Glycobiology vol. 5 no. 5 pp. 463–472, 1995

MINI REVIEW

Molecular cloning of eukaryotic glycoprotein and glycolipid glycosyltransferases: a survey

Mark C.Field[1] and Linda J.Wainwright

Laboratory of Cell Biology, Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AY, UK

[1]To whom correspondence should be addressed

The rapidity with which molecular sequence data are gathered continues to grow. The result is that, for many workers, it is increasingly difficult to keep abreast of the current state of play of molecular cloning, even for those genes that encode proteins of special interest. The clear success of the various worldwide genome projects has made this even more apparent, and by the end of 1996 the complete determination of the nucleotide sequences of the genomes of two eukaryotes, *Saccharomyces cerevisiae* and *Caenorhabditis elegans*, will have either been completed or will be nearing completion. This article is an attempt to provide, in an easily accessible format, a compilation of genes and cDNAs that have been sequenced and deposited in GenBank that encode transferase enzymes involved in eukaryotic glycoprotein or glycolipid biosynthesis. The full sequence information can be easily retrieved from a databank, e.g. GenBank, using the relevant accession number(s).

Key words: database/glycosylation/glycosyltransferase/molecular cloning/protein processing

Introduction

The rate of acquisition of nucleotide data continues to increase. Even with the availability of on-line databases such as GenBank, the EMBL database and others, it is often hard to maintain awareness of the current state of a given field. Simple search strings are frequently inefficient at complete retrieval of the required data; for example, a search for 'glycosyltr*' in GenBank, where '*' is a wild card character, results in very few retrieved entries, and more complex strategies result in a vast increase in the amounts of undesired data being obtained. This is as true of enzymes involved in the construction and processing of eukaryotic complex carbohydrate moieties as it is for any other area.

As a first approach to maintaining a current awareness survey of the progress in molecular cloning of glycosyltransferase genes and cDNAs, we have performed an extensive search of GenBank and the yeast genome sequence database. The present article is restricted to sequences from eukaryotes with a submission date of 1994 or earlier. The search has also been heavily restricted to include only those enzymes involved in processing glycoprotein and glycolipid carbohydrate moieties, and therefore excludes such areas as glucosyltransferases involved in starch and sucrose synthesis, and the modifications of other metabolic products, e.g. flavenoid glucosylation.

Glycosyltransferases are enzymes associated with the exocytotic compartments of eukaryotic cells (125). They are responsible for catalysis of the addition of monosaccharide units either to an existing glycan chain or to a peptide or lipid acceptor initiating a chain (see the box in Figure 1A). All donor monosaccharides are utilized in an activated form, either as a nucleotide sugar, e.g. GDP-mannose or, less frequently but equally important, as a lipid-linked donor, e.g. dolichol-P-glucose (Dol-P-Glc), with the expulsion of the activated carrier upon transfer of the sugar to the acceptor. The majority of these enzymes are lumenally oriented, i.e. with the catalytic domain within a membrane-bounded compartment, as in, for example, the glycoprotein galactosyltransferases and sialyltransferases.

