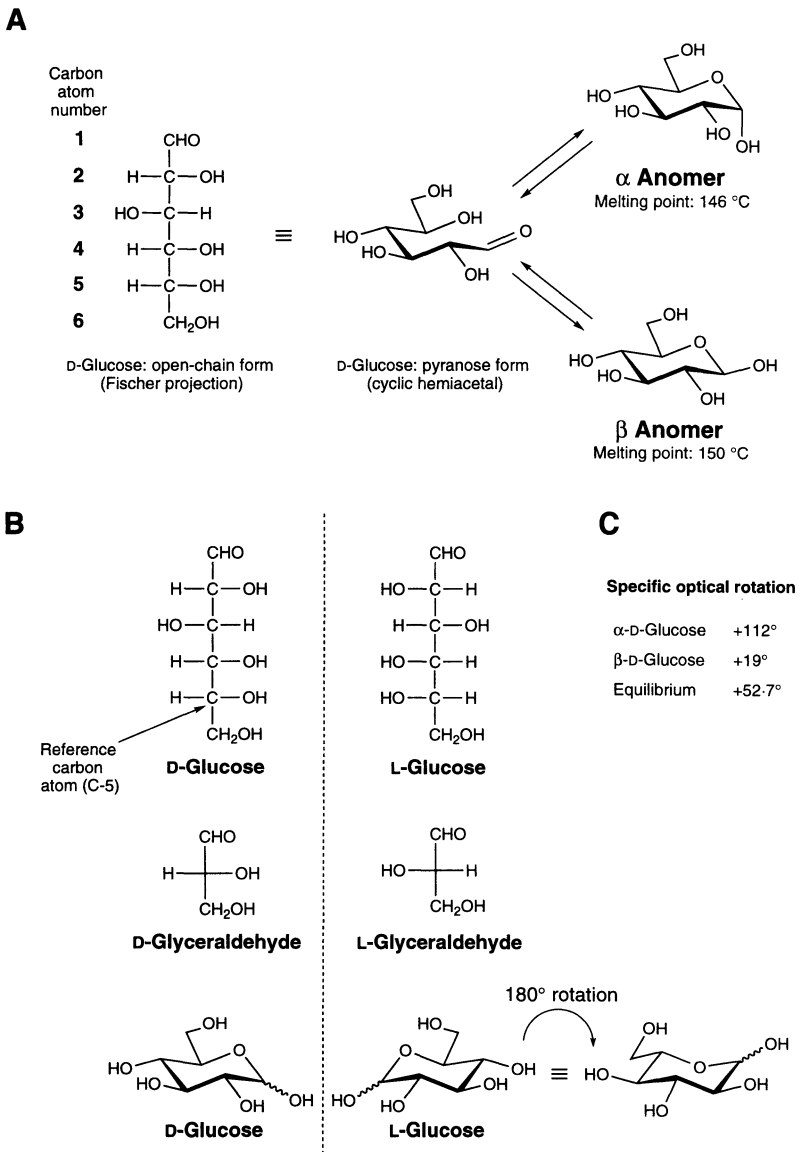
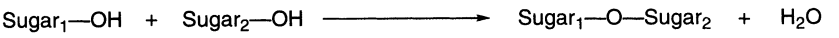


Figure 2. Fundamental properties, structural interrelationships and nomenclatures of carbohydrates as illustrated by the simple monosaccharide, glucose. (**A**) From left to right: carbon-atom numbering system, Fischer projection, and cyclisation of the open-chain form to produce α and β anomers. (**B**) Structural interrelationships defined by the D- and L- nomenclature of Rosanoff, in which carbohydrates are classified by configurational comparison of the reference carbon atom to glyceraldehyde.[16] (**C**) The phenomenon of mutarotation as exemplified by comparison of the specific optical rotation of aqueous D-glucose following initial dissolution of the α anomer and of the β anomer, and at equilibrium.

group of the glycosyl acceptor on the other, resulting in the formation of a glycoside (see Fig. 3A, top right). This reaction seals the anomeric fate of the glycosyl donor, since it has become a cyclic acetal that is resistant to ring-opening save by treatment with strong aqueous acid. Generic glycosidic bonds, one for each anomeric configuration of the glycosyl donor, are illustrated in chemical form in Figure 3B together with their shorthand nomenclatures. However, such linkages do not occur spontaneously in Nature, requiring instead the participation of activated glycosyl donors, in the shape of nucleoside phosphate sugars, and of highly specialised enzymes in order to surmount the considerable kinetic and thermodynamic barriers to bond formation. Of course, chemists have been able to recreate glycosides in the laboratory since the outset of this century,[17] but the inherent polyvalency of the constituent saccharides has required the development of an array of protecting groups designed to mask carbohydrate hydroxyls which would otherwise interfere with the desired coupling reaction and thereby yield a mixture of products. Furthermore, great care must be taken to ensure that glycosides are preferentially formed in the desired anomeric configuration, since the electronic effect of lone-pair electrons residing on the ring oxygen of the donor, adjacent to the anomeric centre, has a significant influence on the relative favourability of forming either the α or β anomers (this phenomenon is termed the 'anomeric effect'). Nevertheless, the synthesis of most types of glycosidic linkage with stereochemical precision is now fairly routine, although still far from trivial in the case of the larger oligosaccharides.[18-20] As will become apparent later (in Chapter 9), the laboratory preparation of synthetic oligosaccharides of adequate size and complexity for mimicking their natural counterparts remains a highly labour-intensive task, and has been the subject of considerable effort by carbohydrate chemists.

The second feature of carbohydrates that is implicit to the existence of protein glycosylation is their ability to form glycosidic linkages not only to other sugars, but also to certain amino acid residues of polypeptides. This type of attachment requires the presence of a suitable nucleophilic moiety in the amino acid residue destined to be the glycosyl acceptor, such as an amino, amido or hydroxyl group, which is capable of condensing with the glycosyl donor. The most common examples of protein glycosylation, at least in eukaryotes, originate from glycosidic linkages between the C-1 position of the carbohydrate's reducing-terminus and either the amido group of asparagine or the hydroxyl group of an hydroxyamino acid such as serine or threonine. Since these two types of glycosylation occur via covalent bonds to nitrogen and oxygen, respectively, the carbohydrate moieties in each case are broadly classified as either '*N*-linked' or '*O*-linked' oligosaccharides based on this distinction (see Fig. 4, bottom right).

A



B

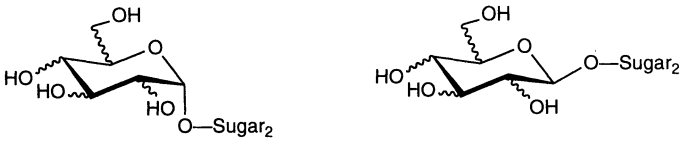
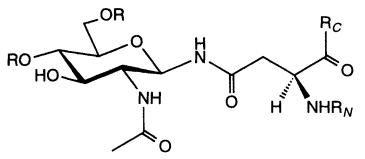


Figure 3. The formation of a glycosidic linkage between two sugars. (**A**) The generic reaction of glycosidic bond formation involves a condensation between two opposing hydroxyl groups of the participating sugars; $Sugar_1$ and $Sugar_2$ may be mono- or oligosaccharides. (**B**) The generic chemical structures of an α- an β-glycoside of a D-hexanopyranose, such as D-glucose.

A



B

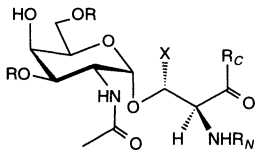


Figure 4. The two major types of protein–carbohydrate linkage found in eukaryotic glycoproteins. (**A**) Asparagine-linked glycosylation via *N*-acetylglucosamine; (**B**) serine-linked (X = H) and threonine-linked (X = Me) glycosylation via *N*-acetylgalactosamine. In each case, R represents an oligosaccharide side-chain; R_N and R_C represent extension of the poly-peptide chain in the *N*-terminal and *C*-terminal directions, respectively.

The inherent capacity of monosaccharides to link together via glycosidic bonds has provided the fundamental, molecular means by which living organisms have evolved to exploit carbohydrates for mediating sophisticated biological tasks. The potential multiplicity and diversity of oligosaccharides is enormous. Theoretically, at least, the sheer number of structural permutations open to carbohydrates dwarfs those of the other biopolymers, even in the case of the smallest oligomeric molecules (see Table 1, below).

Table 1. Isomeric possibilities for comparable sequences of oligopeptides and oligosaccharides. After Schmidt.[19]

Oligomer:	*Composition:*	*Oligopeptide:*	*Oligosaccharide:*
Dimer	AA/AB	1/2	11/20
Trimer	AAA/ABC	1/6	120/720
Tetramer	AAAA/ABCD	1/24	1424/34560
Pentamer	AAAAA/ABCDE	1/120	17872/2144640

One of the reasons for this is that sugars have a much larger palette of different building blocks to draw upon than either nucleic acids or proteins. Another is the possibility of a variety of different attachment points between two monosaccharides. This, in turn, leads to the unique branching capability of carbohydrates apparent in so many biologically-relevant oligosaccharides. Yet another is the option of either α or β anomeric configurations for each of the glycosidic linkages. Rather fortunately for the research scientist, the reality of the matter is that only a small handful of the possible different monosaccharides are commonly utilised in Nature. Figure 5, right, summarises the repertoire of monosaccharides available to vertebrates, including their structures, trivial names, and routine shorthand notations, as well as a selection of others found in non-vertebrate organisms.

Also fortuitous is the finding that the structural variety of oligosaccharides encountered in glycobiology, at least as far as eukaryotic species are concerned, is surprisingly constrained. Only certain branching patterns are 'allowed' by the cell's biosynthetic machinery and only the peripheral sugar moieties of oligosaccharides tend to differ significantly from one molecule to the next, or among different tissues, or even between disparate members of the same phyla. Thus, when comparing the structural relationships of different carbohydrates falling into the same broad category, the invariant regions are commonly referred to as the 'core,' whilst the variable regions which extend away from this core are called 'antennae.'[21] This simple terminology is illustrated in the case of two biological

oligosaccharides that are representative of the *N*-linked and *O*-linked classifications, as shown in Figure 6 (overleaf).

As we shall discover later, in Chapter 3, the martialling of the accumulated structural data on various carbohydrate moieties commonly found in attachment to glycoproteins has led to a more advanced taxonomy of oligosaccharides. Nonetheless, despite the patient efforts of many scientists to develop a roster of important structural motifs, it remains clear that the interlinking of a limited repertoire of simple monosaccharide building blocks leads to a very wide spectrum of biological oligosaccharides

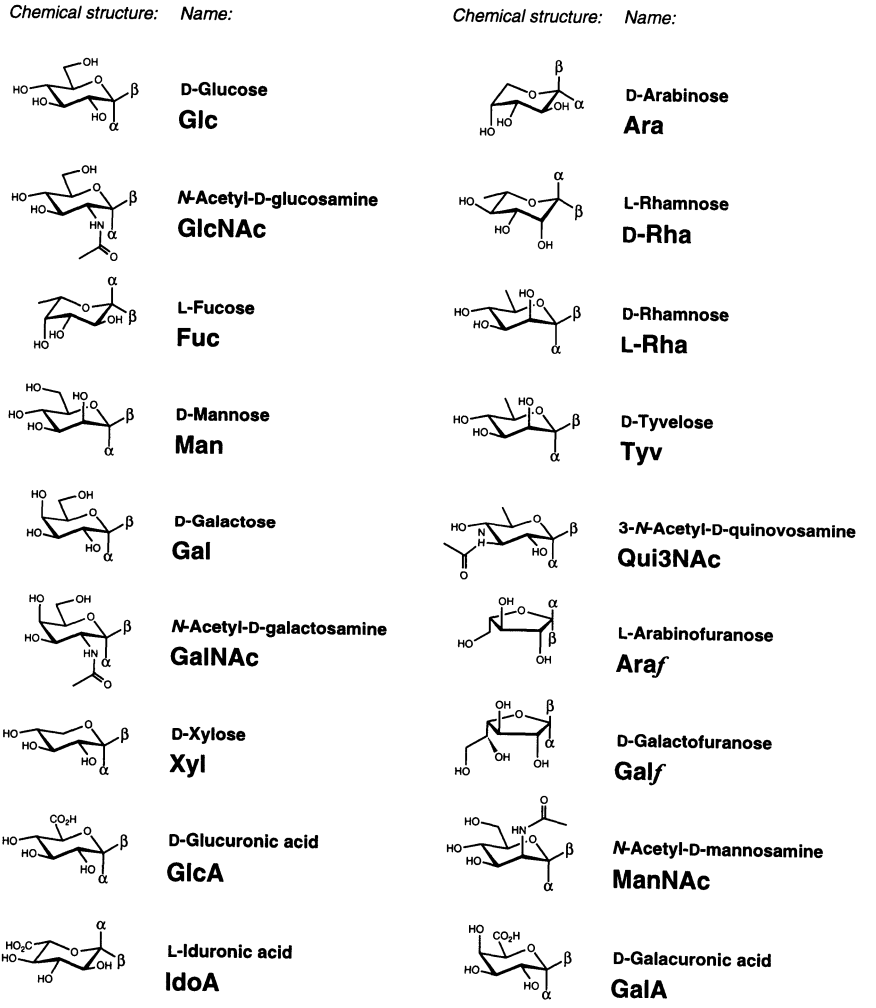


Figure 5. The repertoire of monosaccharides found comprising the glycoconjugates of living organisms.

that will continue to present a formidable challenge to our understanding of the important processes in which they play a role.