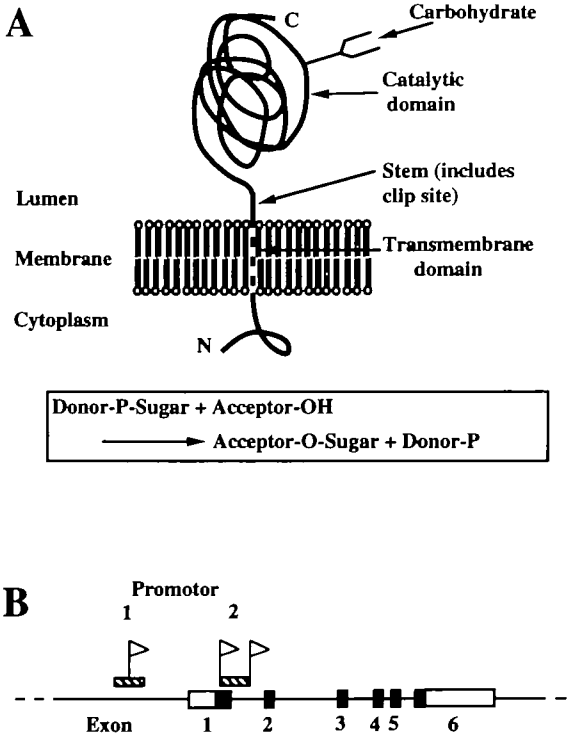

Fig. 1. Polypeptide and genomic organization of the eukaryotic glycosyltransferases. (**A**) Schematic of the domain structure of a glycosyltransferase enzyme, showing the orientation of the polypeptide, the catalytic, stem and transmembrane domains. The cartoon is based on sialyltransferase. The transmembrane region is designated by a dashed line, and the carboxyl and amino termini by C and N, respectively. The presence of an N-linked glycan is also indicated. (**B**) Exon and intron arrangement and transcription start sites for glycosyltransferase. Murine galactosyltransferase is used as the example (60–62,67). Regions containing promoter activity are shown by diagonal shading, and mapped transcriptional start sites are designated by flag symbols. The exon regions are boxed, and the 3′ and 5′ untranslated regions shown by open boxes, whilst the coding sequence is shown filled. The diagram is not to scale, and the precise distances between some of the elements are conjectural as a contiguous DNA fragment containing the entire transcriptional unit has not been isolated.

Conversely, in a number of cases there is evidence to suggest that a number of transferase activities are present on the cytoplasmic face of the secretory compartments. This orientation is particularly common for those enzymes involved in the synthesis of glycolipid glycans, e.g. the galactosyltransferase catalysing the formation of galactosylceramide, and several steps in the synthesis of glycosylphosphatidylinositol (GPI) lipids (126,127). The basic biosynthetic pathways for glycolipids of the sphingomyelin/ganglioside family, the GPI lipids and the N- and O-linked protein processing pathways have all been extensively reviewed (125,128–131,137).

The vast majority of glycosyltransferase enzymes working to elaborate glycan chains are of the type 2 transmembrane structure, i.e. containing a lumenally oriented C-terminus, the catalytic domain, a single membrane-spanning region and a cytoplasmic N-terminus (60–62,67), as shown in Figure 1A. The membrane-spanning region can confer important localization determinants, as elegantly demonstrated for sialyltransferase and galactosyltransferase (132,133). During early studies, many glycosyltransferases were believed to be soluble as any attempt to disrupt the cell to purify them resulted in the production of a clipped form of the enzyme, constituting the lumenal and enzymatically active C-terminal domain, cleaved between the catalytic domain and the transmembrane-spanning domain within the stalk region (Figure 1A). Molecular cloning studies served to explain the somewhat paradoxical observations that soluble enzymes without any discernible mechanism for retention within the secretory pathway could be shown to be tightly localized to specific subcompartments, as demonstrated by numerous ultrastructural studies. These observations also demonstrate that the C-terminal domain contains all the necessary information that is required for catalysis. As few of the enzymes involved in elaboration of ceramide or the GPI-lipids have been cloned, the precise topology of this group of proteins within the microsomal membranes remains to be fully determined.

Most of the lumenal glycosyltransferases are themselves glycoproteins. There is minimal homology between transferases beyond the basic domain structure (134,141). Within families of transferases there is considerable homology and a large number of members of, for example, the fucosyltransferase family, have been cloned by homology-based approaches, e.g. low-stringency screens or polymerase chain reaction (PCR) methods (see Table I for some examples). This latter family is particularly interesting as there is an extensive region, the Lewis X (LeX) locus, on human chromosome 19, that contains several similar α1,3/4-fucosyltransferases that are very closely related to each other at the nucleotide level, but have subtle differences in both precise specificity and tissue-specific expression (see below). Like most other eukaryotic genes, the glycosyltransferase genes are split by the presence of introns, and transcription is often initiated at more than one site, sometimes in a tissue-specific manner (see Figure 1B). In yeast, the presence of introns is characteristically low. Alternative exon usage can also lead to message heterogeneity, and in some cases even to different sized protein products. This topic has been thoroughly reviewed recently (121,141).

What follows is the result of our database searches for the period up to the end of 1994. The retrieved data are summarized with the accession numbers, the citation(s) that are associated with the sequence entries and the EC numbers for the activities, where available. Individual classes of glycosyltransferases are discussed briefly below.

Results and discussion

There were over 130 entries with distinct accession numbers in the GenBank database at the end of 1994. The earliest entries obtained were four galactosyltransferases in 1985, and since then a total of ~130 distinct glycosyltransferases have been cloned (see Figure 2). However, in many cases the homologue from several species has been studied, so that the total number is an overestimate of the repertoire now available in the database, and sequence data for a number of important activities remain completely lacking. The database information is summarized in Table I. It is possible that a small number of entries have been missed, but these are unlikely to alter our overall picture of the current state of molecular cloning of this group of eukaryotic enzymes.