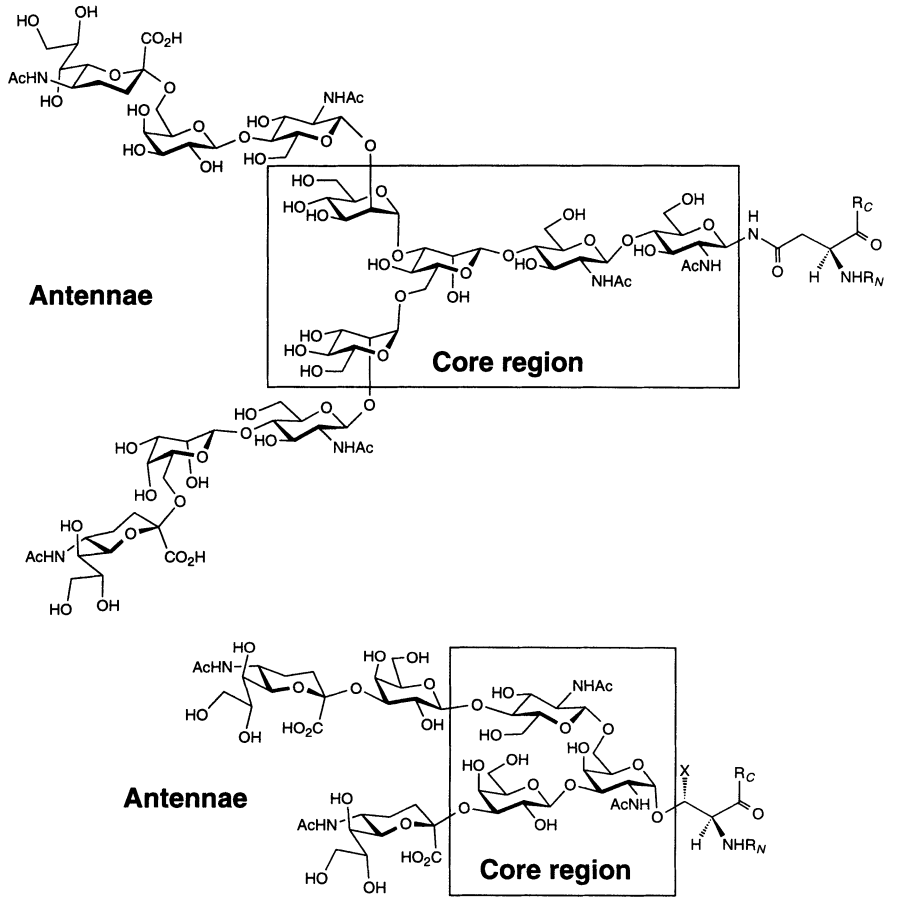


Figure 6. Definition of the 'core' and antennae' regions of typical *N*- and *O*-linked oligosaccharides. The boxed regions are the 'trimannosyl core', $Man_3GlcNAc_2$, of *N*-linked oligosaccharides, and the *O*-linked 'core 2' structure, respectively.

2. NOMENCLATURE: SUGAR-CHAIN SHORTHAND[22,23]

It is appropriate at this point for us to briefly define the carbohydrate terminology and shorthand nomenclature that will be used consistently throughout this book to delineate the various oligosaccharide structures

under discussion. A clear definition of the term 'glycoprotein' was established shortly after the turn of the century[24] by the Committee on Protein Nomenclature of the American Society for Biochemistry, which described this class of biological molecules as 'compounds of the protein molecules with a substance or substances containing a carbohydrate group, other than nucleic acid.' and has since been authoritatively reiterated by Robert Jeanloz[25] to mean 'conjugated proteins containing a prosthetic group of one or more heterosaccharides with a relatively low number of sugar residues, lacking a serially repeating unit and bound covalently to the polypeptide chain.' The subsequent coinage of the umbrella term 'glycoconjugates' to encompass both glycoproteins and glycolipids,[5] has led Jean Montreuil to propose a hierarchical classification scheme, an adapted version of which is presented in Figure 7, below.[26]

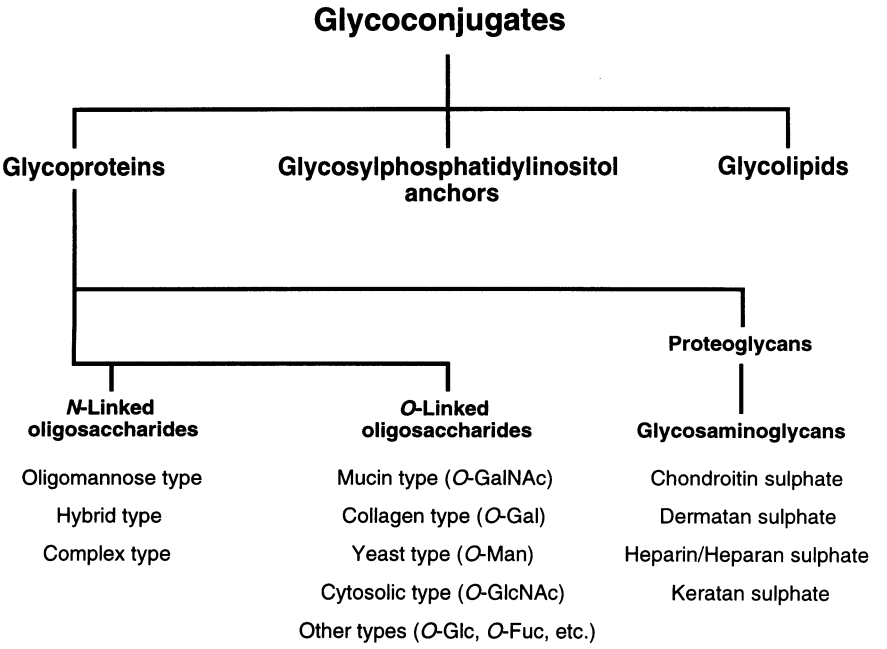


Figure 7. The proposed 'family tree' of glycoconjugates. After Montreuil.[26]

Whilst the use of such terminology is universal, as indeed are the monosaccharide abbreviations listed previously in Figure 5, a cursory inspection of the literature will reveal that the same is not true for the depiction oligosaccharide structures in shorthand notation. Table 2, overleaf, illustrates some of the different ways of representing the simple

disaccharide lactose, which comprises the β anomer of D-galactose and an unspecified anomer of D-glucose coupled together by a single glycosidic bond between the C-1′ and C-4 positions—this is usually referred to in abbreviated form as a 'β1,4-linkage' where the two locants are always separated by a comma and the locant of the reducing-terminal carbohydrate component is on the right.

Table 2. Some common methods of representing the structures of oligosaccharides using shorthand notation, as exemplified for the disaccharide, lactose.

Method 1	β-D-Gal*p*-(1→4)-D-Glc*p*
Method 2	Gal(β1–4)Glc
Method 3	Galβ1→4Glc
Method 4	Gal-β1,4-Glc
Method 5	Galβ4Glc

For clarity, we have chosen Method 2 as the preferred structural shorthand, whereby lactose is written as Gal(β1–4)Glc to depict the requisite β1,4-linkage. Carbohydrate–protein linkages are treated similarly, as in the case of GlcNAc(β1–*N*)Asn to represent the covalent attachment of *N*-acetyl-β-glucosamine to the amido group of asparagine that is common to *N*-linked oligosaccharides. In more complicated structures, square brackets are used to denote oligosaccharide branching, so that for example the representation of the Lewis^a blood-group antigen as Gal(β1–3)[Fuc(α1–4)]GlcNAc indicates the attachment of both galactose and fucose to the inner GlcNAc residue. Alternative linkages and optional modifications are implied respectively by the presence of multiple reducing-terminal locants separated by a slash and by a plus-or-minus sign; hence, ±Sia(α2–3/6)Gal(β1–4)-GlcNAc refers to the three distinct oligosaccharide structures in which the terminal sialic acid (Sia) is present in either α2,3- or α2,6-linkages to the inner galactose, or not at all. Where repetitive motifs are involved, curly brackets are used to indicate the repeating unit, as in Sia(α2–3){Gal(β1–4)-GlcNAc(β1–3)}$_n$Gal(β1–3)GalNAc(α1–*O*)Ser/Thr. Frequently in cases where structural information is lacking, or where complex biosynthetic intermediates are inferred, simpler notations are used, such as Glc(1–*N*)Asn to depict an unspecified glycosidic linkage, or $Man_5GlcNAc_2$ to represent Man(α1–6)[Man(α1–3)]Man(α1–6)[Man(α1–3)]Man(β1–4)GlcNAc(β1–4)-GlcNAc(β1–*N*)Asn, a common heptasaccharide intermediary structure used in eukaryotic *N*-glycosylation. Modifications such as methylation, sulphation and phosphorylation are normally shown in square brackets

preceding the relevant monosaccharide, unless they are at the non-reducing terminus; hence, Me-*O*-3Man(α1–6)[Me-*O*-3]Man(α1–2) refers to 3-*O*-methylation of both component mannose residues, whereas SO_4-6GalNAc-(β1–4)GlcNAc indicates sulphation at the C-6′ position of the disaccharide unit. Finally, in some instances, particularly during discussions of substrate specificity, chemically-correct names are substituted for readability, as in the use of α-2′-fucosyllactose to mean Fuc(α1–2)Gal(β1–4)GlcNAc.

A brief synopsis of enzyme nomenclature is given in Chapter 3; however, to avoid confusion, when the names of each of the major biosynthetic enzymes are first encountered in Chapters 4 to 7, they will be followed by the official Enzyme Commission (EC) designation of the International Union of Biochemistry, given in square brackets.[27*] Note that the majority of glycosyltransferases, the enzymes which form the bulk of the subject matter in this book, are hexosyltransferases and are therefore designated EC 2.4.1.*x*.

3. A BRIEF HISTORY OF GLYCOSCIENCE[28,29]

The emergence of glycoscience as a recognised scientific field has its historical foundations in the carbohydrate chemistry and biochemistry of the mid-to-late nineteenth and early twentieth centuries with the work of such celebrated scientists as Gerardus Mulder, Justus von Liebig (inventor of the eponymous condenser), Eduard Buchner (of funnel fame), Claude Bernard (discoverer of glycogen), and Felix Hoppe-Seyler, among others.[30] The earliest members of the glycoprotein family to be identified as a distinct class of biological compounds were mucins (so-named in 1835 by Nicholas Theodore de Saussure), several of which were isolated in the second half of the nineteenth century (see Table 3, overleaf).

During this period, perhaps the most important studies were those of Olof Hammarsten, a Swede working at the University of Uppsala who enlarged upon the earlier work of Eichwald and Hoppe-Seyler. Eichwald, a Russian clinician working in Scherer's laboratory in Würzburg had already conducted experiments leading him to make the groundbreaking claim that mucins were 'compounds consisting of a moiety with all the properties of a genuine protein and a moiety released under certain conditions as a sugar;' and Hoppe-Seyler in his influential textbook of 1881, *Physiologische Chemie*, had also prescribed that mucins—such as are found in the spittle of the Asian swiftlet, which acts as a nesting cement and which, by the same

* The reader should be beware that this publication from 1992 contains a number of errors. Therefore, in the present work, EC numbers are only given if they have been assigned correctly.

token, is a component of the Chinese delicacy, bird's nest soup—when treated with dilute sulphuric acid, precipitated 'clots' that exhibited the properties of acidic carbohydrates.[40] Hammarsten went further, demonstrating in a very elegant paper that sugars were an integral part of these mucin glycoproteins.[36]

Table 3. Timeline for the isolation of the first mucins. After Gottschalk.[39]

Mucin:	*Author:*	*Year:*	*Reference:*
Mediastinal cyst mucin	J. J. Scherer	1846	31
Helix pomatia mucin	E. Eichwald	1865	32
Submaxillary gland	S. Obolenski	1871	33
Frog egg envelope	P. Giacosa	1882	34
Frog oviduct mucin	P. Giacosa	1882	34
Helix pomatia mucin	O. Hammarsten	1885	35
Bovine submaxillary mucin	O. Hammarsten	1888	36
Seromucoid	C. U. Zanetti	1897	37
Bronchial mucin	F. Müller	1898	38

Meanwhile, much of the structural work on simple sugars was carried out predominantly under the auspices of the great Emil Fischer in Berlin (Nobel Laureate, 1902). In the same year that Hammarsten published his key paper, Fischer was making progress towards solving the configuration of D-glucose, building on the polyhydroxyl aldehyde formulation (see Section 1) proposed by Fittig and Baeyer. By an ingenious application of the theory of the asymmetrical carbon atom, first proposed by Van't Hoff and Le Bel in 1874, Fischer established the relative configurations of all four stereocentres of open-chain glucose and simultaneously determined the isomeric interrelationships between all the other hexoses and pentoses. The necessary proofs came in the form of syntheses between the various different monosaccharides using the chemical techniques for chain-lengthening and shortening developed by fellow countrymen, Heinrich Kiliani and Otto Ruff.[41] However, uncertainty over the ring sizes of pyranose and furanose sugars continued well into this century, exacerbated perhaps by the fact that Fischer (wrongly) championed the proposal that the cyclic form of glucose was five-membered. It was not until 1926 that the debate was finally laid to rest by the work of English chemist, Sir Norman Haworth, and his protegé Edmund Hirst, who used the celebrated technique of methylation analysis first introduced by Thomas Purdie and Sir James Irvine[42] to prove the now-familiar six-membered ring formulation for pyranoses.[43,44]

There followed a fruitful period for carbohydrate research which was highlighted by the correct structural determination of the amino sugars, D-glucosamine[45] and D-galactosamine,[46] and of various of the common disaccharides: sucrose, maltose, cellobiose and lactose. Not least, only two years after Hirst's proof, came the culmination of extensive studies by Phoebus Levene at the Rockefeller Institute in New York, identifying 2-deoxyribose as the carbohydrate component of DNA.[47] In collaboration with LaForge, Levene made another important contribution when he advanced the pioneering work of Schmeideberg[48] by finally identifying the carbohydrate components of the mucopolysaccharide (glycosaminoglycan), chondroitin sulphate, a substance first isolated by Mörner in 1889.[49] But although the study of carbohydrate metabolism was in its heyday with, for example, the elucidation of glycogen catabolism by such luminaries as Carl and Gerty Cori (joint Nobel Prize winners in 1947), progress in the sphere of glycoconjugates remained sluggish. The first major step towards an understanding of glycoproteins was taken by the young Albert Neuberger who, in his landmark paper of 1938, concluded that sugar chains were also integral to proteins other than mucins, in this case ovalbumin.[50,51] However, by the following year Europe was consumed by World War, delaying any further advances by Neuberger; a fuller investigation would have to wait.

Progress in the arena of glycobiology was also significantly hampered for almost a quarter of the century by the continued absence of a structure for the neuraminic, or sialic, acids.[30] As early as 1936, this important family of acidic carbohydrates, now known to be ubiquitous in animal glycoconjugates, had been isolated in Uppsala by Gunnar Blix[52] from bovine submaxillary mucin (who called it sialic acid, after the Greek *sialos* for saliva[53]). Shortly afterwards, in Cologne, Ernst Klenk obtained the same compound from the glycolipids of neural tissue (hence his name for the substance, neuraminic acid).[54] However, it was not until the turn of the Sixties, after considerable efforts on the parts of several outstanding researchers, most notably Alfred Gottschalk, Sir John W. Cornforth,[55] Donald Comb, Saul Roseman[56] and Richard Kuhn, that the basic structure was fully elucidated.[39] The sialic acids are now known to comprise over forty related compounds, the most common of which are listed in Table 4; the shared core structure of sialic acids is shown in Figure 8 (both overleaf).

Towards the end of the 1950s, the covalent attachment of carbohydrate chains to protein was finally put beyond doubt by Helen Muir, working in London,[3] and during the early 1960s by the Uppsala group headed by Ulf Lindahl and Lennart Rodén,[58,59] when they independently demonstrated the base-catalysed release, or so-called β-elimination (see below), of *O*-linked saccharides from the serine residues of proteoglycans. Around the same time, Neuberger returned to the fray and, together with his colleague Robin

Marshall at St. Mary's Hospital in London, he published the discovery of analogous carbohydrate linkages to asparagine: the *N*-linked saccharides.[60] Before long, mucins were shown to possess sugars attached abundantly to both serine and threonine residues, as demonstrated for bovine submaxillary mucin by Ward Pigman's group;[61] and other, less common types of linkages were also identified in collagen and in plant tissue. All of a sudden, it

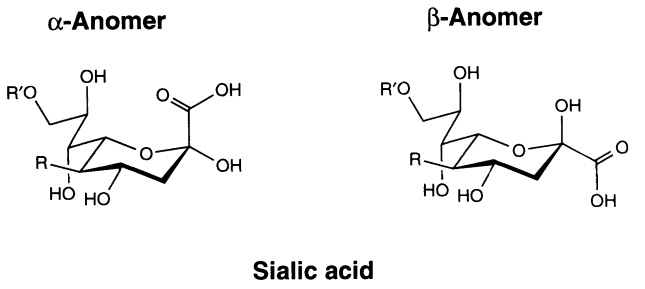


Figure 8. The generic structure of both anomers of sialic acids. The substituents R and R′ vary as indicated in Table 4, below.

Table 4. The nomenclature and structure of selected sialic acids and related sugars. In accordance with the Rules for Carbohydrate Nomenclature recommended by the IUPAC-IUBMB Joint Commision on Biochemical Nomenclature (1995).[57]

Trivial name:	*Full chemical name:*	*Abbreviation:*	*Side chains:*
N-Acetylneuraminic acid	5-Acetamido-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulopyranosonic acid	**Neu5Ac**	R = $NHCOCH_3$ R′ = H
N-Glycolylneuraminic acid	5-Hydroxyacetamido-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulopyranosonic acid	**Neu5Gc**	R = $NHCOCH_2OH$ R′ = H
9-*O*-Acetyl-*N*-acetyl-neuraminic acid	5-Acetamido-9-acetyl-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulopyranosonic acid	**Neu5,9Ac$_2$**	R = $NHCOCH_3$ R′ = $COCH_3$
2-Keto-3-deoxynononic acid	3-Deoxy-D-*glycero*-D-*galacto*-non-2-ulopyranosonic acid	**Kdn**	R = OH R′ = H

seemed, carbohydrates were everywhere: decorating the surface of cells or finding attachment to numerous secreted proteins, most notably in the blood. A selection of proteins that have since turned out to be glycosylated are listed in Table 5 (below and overleaf); the various different categories of linkage that are presently known to occur between carbohydrates and proteins are summarised in Table 6 (overleaf). However, it is important to point out that significant new forms of protein glycosylation were detected for the first time as recently as 1988,[70] and still further varieties have been reported in the last couple of years.[71,72] With this in mind, it seems likely that several other types of linkage may remain to be found in the future.

To begin to understand the significance of glycosylation in biology, we must return to pre-War North America. By 1941, the physiological and medical implications of these newly-uncovered sugar side-chains of proteins and lipids were just beginning to emerge. It was in that year that George Hirst, a clinician working at the Rockefeller Foundation in New York,[73] and two Canadian researchers, Laurella McClelland and Ronald Hare,[74] independently noticed that by treating chicken erythrocytes with infective influenza virus in the cold they could provoke rapid haemagluttination: a phenomenon which ultimately turned out to involve glycosylation. Furthermore, this process was reversed, albeit at a much slower rate, when the temperature was raised to 37° C, but only if live, pathogenic virus was

Table 5. Specific examples of glycoproteins present in animals and plants.

Glycoprotein:	*Present in:*	*Size (M_r):*	*Percentage sugar content:*
Enzymes			
Alkaline phosphatase	Murine liver	130,000	18
Bromelain	Pineapple	33,000	36
Carboxypeptidase Y	Yeast	51,000	17
Hormones and cytokines			
Chorionic gonadotropin	Human urine	38,000	31
Erythropoietin	Human urine	34,000	29
Interferon-γ	Human leukocytes	26,000	20
Lectins			
Potato lectin		50,000	50
Soybean lectin		120,000	6

This table is continued overleaf.

Table 5 (Continued). Specific examples of glycoproteins present in animals and plants.

Glycoprotein:	*Present in:*	*Size (M_r):*	*Percentage sugar content:*
Cell-surface and viral coat glycoproteins			
Glycophorin A	Human erythrocytes	31,000	60
Haemagglutinin	Influenza virus	210,000	25
Rhodopsin	Bovine retina	40,000	7
Serum glycoproteins			
Immunoglobulin G	Human serum	150,000	10
Thyroglobulin	Calf thyroid	670,000	8
Prothrombin	Human serum	72,000	8
Structural glycoproteins			
Collagen	Rat skin	300,000	0·4

Table 6. Types of eukaryotic protein glycosylation involving different sugar–protein linkages.

Linkage:	*Commonly found in:*	*Year:*	*Reference:*
GlcNAc(β1–*N*)Asn	Animal glycoproteins	1964	60
GalNAc(α1–*O*)Ser/Thr	Mucins	1964	61
Xyl(β1–*O*)Ser	Proteoglycans	1964	62
Gal(β1–*O*)Hyl	Collagens	1966	63
Araf(β1–*O*)Hyp	Plant and algal glycoproteins	1967	64
Man(α1–*O*)Ser/Thr	Yeast cell-wall glycoproteins	1968	65
Gal(α1–*O*)Ser	Plant extensins	1973	66
Fuc(α1–*O*)Ser/Thr	Serum clotting factors	1975	67
GlcNAc(β1–*O*)Ser/Thr	Cytoskeletal and nuclear proteins	1984	68
Glc(β1–*O*)Ser	Serum clotting factors	1988	69
Glc(α1–*O*)Tyr	Glycogen	1988	70
GlcNAc(α1–*O*)Ser/Thr	Mucins of *Trypanosoma cruzi*	1994	71
Glc(β1–*N*)Arg	Starch	1995	72

used during the experiment; in contrast, the haemagglutination caused by 'flu samples that had been previously heat-inactivated at 60° C proved to be essentially irreversible.[75] In fact, Hirst showed that following this initial adsorption-elution cycle of the live influenza, the eluted virus could be harvested and re-administered to successive batches of red cells without appreciable loss of haemagglutinating activity. The recycled erythrocytes, on the other hand, would no longer agglutinate even when exposed to fresh virus.

This phenomenon, which as we shall see was the first to truly stimulate interest in the biology of glycoconjugates, was further studied extensively by Sir F. Macfarlane Burnet in Melbourne, Australia. Burnet and his group memorably reported the presence in *Vibrio cholerae* culture filtrates of an enzyme which he coined Receptor-Destroying Enzyme (or RDE), owing to its ability to render red cells inagglutinable by the 'flu virus.[76] Moreover, a variety of human and animal mucins[77] rich in sialic acid, such as the Francis inhibitor in normal serum,[78,79] were already known to block haemagglutination by heat-inactivated (but not live) virus; however, when these substances were pretreated with RDE, or with the live virus, they irrevokably lost their ability as inhibitors.[80,81] Hirst had previously proposed that whilst the agglutination process itself was brought about by a viral haemagglutinin, the apparently inexhaustible capacity of the live influenza virus subsequently to elute from red cells under physiological conditions was characteristic of a viral enzyme, presumably acting on a receptor embedded in the erythrocyte cell membrane. Burnet went on to support this view, proposing that the red-cell receptors and the inhibitory mucins were structurally related because they were all substrates for RDE and for Hirst's proposed viral enzyme; conversely both the bacterial and viral enzymes must belong to the same family because of their similar activities. Since this inactivation of the mucins by RDE was always accompanied by the release of *N*-acetylneuraminic acid, Gottschalk later termed the enzyme neuraminidase [EC 3.2.1.18].[82,83] It turns out that the receptor itself is predominantly glycophorin A (so-named by Vincent Marchesi[84]), a sialylated glycoprotein that occurs abundantly at the surface of red blood cells, and is recognised by the viral haemagglutinin (the latter is therefore classifiable as a lectin, see below).[85] In more recent years, a range of viruses other than influenza have been shown to target variously-substituted sialic acid residues attached to glycoconjugates on the host cell-surface: these include Sendai virus,[86] human encephalomyelitis virus[87] and various others that are pathogenic in birds and domestic animals.[88,89]

Amongst the other early discoveries, of prime importance were those made by Walter Morgan and Winifred Watkins, working at the Lister Institute in London, and by Elvin Kabat of Columbia University in New

York. Culminating extensive studies, they pinpointed sugar molecules, attached to the cell-surface lipids of erythrocytes, as the root cause of the long-known histoincompatibility of different blood types, a fundamental physiological property that was first documented by Karl Landsteiner a half-century before and who is now illustriously commemorated on the Austrian one thousand schilling banknote.[90] In a series of key papers (for a review, see ref. 91) they published the chemical composition and structures corresponding to the three blood group antigens, A, B and H, which give rise to the four common phenotypes of the ABO system.[92] By 1968, detailed carbohydrate structures had also been attributed to the Lewis blood group determinants, Lewis^a (Le^a)[93] and Lewis^b (Le^b),[94] which are the basis for another method of blood typing, previously introduced by Arthur Mourant, an English clinician, and named by him after the blood donor expressing the original antibody against Le^a.[95] Furthermore, in certain cases, the ABO and Lewis antigens were found to be not simply confined to erythrocytes, but were also detected in saliva and other fluids,[96-98] leading to the distinction between individuals who are secretors and those who are non-secretors of these blood group substances.

During the 1980s, the genetic foundations for the ABO blood group system were subsequently revealed through the work of Henrik Clausen, Sen-itiroh Hakomori and their colleagues in Seattle.[99] Of particular note was their finding that the two glycosyltransferases responsible for synthesising the A and B antigens, enzymes which had long since been postulated by Watkins and Morgan,[100] and later characterised,[101-104] share almost identical gene sequences.[99] A difference of a mere four nucleotides between the *A* and *B* alleles, which lead to variations in the amino acid sequences at positions 176, 235, 266 and 268 of the corresponding enzymes, are, it emerged, all that distinguish the two blood-group glycosyltransferases, A and B; and only two of these positions (266 and 268) play a crucial role in determining their respective donor substrates: UDP-galactose and UDP-*N*-acetylgalactosamine.[105] Moreover, it has become apparent that the *O* genotype is characterised by genetic mutations in the *A/B* gene—three such defective alleles are currently known—that give rise to inactive gene products and thereby explain why O individuals do not make the A and B antigens. In the recent past, complementary investigations conducted principly by John Lowe and his co-workers in Michigan have further clarified the situation by successfully cloning genes for a series of fucosyltransferases that synthesise the fucose-dependent H and Lewis antigens;[106] one such enzyme is tied to the secretor status of the individual.[107] Several defective alleles of these genes which are responsible for the absence of H antigens in the so-called 'Bombay phenotype'[108-110] and of Lewis antigens in Lewis-negative individuals[111-113] have also been characterised.

This line of research has had an impressive impact on molecular medicine. Primarily, knowledge of the various blood group determinants, and of the genes that direct their biosynthesis, has enhanced our understanding of the molecular basis of blood typing, which is the key to safe transfusions. Secondly, and somewhat more provocatively, the reappearance, abnormal expression, or complete loss of the ABH and Lewis antigens has been observed in a variety of carcinomas, suggesting that these structures have a role to play in cancer;[114] indeed, screening for such changes would appear to have substantial prognostic value.[92,115] Finally, comes the realisation that nearly all mammals possess yet another type of blood-group-related antigen comprising a terminal galactobiose moiety, Gal(α1–3)Galβ. Strikingly, humans and their close relatives, such as apes and Old World monkeys, do not express this so-called 'straight-chain B' antigen and instead produce antibodies against it. A plausible explanation of this state of affairs is that it has resulted from some catastrophic event in primate evolution; for example, from a pandemic caused by microbes displaying the galactobiose antigen, which applied selective pressure for primates to lose their native ability to synthesise this structure, and hence allowed them to survive by raising the complementary antibodies. Corroboration comes from the finding that two inactive pseudogenes, that presumably once encoded the α1,3-galactosyltransferase responsible for generating this antigen, have been identified in the human genome by homology screening,[116,117] based on our knowledge of the DNA sequences of the bovine[118] and murine[119] counterparts. Exciting work is now in progress to engineer animals, pigs in particular, which will lack the straight-chain B antigen, thus facilitating transplantation of their organs into humans whilst avoiding the otherwise inevitable problem of xenorejection. So far, a number of strategies for achieving this goal have been adopted.[120,121]

With the benefit of hindsight, it is easy to question why, despite the clear implications of Morgan and Watkin's findings, and of Hirst's, such revelations failed to stimulate more widespread interest within the scientific community of the time. Unfortunately, these new insights came just as most researchers were turning their attention towards nucleic acids and the newly-emerging molecular biology. Then, in 1968, Gilbert Ashwell of the National Institutes of Health (NIH) in Bethesda and Anatol Morell of Albert Einstein College of Medicine in New York made a landmark discovery: one which was to become a focus of renewed interest in the biology of carbohydrates.[6,122] By enzymatically removing the sugar-chain-terminating sialic acid residues of [^{64}Cu]radiolabelled caeruloplasmin, a copper-binding serum glycoprotein, and then injecting the trimmed 'asialo'-form into laboratory rabbits, they found that the circulatory lifetime of this substance in the blood was abnormally and dramatically shortened. Closer inspection

revealed that the radiolabelled asialo-caeruloplasmin molecules were accumulating in the liver, where it became apparent there exists a class of specialised receptors which bind such aberrant molecules—in this case by recognising the exposed, underlying galactose residues—and filter them from the blood. Interestingly, a similar behaviour had been reported somewhat earlier by Bertram Gesner and Victor Ginsberg, not for glycoproteins, but for whole cells. They had found that the removal of cell-surface fucose from metabollically-labelled lymphocytes prior to transfusion into animals also led to their subsequently being diverted to the liver, rather than migrating to the spleen as expected, again through the operation of some form of hepatic receptor protein.[123] Such carbohydrate-binding proteins, collectively termed 'lectins' (from the Latin *legere*, to chose) by Boyd and Shapleigh in 1954,[124] and long-known to exist in plants, were of non-immunogenic origin and hence distinct from conventional antibodies raised against saccharides, although they could exhibit similar properties.[125] Indeed many of these plant lectins, termed phytohaemagglutinins in the older literature, were (and still are) invaluable tools for opening the 'black box' that is the cell's glycosylation machinery (see Table 7, opposite).