Simple inspection of the table reveals two major points worthy of comment. Firstly, and most obviously, several enzymes are more heavily studied than others. Particularly well studied are the galactosyltransferases, the fucosyltransferases and the sialyltransferases. Clearly, the interest in these activities stems from the fact that they are responsible for terminal elaborations, and are therefore of more immediate biological relevance. The galactosyltransferases were, in fact, the first group of transferases to come under close scrutiny and, in the 4 years spanning 1986 to 1989, galactosyltransferase sequences accounted for ~70% of the then 23 entries in the database. Since 1989, the situation has become significantly more diverse, certainly in part reflecting interest in such structures as sialyl Lewis X (SLeX), involved in a number of cell adhesion events (122).

Secondly, the pace at which glycosyltransferases are being cloned is increasing. Inspection of Figure 2 suggests an almost exponential increase in the total number of entries in the database with time. If this is to continue, then ~35–40 new sequences will become available by the end of 1995. It is interesting to note that the data from various genome sequence projects have not been fully released to the databases as yet and, rather unexpectedly, only ~15% of the sequences retrieved have been obtained by this 'brute force' route. The impact that yeast sequencing has had on the number of mannosyltransferase sequences has been quite dramatic (but obviously, from the nature of yeast glycosylation, the impact on e.g. galactosyltransferase cloning is not going to be significant), and therefore the database can be expected to expand greatly when data are gathered from organisms with wider repertoires of glycan elaboration, e.g. *Caenorhabditis elegans*. Importantly, these large-scale studies will shed light on the genomic organization of glycosyltransferase genes, particularly poignant for the fucosyltransferase family, members of which are clustered (see below), and provide greater insight into the evolution of these proteins, their transcriptional control and ultimately the function that the glycans they synthesize have in a wider biological context.

We have refrained from providing phylogenetic reconstructions of the various gene families as the data are heavily imbedded within flanking sequences and introns, and would be too numerous for the present article. We hope that the provision of accession numbers and citations will allow interested workers to perform these studies using suitable software routines like PileUp from the GCG package (135) and PAUP (136). We have also completely avoided discussing the cloning strategies employed in obtaining these data. Whilst such information can be of considerable interest to those in the field, it is out of place here and the reader is referred to the excellent recent article by

Table I. Listing of eukaryotic glycosyltransferases that act on glycoproteins and glycolipids. For clarity, the enzyme names have been truncated. Systematic names, with the corresponding EC numbers, are given in Table II. In many cases, the gene or cDNA sequence has been obtained by homology (probing or PCR) or shotgun sequences, and the precise specificity of the encoded enzyme has not been determined. In these cases, an EC number is not given. NK, not known (in cases where the specificity of the transferase has not been defined). Sialyl Lewis X and Lewis X are abbreviated to SLeX and LeX throughout