More recently, Samuel Barondes[126] defined lectins as 'carbohydrate-binding proteins that are neither enzymes nor antibodies.' Animal lectins are now known to play essential tasks in biological processes[127] and a growing number have since been found and classified; particularly notable in this area has been the work of Kurt Drickamer.[128-130] Four such groups have come to especial prominence (see Table 8, overleaf). The first comprises the P-type animal lectins that have been shown to function as the receptors for lysosomal enzymes, and mediate their deployment to the lysosomes;[131-136] the defective biosynthesis of their mannose 6-phosphate ligands is the cause of the rare inherited human disorder known as I-cell disease (see Chapter 7). In the second group are the calcium-dependent (C-type) lectins, of which perhaps the most interesting are those of the so-called 'selectin' family that recognise sialylated ligands such as the sialyl-Lewis antigens and initiate the binding of leukocytes to specific endothelial cells, hence playing essential roles in the normal trafficking of macrophages from the blood into the lymphatic system, and in the inflammatory response.[11,137-139] Since malignant cells are commonly observed to display cell-surface oligosaccharides with abnormally prolific sialic acid content, it is believed such cells may hijack these natural processes of extravasation and thereby provide an avenue for metastasis. Third are the S-type lectins, or 'galectins,' galactose-binding proteins that are implicated in tumour progression.[140,141] Most recently, a fourth group known as I-type lectins, or 'sialoadhesins,' has been reported; these are molecules featuring immunoglobulin-like domains

Table 7. Selected lectins commonly used as tools in the analysis of oligosaccharides.

Plant source:	*Acronym:*	*Monosaccharide inhibitor(s):*	*Oligosaccharide substrates:*
Arachis hypogaea (Peanut)	PNA	β-Gal	Terminal Gal(β1–3)GalNAc, T antigen
Canavalia ensiformis (Jack bean)	Con A	α-Man, α-Glc	Branched mannosides
Datura stramonium (Thorn apple)	DSA	β-GlcNAc	$\{GlcNAc(\beta1–4)\}_n$, Gal(β1–4)-GlcNAc (LacNAc) chains, i antigen
Dolichos biflorus (Horse gram)	DBA	α-GalNAc	Terminal α-GalNAc, Cad antigen
Glycine max (Soybean)	SBA	GalNAc	Terminal GalNAc
Helix pomatia (Roman snail)	HPA	α-GalNAc	Terminal α-GalNAc, A antigen
Lens culinaris (Lentil)	LCA	α-Man, α-Glc	Branched mannosides, esp. with core α1,6-Fuc
Maackia amurensis	MAA	Sialic acid	Terminal Sia(α2–3)Gal
Phaseolus vulgaris (Red kidney bean)	PHA-E	None	GlcNAc(β1–2)[GlcNAc(β1–4)]-Man(α1–6)
" "	PHA-L	None	GlcNAc(β1–2)[GlcNAc(β1–6)]-Man(α1–6)
Ricinus communis (Castor bean)	RCA_{120}	β-Gal	Terminal β-Gal
" "	RCA_{60}	GalNAc, β-Gal	Terminal GalNAc, β-Gal
Sambucus nigra (Elder)	SNA	Sialic acid	Terminal Sia(α2–6)Gal/GalNAc
Tetragonolobus purpurpeas/Lotus tetragonolobus (Asparagus pea)	LTA	α-Fuc	Terminal α-Fuc, H antigen, Lewis antigens
Triticum vulgaris (Wheat germ)	WGA	β-GlcNAc (Neu5Ac)	GlcNAc(β1–4)GlcNAc
Ulex europaeus (Gorse)	UEA-I	α-Fuc	Terminal Fuc(α1–2), H antigen

Table 8. Major families of animal lectins and their carbohydrate ligands

Type:	*Ligands:*	*Binding requirements:*	*Number of members:*
P-Type			
Lysosomal receptor	Mannose 6-phosphate	Calcium dependence*	2
C-Type	Variable	Calcium dependence (most)	> 20
Selectins	Sialylated and sulphated ligands	Calcium dependence	3
S-Type			
Galectins	Galactosyl ligands	Thiol dependence	~ 8
I-Type			
Sialoadhesins, Siglecs†	Sialylated ligands	None	> 5

* Two mannose 6-phosphate receptors are known: the cation-independent and the cation-dependent receptors. For further details see refs. 129 and 142.

† Siglecs are a very recently coined family of sialic-acid binding I-type lectins that comprise sialoadhesin (Siglec-1), CD22 (Siglec-2), CD33 (Siglec-3), myelin-associated glycoprotein, or MAG (Siglec-4a) and Schwann cell myelin protein, or SMP (Siglec-4b). See: Crocker PR, Clark EA, Filbin M, et al. Siglecs: A family of sialic-acid binding lectins. Glycobiology 1998; 8:v.

that are thought to be involved in biological processes as diverse as haematopoiesis, neuronal development and immunity.[142,143]

Ashwell and Morell's paper also served to inject renewed interest in some of the earlier studies highlighting the existence of lectins, other than in plants. In fact, the viral haemagglutinins responsible for Hirst's first observations had already been shown to be lectins by Gottschalk in 1960. A further revelation was the acknowledgement that bacteria were also exploiting lectins to recognise their target host cells and to initiate the infective process. Clear evidence for this came from the work of James P. Duguid at the Ninewells Hospital and Medical School in Dundee, who had reported during the 1950s that strains of *Escherichia coli* and other related bacteria could provoke haemagglutination, very much akin to the influenza virus.[144] However, unlike its virally-mediated counterpart, bacterial haemagglutination was found in many cases to be inhibited, not by sialylated mucins, but by the monosaccharide, mannose, and its derivatives.[145,146] Such

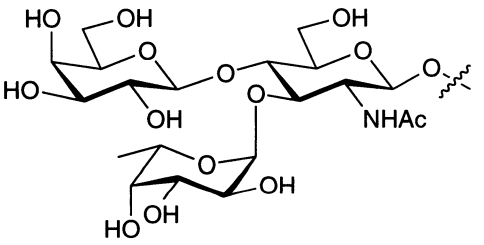
mannose-binding lectins, it transpired, are embedded in the tips of long, hairlike structures, variously called fimbriae, or pili, which protude from the bacterial cell wall.[147-149] Furthermore, in 90 % of the cases where bacterial haemagglutination is *not* prevented by mannose, it transpires that other carbohydrates such as galactobiosides (e.g. Gal(α1–4)Galβ) are effective inhibitors.[149] It has since been demonstrated that a broad range of microorganisms, including protozoa and some fungi, utilise carbohydrate–lectin interactions as a means of latching onto host cells and tissues, and it is now hoped that a fuller understanding of these processes will enable the development of prophylactic agents, such as monoclonal antibodies, or even simple sugars, that will block the route to infection by these types of pathogens.

Another key set of related findings that broke new ground was made by developmental biologists who were interested in using hybridomas to manufacture monoclonal antibodies that would recognise cell-surface antigens occurring during embryogenesis. The intent of these scientists was to highlight molecules with important roles in the early stages of development. Using this approach, Davor Solter and Barbara Knowles drew attention, in 1978, to an antigen appearing to proliferate during the 8-cell stage of embryonic development, which they termed Stage-Specific Embyonic Antigen-1 (SSEA-1);[150] unfortunately, they were unable to resolve the structure of SSEA-1. The subsequent developments in this field are largely attributable to Ten Feizi,[151,152] who demonstrated that carbohydrate structures, and not proteins, were responsible for the SSEA-1 antigenicity. In 1981, following careful examination of the specificity of the anti-SSEA-1 antibody, Feizi in collaboration with Solter and Knowles, published the structure of the antigen (see Fig. 9A, overleaf), which came to be known as Lewis^x (Le^x) because of its isomeric similarity to the Lewis^a determinant (cf. Fig. 9B).[153] In fact, Le^x structures had already been detected in human epithelia by both Kabat's[154] and Morgan's[155] groups during the late 1960s. The Le^x determinant was subsequently confirmed to be Gal(β1–4)-[Fuc(α1–3)]GlcNAc, and the difucosyl derivative, Fuc(α1–2)Gal(β1–4)-[Fuc(α1–3)]GlcNAc, which is analogous to the Lewis^b substance, was termed Lewis^y (Le^y).

Concurrent with Ashwell and Morell's work, Luis Leloir, who had just received the Nobel Prize (1970) for his discovery of the uridine diphosphate sugars,[156,157] began to make progress on the biosynthetic route to asparagine-linked glycoproteins in Buenos Aires: work which was to continue with major contributions from Stuart Kornfeld,[158-161] Frank Hemming,[162,163] Donald Summers,[164,165] Phillips Robbins,[166-169] Robert Spiro,[170-172] Harry Schachter,[173,174] and Saul Roseman, and culminating in Kornfeld's celebrated review of 1985.[175] Leloir made the important

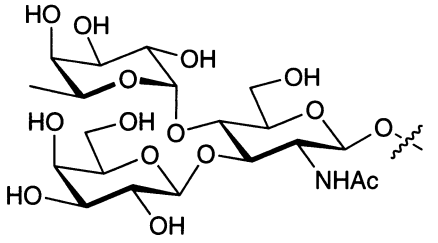
discovery that oligosaccharides destined for *N*-linkage to protein were in fact preassembled on the lipid precursor, dolichol,[176-179] but it was the elegant work of the above group of scientists that enabled us to understand how this precursor, which comprises predominantly mannose, is remodelled to give the sugar chains found on mammalian glycoproteins. In the autumn of 1977, at a touchstone conference held at Woods Hole, Massachusetts, many of the key discoveries in this area were disseminated by some of the field's most important pioneers.[180,181] Much of this work involved the systematic search for the activities of glycosyltransferases in tissue extracts. These enzymes were expected to exist because, logically, they ought to correspond to the biosynthesis of the known oligosaccharide structures; many of these had just been elucidated using NMR techniques through a highly influential collaboration between Johannes Vliegenthart in Utrecht and Jean Montreuil of Lille University.[180] The discovery of certain inhibitors, such as

A **Lewis^x**



Gal(β1–4)[Fuc(α1–3)]GlcNAcβ1-R

B **Lewis^a**



Gal(β1–3)[Fuc(α1–4)]GlcNAcβ1-R

Figure 9. The structures of the two isomeric trisaccharide carbohydrate antigens related to the Lewis blood-group system. (**A**) The Lewis^x blood-group-related antigen, also known as stage-specific embryonic antigen-1 (SSEA); (**B**) the classical Lewis^a blood-group antigen.

tunicamycin and swainsonine, which block specific steps during glycosylation, also provided valuable biosynthetic information.[182-185] Yet further clues came from the characterisation of various yeast[186-188] and, later, mammalian[189-192] cell mutants that possessed defects in the *N*-glycosylation pathway, and were therefore 'frozen' at certain points along the enzymatic assembly line. The fact that all the eukaryotes studied were found to biosynthesise *N*-linked oligosaccharides according to a common blueprint, at least in the early stages, provided clear evidence of an evolutionary link between species. However, important differences were also immediately apparent, with the pathways in unrelated species following widely divergent routes during the later stages of glycoprotein maturation. Yeast, it seemed, had chosen to elaborate its *N*-links prolifically, and almost exclusively, with large numbers of mannose residues,[188,193-195] whereas in man, for example, highly regulated mechanisms for the production of elaborate sugar 'antennae' were apparent.[196-199] Perhaps most importantly, the decoration of these antennae with fucose, galactose and sialic acid, the monosaccharide components of the ABO, Lewis, and sialyl-Lewis antigens, also appear to be a distinguishing trait in birds and mammals.

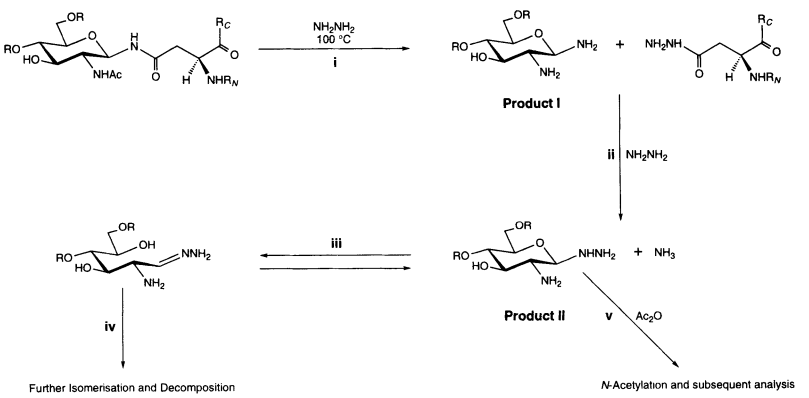


Figure 10. The chemical release of *N*-linked oligosaccharides from glycoproteins and glycopeptides via the method of hydrazinolysis. Reactions (i) and (ii) proceed quickly and reaction (iii) at a rather slower rate. Hydrazinolysis is normally halted at Product II by *N*-acetylation with acetic anhydride. For key, see legend to Fig. 4, above.

The study of the oligosaccharide profiles of glycoproteins was greatly enhanced by the introduction by Akira Kobata and his colleagues in Tokyo of routine hydrazinolysis[200] for the selective cleavage of *N*-glycosidic linkages (see Fig. 10, above), which had up to that point proven more difficult to analyse than their *O*-linked counterparts owing to their resistance to base-catalysed release. The widespread use of this technique,[201] alongside

the method of reductive β-elimination by treatment with alkaline borohydride (see Fig. 11, below),[202] has enabled the structural analysis of countless *N*- and *O*-linked oligosaccharides from glycoproteins. In addition, the discovery of a series of specific endoglycosidases and glycoamidases, the first of which was detected in extracts of *Streptomyces griseus* by Anthony Tarentino and Frank Maley at the New York State Department of Health,[203] has offered researchers alternative methods for releasing *N*-linked oligosaccharides from their parent glycoproteins, without recourse to drastic chemical treatment.[204]

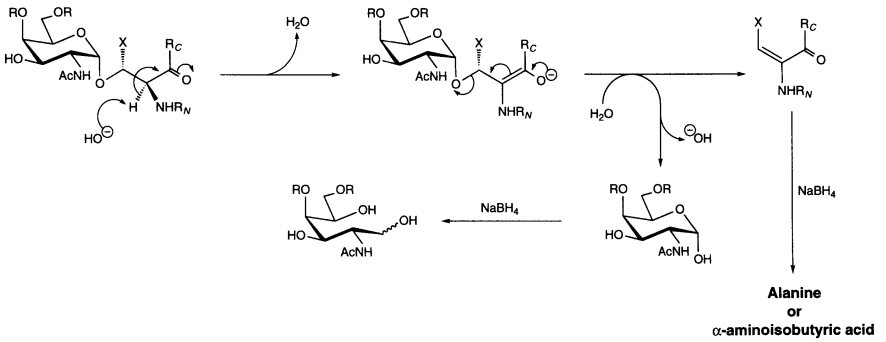


Figure 11. The mechanism of chemical release of *O*-linked oligosaccharides from glycoproteins and glycopeptides via the method of reductive β-elimination. Tritiated sodium borohydride is often used in this approach to facilitate the convenient radiolabelling and subsequent detection of the released oligosaccharides. For key, see legend to Fig. 4, above.

Towards the end of the 1960s, reports were beginning to appear in the literature indicating that the sugar-chains of glycoconjugates displayed at the surface of cancer cells were substantially different from normal.[205-208] One of the singular most important observations linking protein glycosylation with cancer was first made by Robbins' group in 1969.[209] In an effort to reveal the influence of malignancy on cell-surface oligosaccharides, they cultured mouse fibroblasts in the presence of radiolabelled monosaccharides, harvested the glycoproteins from the cell membranes, and digested them with proteases. When the resultant mixture of glycopeptides was separated by column chromatography, it became apparent that those isolated from malignant cell-lines generated by viral transformation had a significantly greater average size than those from normal lines. This switch to larger sugar chains by animal cells undergoing malignant transformation, exemplified in a wide range of cell-types and using different transformation techniques, became popularly known as the Warren-Glick phenomenon after the notable achievements of Leonard Warren and Mary Catherine

Glick.[210,211] This change was linked with the appearance at the cell surface of increasing numbers of branched oligosaccharides;[212-214] and this, in turn, with the hyperactivity of a particular glycosyltransferase, *N*-acetyl-glucosaminyltransferase V, that is responsible for attaching those branches.[215,214] The discovery during the last decade or so of readily-observable correlations between the increased branching of the Warren-Glick phenomenon and tumorgenicity[216] and metastatic potential,[217] most notably through the important work of James Dennis at the Mount Sinai Hospital in Toronto,[217-219] has further provided a strong motivation for studying the key glycosyltransferases which regulate glycoprotein biosynthesis. In particular, the importance of the transcriptional regulation of these enzymes in controlling the surface sugar profile of cells, and the implications for malignancy, have been borne out by the recent discoveries made by Naoyuki Taniguchi in Kyoto[220] and Michael Pierce in Athens, Georgia.[221,222]

Another landmark contribution to the fledgling glycopathology was made by Raj Parekh and his colleagues in Oxford, in collaboration with Kobata, when in 1985 they identified patients suffering from rheumatoid arthritis (RA) as having defective serum immunoglobulin G (IgG).[223] These tetrameric glycoprotein complexes, which play a major role in humoral immunity, were found to be abnormally glycosylated such that they predominantly lacked the galactose residues normally occurring in the oligosaccharide antennae attached to the heavy-chain subunit. A defect in the glycosylation capability of the IgG-producing B lymphocytes appears to be the cause of this galactose deficiency and work is now in progress to determine the precise role of agalactosyl-IgG in the pathogenesis of RA.[224]

One of the remaining challenges faced by glycoscientists since the turn of the 1970s has been the isolation of glycosyltransferases. Whilst glycosidases such as β-galactosidase were among the first enzymes to be obtained in abundance by heterologous expression of their genes in bacteria, glycosyl-transferases remained hopelessly aloof for some considerable time. In order to exploit the newly-emerging gene technologies for the purpose of elucidating structure/function relationships, a sufficient quantity of the desired protein was, and still is, required to determine at least a partial amino acid sequence, or to raise antibodies. Owing to their minute natural concentrations in tissues and body fluids, and to their tendency to be intimately associated with cellular membranes, the purification of the glycosyltransferases was to prove a truly mammoth task. Nevertheless, under the pioneering auspices of Robert Hill, working at the Duke University Medical Centre in North Carolina, the first such enzymes began to become available towards the end of the 1970s through the use of multiple rounds of column chromatography.[225] In 1981, Nunez and Barker reported

the first practical application of a glycosyltransferase for synthetic purposes: they used purified bovine β1,4-galactosyltransferase to prepare lactose.[226] This certainly paved the way for the practical application of these enzymes by chemists and biochemists, but the molecular biologists had yet to find a glycosyltransferase gene. Then at last, in 1986, partial β1,4-galactosyltransferase cDNAs were cloned by three groups, headed by Joel Shaper at the Johns Hopkins in Baltimore,[227] Pradman Qasba at the NIH[228] and Michiko Fukuda in La Jolla.[229] Since that time, a plethora of transferase genes or cDNAs has been obtained:[230,231] some, particularly those which attach the functionally-relevant terminal sugar residues to glycoproteins, such as sialic acid and fucose, have been cloned in a number of different species; others that transfer rare sugars or synthesise the more mundane core structures have been largely ignored. Various attempts to contrive the heterologous expression of glycosyltransferases for the purpose of obtaining large quantities of these enzymes have met with mixed success. Some techniques, especially those involving the production of recombinant enzymes with cleavable signal sequences and polyhistidine or protein A affinity tags, coupled with chromatographic purifications which exploit these tags via divalent metal-ion chelation or binding to IgG-Sepharose, are proving extremely valuable (see later Chapters). However, glycosyltransferases appear to be particularly intractable enzymes for the purposes of crystallography, and a protein structure at this time remains elusive despite considerable efforts by several laboratories.

An important development in the ongoing battle to clone the genes for glycosyltransferases was the introduction to glycobiology in 1989 of expression cloning.[232] In this ingenious approach, carefully chosen antigen-deficient cell-lines are transfected with an exogenous DNA library and the resultant transfectants are then screened, commonly by the methods of panning or cell-sorting, for their ability to synthesise the appropriate cell-surface carbohydrate antigens, which ought logically to correlate with the acquisition of the pertinent glycosyltransferase gene. This powerful technique, which has its roots in the work of Seed and Aruffo,[233,234] was first exploited by John Lowe and his co-workers to facilitate the cloning of the human *H*-gene α1,2-fucosyltransferase.[235,236] Since then, the same method has proved effective for obtaining clones for a host of other genes, including those encoding the xenoantigen-forming α1,3-galactosyltransferase,[237] the transferase forming the so-called I blood-group antigens,[238] the branching β1,6-*N*-acetylglucosaminyltransferase responsible for synthesising *O*-linked oligosaccharides of the 'core 2' variety,[239] and an α2,8-polysialyltransferase present in human foetal brain.[240]

The late 1980s were also marked by the introduction of some important terminology into the scientific vernacular that necessarily acknowledged an

enormous conceptual leap forward by leading researchers in the field. Indeed, only then did scientists officially christen their subject 'glycobiology' and begin to call themselves 'glycobiologists.' This realisation, which swept the scientific community of the time, is exemplified by a simple reality in Nature: not all glycoproteins are created equal. This last statement is worthy of a short digression in order to clarify what we mean by 'equal.' One of the factors of profound importance in the glycosciences is the unique propensity of glycoproteins to exhibit what has been termed 'microheterogeneity.'[241] Elsewhere in biology, it is taken for granted that any two molecules of the same protein, isolated from the same source, share identical amino acid sequences and, of course, the same holds true for two molecules of a particular RNA or DNA with respect to their component nucleotides. This molecular homogeneity simply reflects the template-driven nature of protein and nucleic acid biosynthesis. However, proteins that undergo the biochemical addition of carbohydrate side-chains to become glycoproteins do so in a rather less rigorous fashion. There are apparently no templates to guide the cellular machinery, nor any opportunities for molecular proofreading; instead the organism relies upon the sequential action of a series of biosynthetic enzymes. Such enzymes are called glycosyltransferases or glycosidases, depending on whether they catalyse the addition or the removal, respectively, of carbohydrate, and, as we shall discover in Chapter 3, they are organised into various biochemical 'assembly lines' within the cell. The upshot of this situation is that the action of many of the key enzymes is contingent upon a whole host of highly variable factors, with the final outcome that otherwise identical copies of a glycoprotein will carry different sugar side-chains. For example, a 'pure' sample of a chosen glycoprotein that possesses one or more potential sites for glycosylation will in fact comprise a mixture of compounds, which vary considerably from one molecule to the next in terms of the exact sugar structures found at each of these sites. The range of glycosylation 'possibilities' becomes especially broad in molecules with multiple glycosylation sites such as thyroglobulin and α_2-macroglobulin. In essence, then, the chosen glycoprotein is a heterogeneous 'soup' from which a handful of molecules chosen at random will comprise the same polypeptide backbone but will not necessarily share the same ensembles of sugar side-chains.

In 1988, Thomas Rademacher and his colleagues adopted the terms 'glycoform' and 'glycotype' to describe this phenomenon of microheterogeneity more fully.[4] Two molecules of the same glycoprotein obtained from the same tissue but bearing different ensembles of oligosaccharide side-chains are glycoforms. Thus the macroscopic glycoprotein sample is a composite of the various populations of

contributing glycoforms. Moreover, each glycoform may possess differing physical or biochemical properties, therefore leading to functional diversity. Hence, microheterogeneity may conceivably represent a subtle, complementary mechanism of biological regulation, since host cells possess the ability to control the functioning of a glycoprotein both by modulating expression of the relevant gene and by proportionately altering the overall glycosylation profile. These concepts have significant implications for the use of recombinant glycoproteins as therapeutic agents because observable biological properties in which the carbohydrate moieties are known to participate will be defined by the weighted average of the component glycoform populations.[242-244] The term glycotype, on the other hand, refers to the unique glycosylation capacity of a particular cell type. This reflects the variation in protein glycosylation from one species of organism to the next, or among different tissues in the same animal, or even between the male and the female. A well-known example is the observable difference in the carbohydrate chains of the glycoprotein γ-glutamyl transpeptidase (γ-GT) depending upon whether it has been isolated from liver or kidney. In kidney samples, so-called 'bisected' *N*-linked oligosaccharides predominate that are characterised by the common core structure GlcNAc(β1–2)Man(α1–6)-[**GlcNAc**(β1–4)][GlcNAc-(β1–2)Man(α1–3)]Manβ (the monosaccharide in bold type is the 'bisecting' element), whereas γ-GT from liver tissue exhibits almost exclusively unbisected oligosaccharides (i.e. with the core structure GlcNAc(β1–2)-Man(α1–6)[GlcNAc(β1–2)Man(α1–3)]Manβ).[245] Hence, mammalian liver and kidney are said to have different glycotypes to discriminate the markedly dissimilar aptitudes of these cell types for glycosylating the same protein.

The implications of microheterogeneity cannot be overemphasised. It is now abundantly clear that different glycoforms have varying biological properties and activities,[246,247] as highlighted primarily by the findings of Raymond Dwek's group at Oxford University. They showed that recombinant tissue plasminogen activator, a human anticoagulant, cleaved plasminogen in the same fibrin-dependent manner, but at differential rates depending on the ensemble of glycoforms present in a particular preparation.[248,249] These findings have set a precedent for the necessity to distinguish between different glycoforms in drug patents.[250] More recently, the same group demonstrated the variable catalytic activities exhibited by electrophoretically-pure glycoforms of ribonuclease B isolated from bovine pancreas.[251] It is also clear that the precise nature of the oligosaccharides associated with different glycoforms, and in particular, the type of sugar residues that terminate these structures will determine the relative rates at which they will be cleared from circulation by hepatic lectins, and hence their pharmacokinetics.[242-244]

In closing, it is worth our briefly mentioning a few other important disparate findings that are key to various specialist fields within glycobiology. A prime example is the discovery that *N*-linked oligosaccharides, in particular, have essential roles in mediating the transduction of hormonal signals.[252] Especially notable is the report by Francisco Calvo and Robert Ryan in 1985 that artificially relieving the pituitary hormone, human chorionic gonadotropin, of its *N*-glycosylation switches this molecule's hormonal behaviour from agonist to antagonist.[253] Meanwhile, following discoveries made in 1984 by Gerry Hart and his group in Birmingham, Alabama, it has become increasingly clear throughout the 1980s and 1990s that glycosylation is not limited to surface and extracellular proteins, but is also widespread in the nucleus and cytoplasm of the cell. Current evidence suggests that this novel mode of *O*-glycosylation may have significant roles to play in a number of important biological processes such as the regulation of transcription and maintaining cytoskeletal organisation.[254,255] Finally, it is only now being realised that biological compounds such as glycogen and other storage polysaccharides, which were long thought to comprise wholly carbohydrate, in fact contain covalently-linked proteins which act as self-initiating primers during the early phase of biogenesis.[70,256] Therefore, even today, protein glycosylation is still being found in surprising or unusual circumstances, and the full extent of its distribution in Nature, or of its multifaceted role in biological processes, is likely to remain unknown for some considerable period to come.

4. ONLY THE BEGINNING ...

In this chapter we have introduced several of the major concepts that are the foundation of glycoscience, and have presented the reader with some of the most important relevant discoveries in the field within the broader context of its historical development. However, we have yet to embark upon our journey into the molecular heart of protein glycosylation. Our travels begin in the next chapter, where we discuss the functional implications of oligosaccharides—as far as we presently understand them—on the biology of glycoproteins, thereby rationalising why protein glycosylation has come to merit so much attention from scientists in recent years.

REFERENCES

1. Crick FHC. The structure of the hereditary material. Sci Am 1954; 191:54-61.

2. Sharon N, Lis H. Carbohydrates in cell recognition. Sci Am 1993; 268:74-81.
3. Muir H. The nature of the link between protein and carbohydrate of a chondroitin sulphate complex from hyaline cartilage. Biochem J 1958; 69:195-204.
4. Rademacher TW, Parekh RB, Dwek RA. Glycobiology. Annu Rev Biochem 1988; 57:785-838.
5. Montreuil J. Méthodologie de la structure et du métabolisme des glycoconjugués (glycoprotéines et glycolipides). Actes du Colloque Internationale du CNRS, N° 221, Villeneuve d'Ascq, juin 1973. Paris: CNRS, 1973;1:25-29.
6. Morell AG, Irvine RA, Sternlieb I, et al. Physical and chemical studies on caeruloplasmin. V. Metabolic studies on sialic acid-free ceruloplasmin in vivo. J Biol Chem 1968; 243:155-159.
7. Gribben JG, Devereux S, Thomas NSB, et al. Development of antibodies to unprotected glycosylation sites on recombinant human GM-CSF. Lancet 1990; 335:434-437.
8. Karlsson K-A. Microbial recognition of target-cell glycoconjugates. Curr Opin Struct Biol 1995; 5:622-635.
9. Muramatsu T. Carbohydrate signals in metastasis and prognosis of human carcinomas. Glycobiology 1993; 3:294-296.
10. Kannagi R. Carbohydrate-mediated cell adhesion involved in haematogenous metastasis of cancer. Glycoconj J 1997; 14:577-584.
11. McEver RP. Selectin–carbohydrate interactions during inflammation and metastasis. Glycoconj J 1997; 14:585-591.
12. Lee YC. The interface between glycoscience and glycotechnology. Trends Glycosci Glycotechnol 1997; 9:S1-S7.
13. Alavi A, Axford JS, eds. Glycoimmunology. New York: Plenum Press, 1995. (Back N, Cohen IR, Kritchevsky D, et al, eds. Advances in Experimental Medicine and Biology; vol 376).
14. Kobata A. A retrospective and prospective view of glycopathology. Glycoconj J 1998; 15:323-331.
15. Hudson CS. The significance of certain numerical relations in the sugar group. J Am Chem Soc 1909; 31:66-86.
16. Rosanoff MA. On Fischer's classification of stereo-isomers. J Am Chem Soc 1906; 28:114-121.
17. Koenigs W, Knorr E. Ueber einige Derivate des Traubenzuckers und der Galactose. Chem Ber 1901; 34:957-981.
18. Paulsen H. Advances in selective chemical synthesis of complex oligosaccharides. Angew Chem Int Ed Engl 1982; 21:155-173.
19. Schmidt RR. New methods for the synthesis of glycosides and oligosaccharides—are there alternatives to the Koenigs-Knorr method? Angew Chem Int Ed Engl 1986; 25:212-235.
20. Whitfield DM, Douglas SP. Glycosylation reactions—present status, future directions. Glycoconj J 1996; 13:5-17.
21. Montreuil J. Recent data on the structure of the carbohydrate moiety of glycoproteins. Metabolic and biological implications. Pure Appl Chem 1975; 42:431-477.
22. Kisailus EC, Allen HJ. Nomenclature of monosaccharides, oligosaccharides and glycoconjugates. In: Allen HJ, Kisailus EC, eds. Glycoconjugates: Composition, Structure and Function. New York: Marcel Dekker, 1992:13-32.
23. Vliegenthart JFG, Montreuil J. Primary structure of glycoprotein glycans. In: Montreuil J, Schachter H, Vliegenthart JFG, eds. Glycoproteins. Amsterdam: Elsevier, 1995:13-28. (Neuberger A, van Deenen LLM, eds. New Comprehensive Biochemistry; vol 29a).