Enzyme	EC designation	Species; source	GenBank accession	References	Notes
Sialyltransferases					
α2,3-sialyltransferase	EC 2.4.99.4	*H.sapiens*; cDNA	X74570	5	From WM266-4 melanoma
α2,3-sialyltransferase	EC 2.4.99.4	*H.sapiens*; cDNA	L23768	6	From placental mRNA
α2,3-sialyltransferase		*H.sapiens*; cDNA	L23767	7	Acts on both glycoproteins and glycolipids
α2,3-sialyltransferase	EC 2.4.99.4	*S.scrofa*; cDNA	M97753, M98463	8	Hepatic
α2,3-sialyltransferase	EC 2.4.99.6	*R.rattus*	M97754	13	
α2,3-sialyltransferase	EC 2.4.99.6	*M.musculus*; cDNA	X73523	11	Brain
α2,3-sialyltransferase	EC 2.4.99.4	*M.musculus*; cDNA	X76988, X76988	21	Cloned from brain
α2,3-sialyltransferase	EC 2.4.99.6	*M.musculus*; cDNA	D28941	25	Synthesis of SLeX[a]
α2,3-sialyltransferase	EC 2.4.99.4	*G.gallus*; cDNA	X77775	22	Testis
α2,3-sialyltransferase	EC 2.4.99.3	*G.gallus*; cDNA	X74946	23	From day 6 embryo. O-link specific
α2,6-sialyltransferase	EC 2.4.99.1	*H.sapiens*	A17362	1	
α2,6-sialyltransferase	EC 2.4.99.1	*H.sapiens*; cDNA	X17247	3	
α2,6-sialyltransferase	EC 2.4.99.1	*H.sapiens*; cDNA	X62822	4	From lymphocyte. Probably developmentally regulated
α2,6-sialyltransferase	EC 2.4.99.1	*H.sapiens*; cDNA	S55693, S55689	12	Demonstration of B-cell-specific exon usage
α2,6-sialyltransferase	NK	*H.sapiens*; cDNA	X54363	26	B lymphocyte specific. Responsible for synthesis of CD75 epitope. Cell surface
α2,6-sialyltransferase	EC 2.4.99.1	*H.sapiens*; cDNA	L11720	29	Differences in 5′ ends of transcripts
α2,6-sialyltransferase	EC 2.4.99.1	*H.sapiens*; cDNA	M38193	123	
α2,6-sialyltransferase	EC 2.4.99.1	*R.rattus*; cDNA	M83142, M83143 M83144	14	Demonstration of liver and kidney tissue specific splicing. Main sequence and alternate splices given
α2,6-sialyltransferase	EC 2.4.99.1	*R.rattus*; cDNA	M73985, M73986 M73987	15–17	Kidney and liver-specific splices. Alternate promoters drive tissue-specific transcription. M73985 encodes liver-specific 5′ end
α2,6-sialyltransferase	EC 2.4.99.1	*R.rattus*; genomic	M54999	18	Exon 1 of gene
α2,6-sialyltransferase	EC 2.4.99.1	*R.rattus*; cDNA	M18769	19	Complete coding sequence
α2,6-sialyltransferase	EC 2.4.99.1	*M.musculus*; cDNA	D16106	10	Hepatic. RT-PCR cloning strategy
α2,6-sialyltransferase	EC 2.4.99.1	*G.gallus*; cDNA	X75558	24	From day 6 embryo
α2,8-sialyltransferase		*H.sapiens*	X77922	27	Located on Chr12. Specific for GM3, makes GD3
α2,8-sialyltransferase		*H.sapiens*	D26360	28	Specific for GD3 synthesis
Sialytransferase	NK	*H.sapiens*; cDNA	U14550	9	Mammary epithelia. By homology
Sialyltransferase	EC 2.2.99.2	*H.sapiens*; cDNA	L13972	2	Cloned by homology. No data regarding specificity available
Sialyltransferase homologue	NK	*R.rattus*; cDNA	L13445	20	Degenerate PCR experiment. L13445 is a developmentally regulated newborn brain product
Fucosyltransferases					
α1,2fucosyltransferase	EC 2.4.1.69	*H.sapiens*; cDNA	M35531	36	Blood group H FT
α1,2fucosyltransferase	EC 2.1.4.69	*R.norwegicus*; cDNA	L26009,L26010	39	Isolated from a colon epithelial tumour. Evidence for two genes
α1,3fucosyltransferase III	EC 2.4.1.65	*H.sapiens*; cDNA	X53578	31	Synthesizes SSEA structure. Mapped to Chr19
α1,3fucosyltransferase IV	EC 2.4.1.65	*H.sapiens*; genomic	S52967, S52968	37	Mapped to Chr19. Linked within S52874, S52969 13 kb to LeX FTIII
α1,3fucosyltransferase IV	EC 2.4.1.65	*H.sapiens*	L01698	40	Part of LeX cluster on Chr19. Can make LeX, SLeX and difucosyl LeX
α1,3fucosyltransferase VI	EC 2.4.1.65	*H.sapiens*; cDNA	M98825	34	By RT-PCR. Can synthesize both SLeX and LeX *in vivo*
α1,3fucosyltransferase VII	EC 2.4.1.65	*H.sapiens*; cDNA	X78031	41	Forms SLeX structure. Mapped to Chr9
α1,3fucosyltransferase VII	EC 2.4.1.65	*H.sapiens*; cDNA	U08112, U11282	32	Makes SLeX, but not LeX *in vitro*
α1,3fucosyltransferase	EC 2.4.1.65	*H.sapiens*	M65030	35	Makes LeX and VIM-2 structures, but not the E-selectin ligand *in vivo*
α1,3fucosyltransferase	EC 2.4.1.65	*H.sapiens*; genomic	S65161	38	Probable identity with ELFT. Report of failure to make SLeX *in vivo* due to cell background
α1,3fucosyltransferase	EC 2.4.1.65	*H.sapiens*	M81485	30	Homologous to LeX FT, and can synthesize LeX. Part of Lewis locus on Chr19. Plasma type FT
α1,3fucosyltransferase	EC 2.4.1.65	*H.sapiens*; cDNA	M58596, M58597	33	Fucosyltransferase for formation of the E-selectin ligand (ELFT)
Galactosyltransferases					
α1,2galactosyltransferase		*S.pombe*	Z30917	75	Golgi resident
α1,3galactosyltransferase	EC 2.4.1.151	*H.sapiens*; genomic	J05421	70	Mapped to 12q14-q15. Pseudogene
α1,3galactosyltransferase	EC 2.4.1.151	*H.sapiens*	M65082	58	Homologous, but distinct to ABH or α1,3galactosyltransferase pseudogene