24. Joint Recommendations of the Physiological and Biochemical Committees on Protein Nomenclature. J Biol Chem 1908; 4:xlviii-li.
25. Jeanloz RW. The nomenclature of mucopolysaccharides. Arthritis Rheum 1960; 3:233-237.
26. Montreuil J. Primary structure of glycoprotein glycans. Basis for the molecular biology of glycoproteins. Adv Carbohydr Chem Biochem 1980; 37:157-223.
27. Webb EC, ed. Enzyme Nomenclature 1992: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes. San Diego: Academic Press, Inc., 1992.
28. Kent PW. Some landmarks in carbohydrate biochemistry: 1921-1986. Bull Biochem Soc 1988; 10:4-10.
29. Kent PW. Some landmarks in carbohydrate biochemistry: 1921-1986. Part II. Some biological functions of carbohydrates and their biochemical transformations. Bull Biochem Soc 1988; 11:4-13.
30. Gottschalk A. Historical introduction. In: Gottschalk A, ed. Glycoproteins: Their Composition, Structure and Function. 2nd ed. Amsterdam: Elsevier, 1972:1-23; vol A.
31. Scherer JJ. Ueber den flüssigen Schleimstoff des tierischen Körpers. Ann Chem Pharm 1846; 57:196-201.
32. Eichwald E. Beiträge zur Chemie der gewebbildenden Substanzen und ihrer Abkömmlinge. I. Ueber das Mucin, besonders der Weinbergschnecke. Ann Chem Pharm 1865; 134:177-211.
33. Obolenski S. Ueber das Mucin aus der Submaxillardrüse. In: Hoppe-Seyler F, ed. Medizinisch-chemische Untersuchungen. Berlin: A Hirschwald, 1871:590-593; vol 4.
34. Giacosa P. Études sur la composition chimique de l'oeuf et de ses enveloppes chez la grenouille commune. I. Sur l'enveloppe muqueuse de l'oeuf. Hoppe-Seyler's Z Physiol Chem 1882; 7:40-56.
35. Hammarsten O. Studien über Mucin und mucinähnliche Substanzen. Pflüger's Arch Physiol 1885; 36:373-456.
36. Hammarsten O. Ueber das Mucin der Submaxillardrüse. I. Darstellung, Zusammensetzung und Eigenschaften des Submaxillarismucins. Hoppe-Seyler's Z Physiol Chem 1888; 12:163-195.
37. Zanetti CU. Sull'ovimucoide e sopra un nuovo glicoproteide contenuto nel siero di sangue. Ann Chim Farmac 1897; 24:529-534.
38. Müller F. Die Chemie des Mucins und der Mucoide. Sitzungsber Gesell Beförderung Naturwiss Marburg 1898; 6:117-136.
39. Gottschalk A. The Chemistry and Biology of Sialic Acids and Related Substances. Cambridge: Cambridge University Press, 1960.
40. Hoppe-Seyler F. Physiologische Chemie. Berlin: A Hirschwald, 1881:94-95; vol 1.
41. Fischer E. Ueber die Configuration des Traubenzuckers und seiner Isomeren. Chem Ber 1891; 24:1836-1845, 2683-2687.
42. Purdie T, Irvine JC. The alkylation of sugars. J Chem Soc 1903; 1021-1037.
43. Hirst EL. The structure of the normal monosaccharides. Part IV. Glucose. J Chem Soc 1926:350-357.
44. Haworth WN. The Constitution of Sugars. London: E Arnold & Co., 1929.
45. Haworth NW, Lake WHG, Peat S. The configuration of glucosamine (chitosamine). J Chem Soc 1939:271-274.
46. James SP, Smith F, Stacey M, et al. A constitutional synthesis of chondrosamine. Nature 1945; 156:308-309.

47. Levene PA, Mikeska LA, Mori T. On the carbohydrate of thymonucleic acid. J Biol Chem 1929-30; 85:785-787.
48. Schmeideberg O. Ueber die chemische Zusammensetzung des Knorpels. Arch Exp Path Pharmakol 1891; 28:355-404.
49. Levene PA, La Forge FB. On chondroitin sulphuric acid. J Biol Chem 1913-1915; 15:69-79, 15:155-160; 19:123-130; 20:433-444.
50. Neuberger A. Carbohydrates in proteins. I. The carbohydrate component of crystalline egg albumin. Biochem J 1938; 32:1435-1451.
51. Neuberger A. Early work on the glycoprotein of egg albumin. Trends Biochem Sci 1988; 13:398-399.
52. Blix G. Ueber die Kohlenhydratgruppen des Submaxillarismucins. Hoppe-Seyler's Z Physiol Chem 1936; 240:43-54.
53. Blix G, Svennerholm L, Werner I. The isolation of chondrosamine from gangliosides and from submaxillary mucin. Acta Chem Scand 1952; 6:358-362.
54. Klenk E. Neuraminsäure, das Spaltprodukt eines neuen Gehirnlipoids. Hoppe-Seyler's Z Physiol Chem 1941; 268:50-58.
55. Cornforth JW, Firth ME, Gottschalk A. The synthesis of *N*-acetylneuraminic acid. Biochem J 1958; 68:57-61.
56. Comb DG, Roseman S. Composition and enzymatic synthesis of *N*-acetylneuraminic acid (sialic acid). J Am Chem Soc 1958; 80:497-499.
57. Schauer R, Kamerling JP. Chemistry, biochemistry and biology of sialic acids. In: Montreuil J, Vliegenthart JFG, Schachter H, eds. Glycoproteins II. Amsterdam: Elsevier, 1997:241-400. (Neuberger A, van Deenen LLM, eds. New Comprehensive Biochemistry; vol 29b).
58. Gregory JD, Laurent TC, Rodén L. Enzymatic degradation of chondromucoprotein. J Biol Chem 1964; 239:3312-3320.
59. Lindahl U, Rodén L. The chondroitin 4-sulphate–protein linkage. J Biol Chem 1966; 241:2113-2119.
60. Marshall RD, Neuberger A. Carbohydrates in protein. VIII. The isolation of 2-acetamido-1-(1-β-aspartamido)-1,2-dideoxy-β-D-glucose from hen's egg albumin. Biochemistry 1964; 3:1596-1600.
61. Tanaka K, Bertolini M, Pigman W. Serine and threonine glycosidic linkages in bovine submaxillary mucin. Biochem Biophys Res Commun 1964; 16:404-409.
62. Lindahl U, Rodén L. The linkage of heparin to protein. Biochem Biophys Res Commun 1964; 17:254-259.
63. Butler WT, Cunningham LW. Evidence for the linkage of a disaccharide to hydroxylysine in tropocollagen. J Biol Chem 1966; 241:3882-3888.
64. Lamport DTA. Hydroxyproline-*O*-glycosidic linkage of the plant cell wall glycoprotein extensin. Nature 1967; 216:1322-1324.
65. Sentandreu R, Northcote DH. The structure of a glycopeptide isolated from the yeast cell wall. Biochem J 1968; 109:419-432.
66. Lamport DT, Katona L, Roerig S. Galactosylserine in extensin. Biochem J 1973; 133:125-132.
67. Hallgren P, Lundblad A, Svensson S. A new type of carbohydrate–protein linkage in a glycopeptide from normal human urine. J Biol Chem 1975; 250:5312-5314.
68. Torres C-R, Hart GW. Topography and polypeptide distribution of terminal *N*-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for *O*-linked GlcNAc. J Biol Chem 1984; 259:3308-3317.

69. Hase S, Kawabata S, Nishimura H, et al. A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. J Biochem 1988; 104:867-868.
70. Lomako J, Lomako WM, Whelan WJ. A self-glucosylating protein is the primer for rabbit muscle glycogen biosynthesis. FASEB J 1988; 2:3097-3103.
71. Almeida IC, Ferguson MAJ, Schenkman S, et al. Lytic anti α-galactosyl antibodies from patients with chronic Chagas' disease recognise novel *O*-linked oligosaccharides on mucin-like glycosylphosphatidylinositol-anchored glycoproteins of *Trypanosoma brucei*. Biochem J 1994; 304:793-802.
72. Singh DG, Lomako J, Lomako WM, et al. β-Glucosylarginine: A new glucose-protein bond in a self-glucosylating protein from sweet corn. FEBS Lett 1995; 376:61-64.
73. Hirst GK. The agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus. Science 1941; 94:22-23.
74. McClelland L, Hare R. The adsorption of influenza virus by red cells and a new in vitro method of measuring antibodies for influenza virus. Can J Public Health 1941; 32:530-538.
75. Hirst GK. Adsorption of influenza haemagglutinins and virus by red blood cells. J Exp Med 1942; 76:195-209.
76. Burnet FM, Stone JD. The receptor-destroying enzyme of *V. cholerae*. Aust J Exp Biol Med Sci 1947; 25:227-233.
77. Burnet FM. Mucins and mucoids in relation to influenza virus action. III. Inhibition of virus haemagglutination by glandular mucins. Aust J Exp Biol Med Sci 1948; 26:371-379.
78. Francis T. Dissociation of the haemagglutinating and antibody-measuring capacities of influenza virus. J Exp Med 1947; 85:1-7.
79. McCrae JF. Mucins and mucoids in relation to influenza virus action. II. Isolation and characterization of the serum mucoid inhibitor of heated influenza virus. Aust J Exp Biol Med Sci 1948; 26:355-370.
80. Anderson SG. Mucins and mucoids in relation to influenza virus action. I. Inactivation by RDE and by viruses of the influenza group, of the serum inhibitor of haemagglutination. Aust J Exp Biol Med Sci 1948; 26:347-354.
81. Burnet FM. The initiation of cellular infection by influenza and related viruses. Lancet 1948; 254:7-11.
82. Gottschalk A. Neuraminidase: The specific enzyme of influenza virus and *Vibrio cholerae*. Biochim Biophys Acta 1957; 23:645-646.
83. Gottschalk A. Correlation between composition, structure, shape and function of a salivary mucoprotein. Nature 1960; 186:949-951.
84. Marchesi VT, Tillack TW, Jackson RL, et al. Chemical characterization and surface orientation of the major glycoprotein of the human erythrocyte membrane. Proc Natl Acad Sci USA 1972; 69:1445-1449.
85. Springer GF, Nagai Y, Tegtmeyer H. Isolation and properties of human blood-group NN and meconium-Vg antigens. Biochemistry 1966; 5:3254-3272.
86. Markwell M, Paulson J. Sendai virus utilizes specific sialyloligosaccharides as host cell receptor determinants. Biochemistry 1980; 77:5693-5697.
87. Tavakkol A, Burness ATH. Evidence for a direct role for sialic acid in the attachment of encephalomyocarditis virus to human erythrocytes. Biochemistry 1990; 29:10684-10690.

88. Schultze B, Herrler G. Bovine coronavirus uses *N*-acetyl-9-*O*-acetylneuraminic acid as a receptor determinant to initiate the infection of cultured cells. J Gen Virol 1992; 73:901-906.
89. Schultze B, Enjuanes L, Cavanagh D, et al. *N*-Acetylneuraminic acid plays a critical role for the haemagglutinating activity of avian infectious bronchitis virus and porcine transmissible gastroenteritis virus. Adv Exp Med Biol 1993; 342:305-310.
90. Landsteiner K. Zur Kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutserums und der Lymphe. Zentr Bakteriol Parasitenk 1900; 27:357-362.
91. Watkins WM. Biochemistry and genetics of the ABO, Lewis, and P blood group systems. In: Harris H, Hirschhorn K, eds. Advances in Human Genetics. New York: Plenum Press, 1980:1-136, 379-185; vol 10.
92. Greenwell P. Blood group antigens: Molecules seeking a function? Glycoconj J 1997; 14:159-173.
93. Annison EF, Morgan WTJ. Studies in immunochemistry. 10. The isolation and properties of Lewis (Le^a) human blood-group substance. Biochem J 1952; 50:460-472.
94. Marr AMS, Donald ASR, Watkins WM, et al. Molecular and genetic aspects of human blood-group Le^b specificity. Nature 1967; 215:1345-1349.
95. Mourant AE. A 'new' human blood group antigen of frequent occurrence. Nature 1946; 158:237-238.
96. Yamakami K. Discovery of secretion in seminal plasma and saliva. J Immunol 1926; 12:185-189.
97. Putkonen T. Über die gruppenspezifischen Eigenschaften verschiedener Korperflüssigkeiten. Acta Soc Med 'Duodecim' Ser A 1930; 14:107-121.
98. Lehrs H. Über gruppenspezifische Eigenschaften des menschlichen Speichels. Z Immunitätsforsch 1930; 66:175-192.
99. Yamamoto F, Clausen H, White T, et al. Molecular genetic basis of the histo-blood group ABO system. Nature 1990; 345:229-233.
100. Watkins WM, Morgan WTJ. Possible genetical pathways for the biosynthesis of blood group mucopolysaccharides. Vox Sang 1959; 4:97-119.
101. Kobata A, Grollman EF, Ginsburg V. An enzymic basis for blood type A in humans. Arch Biochem Biophys 1968; 124:609-612.
102. Kobata A, Grollman EF, Ginsburg V. An enzymatic basis for blood type B in humans. Biochem Biophys Res Commun 1968; 32:272-277.
103. Hearn VM, Smith ZG, Watkins WM. An α-*N*-acetyl-D-galactosaminyltransferase associated with the human blood-group A character. Biochem J 1968; 109:315-317.
104. Race C, Watkins WM. Properties of an α-D-galactosyltransferase in human tissues obtained in blood-group B donors. Biochem J 1969; 114:86P.
105. Yamamoto F, Hakomori S. Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. J Biol Chem 1990; 265:19257-19262.
106. Lowe JB. Selectin ligands, leukocyte trafficking, and fucosyltransferase genes. Kidney Int 1997; 51:1418-1426.
107. Rouquier S, Lowe JB, Kelly RJ, et al. Molecular cloning of a human genomic region containing the H blood group α(1,2)fucosyltransferase gene and two *H* locus-related DNA restriction fragments. Isolation of a candidate for the human Secretor blood group locus. J Biol Chem 1995; 270:4632-4639.

108. Kaneko M, Nishihara S, Shinya N, et al. Wide variety of point mutations in the *H* gene of Bombay and para-Bombay individuals that inactivate H enzyme. Blood 1997; 90:839-849.
109. Wang B, Koda Y, Soejima M, et al. Two missense mutations of H type α(1,2)fucosyltransferase gene (*FUT1*) responsible for para-Bombay phenotype. Vox Sang 1997; 72:31-35 [Correction: Vox Sang 1997;72:147].
110. Koda Y, Soejima M, Johnson PH, et al. Missense mutation of *FUT1* and deletion of *FUT2* are responsible for Indian Bombay phenotype of ABO blood group system. Biochem Biophys Res Commun 1997; 238:21-25.
111. Nishihara S, Yazawa S, Iwasaki H, et al. α(1,3/1,4)-Fucosyltransferase (FucT-III) gene is inactivated by a single amino acid substitution in Lewis histo-blood type negative individuals. Biochem Biophys Res Commun 1993; 196:624-631.
112. Nishihara S, Narimatsu H, Iwasaki H, et al. Molecular genetic analysis of the human Lewis histo-blood group system. J Biol Chem 1994; 269:29271-29278.
113. Kudo T, Iwasaki H, Nishihara S, et al. Molecular genetic analysis of the human Lewis histo-blood group system. II. Secretor gene inactivation by a novel single missense mutation A385T in Japanese nonsecretor individuals. J Biol Chem 1996; 271:9830-9837.
114. Fukuda M. Possible roles of tumor-associated carbohydrate antigens. Cancer Res 1996; 56:2237-2244.
115. Hakomori S. Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens. Adv Cancer Res 1989; 52:257-331.
116. Joziasse DH, Shaper JH, Jabs EW, et al. Characterization of an α1→3-galactosyltransferase homologue on human chromosome 12 that is organized as a processed pseudogene. J Biol Chem 1991; 266:6991-6998.
117. Shaper NL, Lin S, Joziasse DH, et al. Assignment of two human α-1,3-galactosyltransferase gene sequences (*GGTA1* and *GGTA1P*) to chromosomes 9q33-q34 and 12q14-q15. Genomics 1992; 12:613-615.
118. Joziasse DH, Shaper JH, van den Eijnden DH, et al. Bovine α1→3-galactosyltransferase: Isolation and characterization of a cDNA clone. Identification of homologous sequences in human genomic DNA. J Biol Chem 1989; 264:14290-14297.
119. Larsen RD, Rivera-Marrero CA, Ernst LK, et al. Frame-shift and nonsense mutations in a human genomic sequence homologous to a murine UDP-Gal:β-D-Gal(1,4)-D-GlcNAc α(1,3)-galactosyltransferase cDNA. J Biol Chem 1990; 265:7055-7061.
120. McKenzie IFC, Osman N, Cohney S, et al. Strategies to overcome the anti-Galα(1-3)-Gal reaction in xenotransplantation. Transplant Proc 1996; 28:537.
121. McKenzie IF, Osman N, Cohney S, et al. Strategies to overcome hyperacute rejection of pig-to-human xenografts. Transplant Proc 1997; 29:973-974.
122. Ashwell G, Morell AG. The role of surface carbohydrates in the hepatic recognition and transport of circulating glycoproteins. Adv Enzymol Related Areas Mol Biol 1974; 41:99-128.
123. Gesner BM, Ginsberg V. Effect of glycosidases on the fate of transfused lymphocytes. Proc Natl Acad Sci USA 1964; 52:750-755.
124. Boyd WC, Shapleigh E. Specific precipitating activity of plant agglutinins (lectins). Science 1954; 119:419.
125. Goldstein IJ, Hughes RC, Monsigny M, et al. What should be called a lectin? Nature 1980; 285:66.
126. Barondes SH. Bifunctional properties of lectins: Lectins redefined. Trends Biochem Sci 1988; 13:480-482.

127. Sharon N, Lis H. Lectins as cell recognition molecules. Science 1989; 246:227-234.
128. Drickamer K, Taylor ME. Biology of animal lectins. Annu Rev Cell Biol 1993; 9:237-264.
129. Drickamer K. Molecular structure of animal lectins. In: Fukuda M, Hindsgaul O, eds. Molecular Glycobiology. Oxford: IRL Press/Oxford University Press, 1994:53-87. (Hames BD, Glover DM, eds. Frontiers in Molecular Biology).
130. Drickamer K. Multiplicity of lectin-carbohydrate interactions. Nature Struct Biol 1995; 2:437-439.
131. Von Figura K, Hasilik A. Lysosomal enzymes and their receptors. Annu Rev Biochem 1986; 55:167-193.
132. Pfeffer SR. Mannose-6-phosphate receptors and their role in targeting proteins to lysosomes. J Membr Biol 1988; 103:7-16.
133. Dahms NM, Lobel P, Kornfeld S. Mannose 6-phosphate receptors and lysosomal enzyme targeting. J Biol Chem 1989; 264:12115-12118.
134. Kornfeld S, Mellman I. The biogenesis of lysosomes. Annu Rev Cell Biol 1989; 5:483-525.
135. Kornfeld S. Lysosomal enzyme targeting. Biochem Soc Trans 1990; 18:367-374.
136. Pfeffer SR. Mannose-6-phosphate receptors and their role in protein sorting along the pathway to lysosomes. Cell Biophys 1991; 19:131-140.
137. Lasky LA. Selectin–carbohydrate interactions and the initiation of the inflammatory response. Annu Rev Biochem 1995; 64:113-139.
138. McEver RP, Moore KL, Cummings RD. Leukocyte trafficking mediated by selectin–carbohydrate interactions. J Biol Chem 1995; 270:11025-11028.
139. Kansas GS. Selectins and their ligands: Current concepts and controversies. Blood 1996; 88:3259-3287.
140. Barondes SH, Cooper DNW, Gitt MA, et al. Galectins. Structure and function of a large family of animal lectins. J Biol Chem 1994; 269:20807-20810.
141. Barondes SH, Castronovo V, Cooper DNW, et al. Galectins: A family of animal β-galactoside-binding lectins. Cell 1994; 76:597-598.
142. Powell LD, Varki A. I-Type lectins. J Biol Chem 1995; 270:14243-14246.
143. Kelm S, Schauer R, Crocker PR. The sialoadhesins: A family of sialic acid-dependent cellular recognition molecules within the immunoglobulin superfamily. Glycoconj J 1996; 13:913-926.
144. Duguid JP, Smith JW, Dempster G, et al. Non-flagellar filamentous appendages ("fimbriae") and haemagglutinating activity in *Bacterium coli*. J Path Bacteriol 1955; 70:335-348.
145. Duguid JP, Gillies RR. Fimbriae and adhesive properties in dysentery bacilli. J Path Bacteriol 1957; 74:397-411.
146. Old DC. Inhibition of the interaction between fimbrial haemagglutinins and erythrocytes by D-mannose and other carbohydrates. J Gen Microbiol 1972; 71:149-157.
147. Duguid JP. The function of bacterial fimbriae. Arch Immunol Therap Exp 1968; 16:173-188.
148. Duguid JP, Clegg S, Wilson MI. The fimbrial and non-fimbrial haemagglutinins of *Escherichia coli*. J Med Microbiol 1979; 12:213-227.
149. Klemm P. Fimbrial adhesions of *Escherichia coli*. Rev Infect Dis 1985; 7:321-340.
150. Solter D, Knowles BB. Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). Proc Natl Acad Sci USA 1978; 75:5565-5569.

151. Feizi T. Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens. Nature 1985; 314:53-57.
152. Feizi T. Blood group-related oligosaccharides are ligands in cell-adhesion events. Biochem Soc Trans 1992; 20:274-278.
153. Gooi HC, Feizi T, Kapadia A, et al. Stage-specific embryonic antigen involves α1→3 fucosylated type 2 blood group chains. Nature 1981; 292:156-158.
154. Lloyd KO, Kabat EA, Licerio E. Immunochemical studies on blood groups. 38. Structures and activities of oligosaccharides produced by alkaline degradation of blood-group Lewis^a substance. Proposed structure of the carbohydrate chains of human blood-group A, B, H, Le^a, and Le^b substances. Biochemistry 1968; 7:2976-2990.
155. Marr AM, Donald AS, Morgan WT. Two new oligosaccharides obtained from a Le^a-active glycoprotein. Biochem J 1968; 110:789-791.
156. Caputto R, Leloir LF, Cardini CE, et al. Isolation of the coenzyme of the galactose phosphate-glucose phosphate transformation. J Biol Chem 1950; 184:333-350.
157. Leloir LF. The enzymatic transformation of uridine diphosphate glucose into a galactose derivative. Arch Biochem Biophys 1951; 33:186-190.
158. Tabas I, Schlesinger S, Kornfeld S. Processing of high mannose oligosaccharides to form complex type oligosaccharides on the newly synthesized polypeptides of the vesicular stomatitis virus G protein and the IgG heavy chain. J Biol Chem 1978; 253:716-722.
159. Li E, Tabas I, Kornfeld S. The synthesis of complex-type oligosaccharides. I. Structure of the lipid-linked oligosaccharide precursor of the complex-type oligosaccharides of the vesicular stomatitis virus G protein. J Biol Chem 1978; 253:7762-7770.
160. Kornfeld S, Li E, Tabas I. The synthesis of complex-type oligosaccharides. II. Characterization of the processing intermediates in the synthesis of the complex oligosaccharide units of the vesicular stomatitis virus G protein. J Biol Chem 1978; 253:7771-7778.
161. Tabas I, Kornfeld S. The synthesis of complex-type oligosaccharides. III. Identification of an α-D-mannosidase activity involved in a late stage of processing of complex-type oligosaccharides. J Biol Chem 1978; 253:7779-7786.
162. Hemming FW. The role of polyprenol-linked sugars in eukaryotic macromolecular synthesis. Biochem Soc Trans 1977; 5:1682-1687.
163. Hemming FW. Dolichol phosphate, a coenzyme in the glycosylation of animal membrane-bound glycoproteins. Biochem Soc Trans 1977; 5:1223-1231.
164. Hunt LA, Summers DF. Glycosylation of vesicular stomatitis virus glycoprotein in virus-infected HeLa cells. J Virol 1976; 20:646-657.
165. Hunt LA, Etchison JR, Summers DF. Oligosaccharide chains are trimmed during synthesis of the envelope glycoprotein of vesicular stomatitis virus. Proc Natl Acad Sci USA 1978; 75:754-758.
166. Robbins PW, Hubbard SC, Turco SJ, et al. Proposal for a common oligosaccharide intermediate in the synthesis of membrane glycoproteins. Cell 1977; 12:893-900.
167. Turco SJ, Robbins PW. The initial stages of processing of protein-bound oligosaccharides in vitro. J Biol Chem 1979; 254:4560-4567.
168. Hubbard SC, Robbins PW. Synthesis and processing of protein-linked oligosaccharides in vivo. J Biol Chem 1979; 254:4568-4576.
169. Hubbard SC, Robbins PW. Synthesis of the *N*-linked oligosaccharides of glycoproteins. Assembly of the lipid-linked precursor oligosaccharide and its relation to protein synthesis in vivo. J Biol Chem 1980; 255:11782-11793.

170. Spiro MJ, Spiro RG, Bhoyroo VD. Lipid-saccharide intermediates in glycoprotein biosynthesis. I. Formation of an oligosaccharide-lipid by thyroid slices and evaluation of its role in protein glycosylation. J Biol Chem 1976; 251:6400-6408.
171. Spiro RG, Spiro MJ, Bhoyroo VD. Lipid-saccharide intermediates in glycoprotein biosynthesis. II. Studies on the structure of an oligosaccharide-lipid from thyroid. J Biol Chem 1976; 251:6409-6419.
172. Spiro MJ, Spiro RG, Bhoyroo VD. Lipid-saccharide intermediates in glycoprotein biosynthesis. III. Comparison of oligosaccharide-lipids formed by slices from several tissues. J Biol Chem 1976; 251:6420-6425.
173. Harpaz N, Schachter H. Control of glycoprotein synthesis. IV. Bovine colostrum UDP-*N*-acetylglucosamine:α-D-mannoside β-2-*N*-acetylglucosaminyltransferase I. Separation from UDP-*N*-acetylglucosamine:α-D-mannoside β-2-*N*-acetylglucosaminyltransferase II, partial purification and substrate specificity. J Biol Chem 1980; 255:4885-4893.
174. Harpaz N, Schachter H. Control of glycoprotein synthesis. V. Processing of asparagine-linked oligosaccharides by one or more rat liver Golgi α-D-mannosidases dependent on the prior action of UDP-*N*-acetylglucosamine:α-D-mannoside β-2-*N*-acetylglucosaminyltransferase I. J Biol Chem 1980; 255:4894-4902.
175. Kornfeld R, Kornfeld S. Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem 1985; 54:631-664.
176. Behrens NH, Leloir LF. Dolichol monophosphate glucose: An intermediate in glucose transfer in liver. Proc Natl Acad Sci USA 1970; 66:153-159.
177. Parodi AJ, Behrens NH, Leloir LF, et al. The role of polyprenol-bound saccharides as intermediates in glycoprotein synthesis in liver. Proc Natl Acad Sci USA 1972; 69:3268-3272.
178. Behrens NH, Carminatti H, Staneloni RJ, et al. Formation of lipid-bound oligosaccharides containing mannose. Their role in glycoprotein synthesis. Proc Natl Acad Sci USA 1973; 70:3390-3394.
179. Leloir LF. The role of dolichol in protein glycosylation. Adv Exp Med Biol 1977; 83:9-19.
180. Gregory JD, Jeanloz RW, eds. Glycoconjugate Research. Proceedings of the Fourth International Symposium on Glycoconjugates. New York: Academic Press, 1979; vol 1.
181. Gregory JD, Jeanloz RW, eds. Glycoconjugate Research. Proceedings of the Fourth International Symposium on Glycoconjugates. New York: Academic Press, 1979; vol 2.
182. Elbein AD. Glycosylation inhibitors for *N*-linked glycoproteins. Methods Enzymol 1987; 138:661-709.
183. Elbein AD. Inhibitors of the biosynthesis and processing of *N*-linked oligosaccharide chains. Annu Rev Biochem 1987; 56:497-534.
184. Elbein AD. Glycosidase inhibitors: Inhibitors of *N*-linked oligosaccharide processing. FASEB J 1991; 5:3055-3063.
185. Kaushal GP, Elbein AD. Glycosidase inhibitors in study of glycoconjugates. Methods Enzymol 1994; 230:316-329.
186. Huffaker TC, Robbins PW. Temperature-sensitive yeast mutants deficient in asparagine-linked glycosylation. J Biol Chem 1982; 257:3203-3210.
187. Huffaker TC, Robbins PW. Yeast mutants deficient in protein glycosylation. Proc Natl Acad Sci USA 1983; 80:7466-7470.