Continued

Table I. Continued

Enzyme	EC designation	Species; source	GenBank accession	References	Notes
α1,3galactosyltransferase	EC 2.4.1.151	*H.sapiens*	M60263	59	Pseudogene. Maps to Chr12
α1,3galactosyltransferase	EC 2.4.1.151	*H.sapiens*	J05421	55	Pseudogene. Not a reverse transcript but ORF corrupted at several points
α1,3galactosyltransferase	EC 2.4.1.151	*B.taurus*; cDNA	J04989	45	
α1,3galactosyltransferase	EC 2.4.1.151	*M.musculus*	M26925	73	
α1,3galactosyltransferase	EC 2.4.1.151	*M.musculus*; cDNA	M85153	76	Alternate splice site
α1,3galactosyltransferase	EC 2.4.1.151	*C.aethiops*	M73307	57	3′ genomic flank sequence given in each case
		P.paniscus	M72526		
		E.patas	M73308		
		G.gorilla	M73304		
		M.mulatta	M73306		
		A.geoffroyi	M73309		
		P.pygmaeus	M73305		
		S.sciureus	M73310		
		A.caraya	M73311		
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*H.sapiens*; cDNA	M22921,X14085	49	Placental. Complete coding sequence
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*H.sapiens*; cDNA	X51589	50	5′ end
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*H.sapiens*; cDNA	X55415	51	From adenocarcinoma (HeLa)
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*H.sapiens*; cDNA	X13223	52	Placental. Related to M13701
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*H.sapiens*; cDNA	U10472, U10473 U10474	53	From fetal liver
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*H.sapiens*	M14624	54	
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*H.sapiens*; genomic	M70427, M70428 M70429, M70430 M70432, M70433	56	From preB cells. Exons 16 given as individual accessions
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*B.taurus*	J05217	42	Two different N-termini encoded
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*B.taurus*; cDNA	M25398	44	
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*B.taurus*; cDNA	M13569	46	
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*B.taurus*; cDNA	X14558	47	Hepatic. Overlaps with M13214 & M13569
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*M.musculus*	L16840	60–62	Two ATGs possible. Sperm-specific transcripts contain a 5′ extension. Presence of regulatory elements at the 3′ end
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*M.musculus*; cDNA	D00314, D00315	63	From F9 cells
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*M.musculus*; cDNA	D37790, D37791	77	
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*M.musculus*; genomic	M27917, M27918 M27919, M27920 M27921, M27922 M27923	66	Exon 1, 2 (alternate splice), 3, 4, 5, 6 and 3′ UTR as separate accessions
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*M.musculus*	M36289	67	Evidence for alternate splicing creating long and short forms
β1,4galactosyltransferase	EC 2.4.1.38	*M.musculus*	J03880	71, 72	Two possible ATGs, and alternate to mRNA in spermatogenesis
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*G.gallus*; cDNA	L12565	68	
β1,4galactosyltransferase	EC 2.4.1.38, 2.4.1.90	*G.gallus*; cDNA	X16336	69	
Galactosyltransferase	EC 2.1.4.22	*H.sapiens*; cDNA	M13701	48	
Galactosyltransferase	NK	*B.taurus*; cDNA	M13214	43	
Galactosyltransferase	EC 2.4.1.45	*R.norwegicus*; cDNA	L21698, S66862	64	Acts on ceramide. From myelin
Galactosyltransferase	EC 2.4.1.62	*R.norwegicus*; cDNA	U07683	65	Acts on ceramide. From 18-day-old brain
Galactosyltransferase	NK	*C.elegans*; genomic	Z29095	74	Cosmid R10E11
Galactosyltransferase		*L.donovani*; genomic	L11348	78	Obtained by complementation (LPG1). Transfers galactofuranose. Expressed in promastigote only
N-Acetylglucosaminyltransferases					
GlcNActransferaseNAG1PTase		*M.musculus*; genomic	U03603	85	Transfers GlcNAc1P to dolichol
GlcNActransferaseNAG1PTase		*M.musculus*; cDNA	X65603, S41875	89	Transfers GlcNAc1P to dolichol. From mammary gland. Mapped to Chr17
GlcNActransferaseNAG1PTase		*C.longicaudatus*; cDNA	J05590, M22755	87, 88	Transfers GlcNAc1P to dolichol. Tunicamycin sensitive
GlcNActransferaseNAG1PTase		*L.mexicana*	M96635	101	Transfers GlcNAc1P to dolichol
GlcNActransferaseNAG1PTase		*S.cerevisiae*	Y00126	102	Transfers GlcNAc1P to dolichol
GlcNActransferaseNAG1PTase		*S.pombe*	U09454	103	Transfers GlcNAc1P to dolichol
β1,2GlcNActransferase	NK	*C.elegans*; genomic	Z46381	92	On Ch3, cosmid M01F1
β1,4GlcNActransferase		*L.stagnalis*; cDNA	X80228	100	Prostate. Adds GlcNAc to βGlcNAc
β1,6GlcNActransferase	EC 2.4.1.102	*H.sapiens*; cDNA	M97347	84	
β1,6GlcNActransferase		*H.sapiens*; cDNA	Z19550, L19656	94, 95	Forms blood group I. Mapped to Chr9
β1,6GlcNActransferase	EC 2.4.1.102	*M.musculus*; cDNA	U19295	79	Responsible for synthesis of the O-link core 2