188. Orlean P. Biogenesis of yeast wall and surface components. In: Pringle JR, Broach JR, Jones EW, eds. The Molecular and Cellular Biology of the Yeast *Saccharomyces*. Cell Cycle and Cell Biology. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1997:229-362.
189. Stanley P, Narasimhan S, Siminovitch L, et al. Chinese hamster ovary cells selected for resistance to the cytotoxicity of phytohemagglutinin are deficient in a UDP-*N*-acetylglucosamine:glycoprotein *N*-acetylglucosaminyltransferase activity. Proc Natl Acad Sci USA 1975; 72:3323-3327.
190. Stanley P. Glycosylation engineering. Glycobiology 1992; 2:99-107.
191. Stanley P, Ioffe E. Glycosyltransferase mutants: Key to new insights in glycobiology. FASEB J 1995; 9:1436-1444.
192. Stanley P, Raju TS, Bhaumik M. CHO cells provide access to novel *N*-glycans and developmentally regulated glycosyltransferases. Glycobiology 1996; 6:695-699.
193. Tanner W, Lehle L. Protein glycosylation in yeast. Biochim Biophys Acta 1987; 906:81-99.
194. Kukuruzinska MA, Bergh ML, Jackson BJ. Protein glycosylation in yeast. Annu Rev Biochem 1987; 56:915-944.
195. Herscovics A, Orlean P. Glycoprotein biosynthesis in yeast. FASEB J 1993; 7:540-550.
196. Schachter H. Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides. Adv Exp Med Biol 1986; 205:53-85.
197. Schachter H. The 'yellow brick road' to branched complex *N*-glycans. Glycobiology 1991; 1:453-461.
198. Schachter H. Biosynthesis. 2c. Glycosyltransferases involved in the synthesis of the *N*-glycan antennae. In: Montreuil J, Vliegenthart JFG, Schachter H, eds. Glycoproteins. Amsterdam: Elsevier, 1995:153-199. (Neuberger A, van Deenen LLM, eds. New Comprehensive Biochemistry; vol 29a).
199. Schachter H. Biosynthesis. 4b. Substrate level controls for *N*-glycan assembly. In: Montreuil J, Vliegenthart JFG, Schachter H, eds. Glycoproteins. Amsterdam: Elsevier, 1995:281-286. (Neuberger A, van Deenen LLM, eds. New Comprehensive Biochemistry; vol 29a).
200. Takasaki S, Mizuochi T, Kobata A. Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. Methods Enzymol 1982; 83:263-268.
201. Patel TP, Parekh RB. Release of oligosaccharides from glycoproteins by hydrazinolysis. Methods Enzymol 1994; 230:57-66.
202. Spiro RG. Study of the carbohydrates of glycoproteins. Methods Enzymol 1970; 28:3-43.
203. Tarentino AL, Maley F. Purification and properties of an endo-β-*N*-acetylglucosaminidase from *Streptomyces griseus*. J Biol Chem 1974; 249:811-817.
204. O'Neill RA. Enzymatic release of oligosaccharides from glycoproteins for chromatographic and electrophoretic analysis. J Chromatography A 1996; 720:201-215.
205. Aub JC, Tieslau C, Lankester A. Reactions of normal and tumour cell surfaces to enzymes. I. Wheat-germ lipase and associated mucopolysaccharides. Proc Natl Acad Sci USA 1963; 50:613-619.
206. Burger MM, Goldberg AR. Identification of a tumour-specific determinant on neoplastic cell surfaces. Proc Natl Acad Sci USA 1967; 57:359-366.
207. Inbar M, Sachs L. Structural difference in sites on the surface membrane of normal and transformed cells. Nature 1969; 223:710-712.
208. Inbar M, Sachs L. Interaction of the carbohydrate-binding protein concanavalin A with normal and transformed cells. Proc Natl Acad Sci USA 1969; 63:1418-1425.

209. Meezan E, Wu HC, Black PH, et al. Comparative studies on the carbohydrate-containing membrane components of normal and virus-transformed mouse fibroblasts. II. Separation of glycoproteins and glycopeptides by Sephadex chromatography. Biochemistry 1969; 8:2518-2524.
210. Buck CA, Glick MC, Warren L. Glycopeptides from the surface of control and virus-transformed cells. Science 1971; 172:169-171.
211. Warren L, Critchley D, Macpherson I. Surface glycoproteins and glycolipids of chicken embryo cells transformed by a temperature-sensitive mutant of Rous sarcoma virus. Nature 1972; 235:275-278.
212. Yamashita K, Ohkura T, Tachibana Y, et al. Comparative study of the oligosaccharides released from baby hamster kidney cells and their polyoma transformant by hydrazinolysis. J Biol Chem 1984; 259:10834-10840.
213. Pierce M, Arango J. Rous sarcoma virus-transformed baby hamster kidney cells express higher levels of asparagine-linked tri- and tetraantennary glycopeptides containing [GlcNAc-β(1,6)Man-α(1,6)Man] and poly-*N*-acetyllactosamine sequences than baby hamster kidney cells. J Biol Chem 1986; 261:10772-10777.
214. Kobata A. Structural changes induced in the sugar chains of glycoproteins by malignant transformation of producing cells and their clinical application. Biochimie 1988; 70:1575-1585.
215. Yamashita K, Tachibana Y, Ohkura T, et al. Enzymatic basis for the structural changes of asparagine-linked sugar chains of membrane glycoproteins of baby hamster kidney cells induced by polyoma transformation. J Biol Chem 1985; 260:3963-3969.
216. Hiraizumi S, Takasaki S, Shiroki K, et al. Transfection with fragments of the adenovirus *12* gene induces tumorigenicity-associated alteration of *N*-linked sugar chains in rat cells. Arch Biochem Biophys 1990; 280:9-19.
217. Dennis JW, Laferté S, Waghorne C, et al. β1-6 Branching of Asn-linked oligosaccharides is directly associated with metastasis. Science 1987; 236:582-585.
218. Dennis JW, Laferté S. Oncodevelopmental expression of -GlcNAcβ1–6Manα1–6-Manβ1- branched asparagine-linked oligosaccharides in murine tissues and human breast carcinomas. Cancer Res 1989; 49:945-950.
219. Fernandes B, Sagman U, Auger M, et al. β1–6 Branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. Cancer Res 1991; 51:718-723.
220. Kang R, Saito H, Ihara Y, et al. Transcriptional regulation of the *N*-acetylglucosaminyl-transferase V gene in human bile duct carcinoma cells (HuCC-T1) is mediated by Ets-1. J Biol Chem 1996; 271:26706-26712.
221. Buckhaults P, Chen L, Fregien N, et al. Transcriptional regulation of *N*-acetyl-glucosaminyltransferase V by the *src* oncogene. J Biol Chem 1997; 272:19575-19581.
222. Pierce M, Buckhaults P, Chen L, et al. Regulation of *N*-acetylglucosaminyl-transferase V and Asn-linked oligosaccharide β(1,6) branching by a growth factor signaling pathway and effects on cell adhesion and metastatic potential. Glycoconj J 1997; 14:623-630.
223. Parekh RB, Dwek RA, Sutton BJ, et al. Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. Nature 1985; 316:452-457.
224. Axford JS. The importance of oligosaccharides to rheumatic disease: A personal perspective. Glycoconj J 1997; 14:863-866.
225. Sadler JE, Beyer TA, Oppenheimer CL, et al. Purification of mammalian glycosyl-transferases. Methods Enzymol 1982; 83:458-514.

226. Nunez HA, Barker R. Enzymatic synthesis and carbon-13 nuclear magnetic resonance conformational studies of disaccharides containing β-D-galactopyranosyl and β-D-[1-^{13}C]galactopyranosyl residues. Biochemistry 1980; 19:489-495.
227. Shaper NL, Shaper JH, Meuth JL, et al. Bovine galactosyltransferase: Identification of a clone by direct immunological screening of a cDNA expression library. Proc Natl Acad Sci USA 1986; 83:1573-1577.
228. Narimatsu H, Sinha S, Brew K, et al. Cloning and sequencing of cDNA of bovine *N*-acetylglucosamine (β1–4)galactosyltransferase. Proc Natl Acad Sci USA 1986; 83:4720-4724.
229. Appert HE, Rutherford TJ, Tarr GE, et al. Isolation of a cDNA coding for human galactosyltransferase. Biochem Biophys Res Commun 1986; 139:163-168.
230. Narimatsu H. Recent progress in molecular cloning of glycosyltransferase genes of eukaryotes. Microbiol Immunol 1994; 38:489-504.
231. Field MC, Wainwright LJ. Molecular cloning of eukaryotic glycoprotein and glycolipid glycosyltransferases: A survey. Glycobiology 1995; 5:463-472 [Correction: Glycobiology 6:5].
232. Fukuda M, Bierhuizen MFA, Nakayama J. Expression cloning of glycosyltransferases. Glycobiology 1996; 6:683-689.
233. Seed B, Aruffo A. Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. Proc Natl Acad Sci USA 1987; 84:3365-3369.
234. Aruffo A, Seed B. Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. Proc Natl Acad Sci USA 1987; 84:8573-8577.
235. Rajan VP, Larsen RD, Ajmera S, et al. A cloned human DNA restriction fragment determines expression of a GDP-L-fucose:β-D-galactoside 2-α-L-fucosyltransferase in transfected cells. Evidence for isolation and transfer of the human *H* blood group locus. J Biol Chem 1989; 264:11158-11167.
236. Ernst LK, Rajan VP, Larsen RD, et al. Stable expression of blood group H determinants and GDP-L-fucose:β-D-galactoside 2-α-L-fucosyltransferase in mouse cells after transfection with human DNA. J Biol Chem 1989; 264:3436-3447.
237. Larsen RD, Rajan VP, Ruff MM, et al. Isolation of a cDNA encoding a murine UDP-galactose:β-D-galactosyl-1,4-*N*-acetyl-D-glucosaminide α-1,3-galactosyltransferase: Expression cloning by gene transfer. Proc Natl Acad Sci USA 1989; 86:8227-8231.
238. Bierhuizen M, Mattei MG, Fukuda M. Expression of the developmental I antigen by a cloned human cDNA encoding a member of a β-1,6-*N*-acetylglucosaminyltransferase gene family. Genes Dev 1993; 7:468-478.
239. Bierhuizen M, Fukuda M. Expression cloning of a cDNA encoding UDP-GlcNAc: Galβ1-3-GalNAc-R (GlcNAc to GalNAc) β1-6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen. Proc Natl Acad Sci USA 1992; 89:9326-9330.
240. Nakayama J, Fukuda MN, Fredette B, et al. Expression cloning of a human polysialyl-transferase that forms the polysialylated neural cell adhesion molecule present in embryonic brain. Proc Natl Acad Sci USA 1995; 92:7031-7035.
241. Kobata A. The carbohydrates of glycoproteins. In: Ginsberg V, Robbins PW, eds. Biology of Carbohydrates. 2nd ed. New York: John Wiley & Sons, 1984:87-161; vol 2.
242. Cumming DA. Glycosylation of recombinant protein therapeutics: Control and functional implications. Glycobiology 1991; 1:115-130.
243. Cumming DA. Structural heterogeneity and functional differentiation: A rationale for glycosylation analysis of recombinant therapeutics. Semin Cell Biol 1991; 2:273-279.

244. Cumming DA. Physiological relevance of protein glycosylation. Dev Biol Standard 1992; 76:83-94.
245. Yamashita K, Hitoi A, Tateishi N, et al. Organ-specific difference in the sugar chains of γ-glutamyltranspeptidase. Arch Biochem Biophys 1983; 225:993-996.
246. Rudd PM, Woods RJ, Wormald MR, et al. The effects of variable glycosylation on the functional activities of ribonuclease, plasminogen and tissue plasminogen activator. Biochim Biophys Acta 1995; 1248:1-10.
247. Rudd PM, Dwek RA. Glycosylation: Heterogeneity and the 3D structure of proteins. Crit Rev Biochem Mol Biol 1997; 32:1-100.
248. Parekh RB, Dwek RA, Thomas JR, et al. Cell-type-specific and site-specific *N*-glycosylation of type I and type II human tissue plasminogen activator. Biochemistry 1989; 28:7644-7662.
249. Parekh RB, Dwek RA, Rudd PM, et al. *N*-Glycosylation and in vitro enzymatic activity of human recombinant tissue plasminogen activator expressed in Chinese hamster ovary cells and a murine cell line. Biochemistry 1989; 28:7670-7679.
250. Friedman E. The future of glycobiology. Bio/Technology 1991; 9:610-611.
251. Rudd PM, Joao HC, Coghill E, et al. Glycoforms modify the dynamic stability and functional activity of an enzyme. Biochemistry 1994; 33:17-22.
252. Baenziger JU, Green ED. Pituitary glycoprotein hormone oligosaccharides. Structure, synthesis and function of the asparagine-linked oligosaccharides on lutropin, follitropin and thyrotropin. Biochim Biophys Acta 1988; 947:287-306.
253. Calvo FO, Ryan RJ. Inhibition of adenylyl cyclase activity in rat corpora luteal tissue by glycopeptides of human chorionic gonadotropin and the α-subunit of human chorionic gonadotropin. Biochemistry 1985; 24:1953-1959.
254. Hart GW, Kreppel LK, Comer FI, et al. *O*-GlcNAcylation of key nuclear and cytoskeletal proteins: reciprocity with *O*-phosphorylation and putative roles in protein multimerization. Glycobiology 1996; 6:711-716.
255. Hart GW. Dynamic *O*-linked glycosylation of nuclear and cytoskeletal proteins. Annu Rev Biochem 1997; 66:315-335.
256. Alonso MD, Lomako J, Lomako WM, et al. A new look at the biogenesis of glycogen. FASEB J 1995; 9:1126-1137.