Continued

Table I. Continued

Enzyme	EC designation	Species; source	GenBank accession	References	Notes
β1,6GlcNActransferase	NK	*C.elegans*; genomic	Z36752	92	On Ch3, cosmid F35H8
β1,6GlcNActransferase	NK	*C.elegans*; genomic	Z37092	92	On Ch3, cosmid F44F4
GlcNActransferaseI	EC 2.4.1.55	*H.sapiens*; cDNA	M55621	83	From A431 cells. Medial Golgi
GlcNActransferaseI	EC 2.4.1.101	*H.sapiens*; cDNA	T08544	91	EST06436 clone HIBBF95
GlcNActransferaseI	EC 2.4.1.101	*H.sapiens*	M61829	97	Located on Chr5
GlcNActransferaseI	EC 2.4.1.101	*R.rattus*; cDNA	D16302	98	Hepatic
GlcNActransferaseI	EC 2.4.1.101	*M.musculus*; genomic	M73491	81	
GlcNActransferaseI	EC 2.4.1.101	*M.musculus*; genomic	L07037	82	
GlcNActransferaseI	EC 2.4.1.101	*O.cuniculus*; cDNA	M57301	93	
GlcNActransferaseI	EC 2.4.1.143	*C.elegans*	Z46381	92	Located on Chr3
GlcNActransferaseIII	EC 2.4.1.144	*H.sapiens*	D13789	96	
GlcNActransferaseIII	EC 2.4.1.144	*R.rattus*; cDNA	D10852	99	Kidney
GlcNActransferaseV		*H.sapiens*; cDNA	D17716	90	Hepatic
GlcNActransferaseV	EC 2.4.1.155	*R.rattus*; cDNA	L14284	80	
N-Acetylgalactosaminyltransferases					
α1,3GalNActransferase	EC 2.4.1.40	*H.sapiens*; cDNA	J05175	104	Formation of blood group A
α1,3GalNActransferase	EC 2.4.1.40	*H.sapiens*; cDNA	S44054	124	A2 allele of blood group A
β1,4GalNActransferase	EC 2.4.1.92	*H.sapiens*; cDNA	M83651	105	Acts on GM2 and GD2
Protein;GalNActransferase	EC 2.4.1.41	*B.taurus*; cDNA	L07780	106	Initiation of O-link chain
Protein;GalNActransferase	EC 2.4.1.41	*B.taurus*; cDNA	L17437, L16925	107	Placental. Initiation of O-link chain
Mannosyltransferases					
Mannosyltransferase		*C.elegans*; cDNA	M75918	116	
Mannosyltransferase	EC 2.4.1.109	*S.cerevisiae*; genomic	L19169, L19700	108	PMT 1. Addition of O-linked mannosyl residues to protein from Dol-P-man
Mannosyltransferase	EC 2.4.1.109	*S.cerevisiae*; genomic	L05146	140	PMT 2
Mannosyltransferase	EC 2.4.1.83	*S.cerevisiae*	J04184	109	DPM 1. Synthesis of Dol-P-man. Also a 'flippase'
α1,2mannosyltransferase		*S.cerevisiae*; cDNA	M81110	114	MNT 1
α1,6mannosyltransferase		*S.cerevisiae*; genomic	D11095	115	OCH 1
β1,4mannosyltransferase		*S.cerevisiae*	J05416, Z35979	112	Alg1. Located on Chr9. Acts on Dol-P-$GlcNAc_2$
Mannosyltransferase		*S.cerevisiae*; genomic	L17083	113	KTR 2, part of the KRE 2 family
Mannosyltransferase		*S.cerevisiae*	X62941	139	KTR 1. On Chr 15
Mannosyltransferase		*S.cerevisiae*	X62647	139	KRE 2; allelic to MNT 1. On Chr 4
Mannosyltransferase		*S.cerevisiae*		117	Alg 2. Reported to have been cloned, but no GenBank entry
Mannosyltransferase		*S.cerevisiae*	L23753	138	MNN 1
Mannosyltransferase		*S.cerevisiae*	L23752	138	MNN 9; homologue of VAN 1
Mannosyltransferase		*S.cerevisiae*	Z38059	110	Located on Chr9. Putative
Mannosyltransferase		*S.cerevisiae*	Z46728	111	Located on Chr9
Glucosyltransferases					
Glucosyltransferase	EC 2.4.1.117	*S.cerevisiae*	X77573	119	Alg5. Synthesis of Dol-P-Glc
Glucosyltransferase		*S.cerevisiae*	X75929	118	Alg 8. Involved in synthesis of Dol-P-$GlcNAc_2Man_9Glc_3$

[a]See the list of abbreviations for this and others used in the table.

Schachter which covers this area in quite some depth (120). The remaining discussion will focus on particular glycosyltransferase classes/families.

Sialyltransferases

There are sequences for three major types of sialytransferase in the database: those forming α2,3, α2,6 and α2,8 linkages. The donor sugar is linked to CMP. The examples of the α2,3- and α2,6-forming transferases are restricted to those acting on N- and O-linked glycans of glycoproteins, whilst the two α2,8-specific transferases act on glycolipids and form polysialic acids. There is considerable evidence for tissue-specific expression of some of the enzymes, together with the use of alternative splicing, e.g. M83142 etcetera and M73985. In the latter case, there is also evidence for the use of alternate promoter elements for the initiation of transcription in a tissue-specific manner. Examples of sialyltransferases with defined high levels of specificity include X74946, for *O*-glycan sialylation, D28941 for synthesis of SLeX, and X77922 for production of GD2. A further curiosity is X54363, which was isolated in an attempt to clone the CD75 antigen from B cells. This clone turned out to encode a sialyltransferase responsible for synthesis of the CD75 antibody epitope (26).

Fucosyltransferases

The fucosyltransferases are perhaps the most complex group at the genomic level. They transfer fucose from UDP in α1-2, α1-3, α1-4 and α1-6 linkage. The latter two specificities can apparently be encoded by the same polypeptide. In *Homo sapiens*, there are a number of these transferase genes present as part of a cluster, the LeX locus, on chromosome 19. At least two of the genes in this locus are positioned very close to each

Table II. Enzyme commission (EC) nomenclature and full enzyme name. In formal notation the name of an enzyme is given as substrate 1 (donor): substrate 2 (acceptor) followed by the activity. All the glycosyltransferases considered here have EC designations beginning 2.4, designating them as glycosyltransferases. EC 2.4.1 are hexosyltransferases, and EC 2.4.99 are nonhexosyl- nonpentosyl- transferases, e.g. the sialyltransferases (121). As these are highly cumbersome, most enzymes are commonly known by trivial names, and this is equally the case for glycosyltransferases. This table allows the full name to be correlated with the accession number and other information given in Table I

EC number	Systematic enzyme name
2.4.1.38	UDP-Gal:GlcNAc-R β1,4-Galactosyltransferase
2.4.1.40	UDP-GalNAc.Fucα1-2Gal-R α1,3-GalNActransferase
2.4.1.41	UDP-GalNAc:Polypeptide GalNActransferase
2.4.1.45	UDP-Gal:Ceramide Galactosyltransferase
2.4.1.62	UDP-Gal:Ceramide Galactosyltransferase
2.4.1.65	Lewis GDP-Fuc:Galβ1-3(4)GlcNAc-R α4(3)-fucosyltransferase III
2.4.1.83	GDP-Man:Dolichol Phosphorylmannosyltransferase
2.4.1.90	UDP-Gal:GlcNAc-R β1,4-Galactosyltransferase
2.4.1.92	UDP-GalNAc:GM3/GD3 β1,4-GalNActransferase
2.4.1.102	UDP-GlcNAc:Galβ1-3GalNAc-R β1,6-GlcNActransferase
2.4.1.109	Dol-P-Man:Polypeptide Mannosyltransferase
2.4.1.117	UDP-Glc:Dolichol Phosphorylglucosyltransferase
2.4.1.143	UDP-GlcNAc:Manα1-6R β1,2-GlcNActransferase
2.4.1.144	UDP-GlcNAc:R1-Manα1-6-[GlcNAcβ1-2Manα1-3]Manβ1-4R2 β1,4-GlcNActransferase
2.4.1.151	UDP-Gal:Galβ1-4GlcNAc-R α1,3-Galactosyltransferase
2.4.1.155	UDP-GlcNAc:Manα1-6R β1-6GlcNActransferase
2.4.99.1	CMP-sialic acid:Galβ1-4GlcNAc-R α2,6-sialyltransferase
2.4.99.3	CMP-sialic acid:Galβ1-3GalNAc-R α2,3-sialyltransferase
2.4.99.4	CMP-sialic acid:Galβ1-3GalNAc-R α2,3-sialyltransferase
2.4.99.6	CMP-sialic acid:Galβ1-3(4)GlcNAc-R α2,3-sialyltransferase

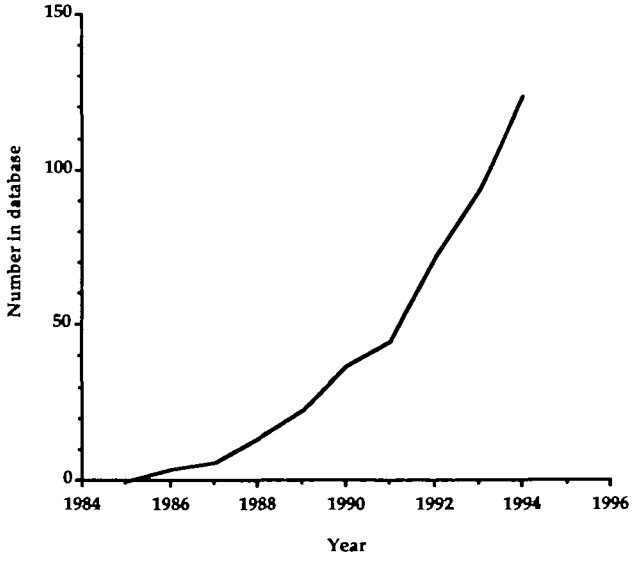

Fig. 2. Glycosyltransferase sequence submissions to GenBank by date. The period 1984–1994 is covered. According to this search, the first glycosyltransferases to be cloned, sequenced and submitted to GenBank were four galactosyltransferases, in 1986. Dates are taken from the date of publication of the sequence in a refereed journal and therefore are not necessarily the same as the submission date (a few sequences unpublished at the time of writing are included by submission date). A score of one per publication has been used as a more reliable index of activity than scoring accession numbers, as several reports have generated large numbers of sequence submissions which would artificially inflate the apparent rate of progress.

other; fucosyltransferase IV (S52987) and fucosyltransferase III (X53578) are within 13 kb of each other (37), and several more genes are probably located in this region, e.g. M81485. Interest in this particular area has arisen because the E-selectin cell adhesion molecule requires SLeX to be present on its ligand for recognition (122), although other structures also appear to be functional, and therefore most of the fucosyltransferases have been examined for their potential to form SLeX and LeX in different *in vivo* and *in vitro* assays. Results are occasionally contradictory, with subtle differences being displayed between the various α1,3 fucosyltransferases and the cell type within which they are expressed. The enzyme responsible for addition of core fucose (α1-6 linked) has not been cloned. A final enzyme of note within this group is the α1,2 fucosyltransferase responsible for formation of the blood group H structure (M35531).

Galactosyltransferases

The galactosyltransferases are an ever popular group of enzymes, catalysing the addition of galactose from UDP-gal in α1-3, β1-3, β1-4 and a variety of other linkages. They were the first group to be studied at the molecular genetic level, new sequences are still being generated, and they represent the largest single group in the database, with 37 entries. Many of these are redundant. In common with the sialyltransferases, there is substantial evidence for developmental regulation and alternative exon use with complex control of transcriptional initiation (see Figure 1). Unique for this class is the observation of pseudogenes in the human genome, related to the α1,3-type transferase. This is most probably involved in the interesting evolutionary phenomenon of the α-gal/anti-α-gal system, whereby Old World higher primates, including humans, have high levels of antibodies recognizing α1,3-linked terminal galactose, but do not synthesize the linkage, whereas in most other mammals the linkage is present and the antibody specificity is absent (57). Two sequences of glycolipid-specific enzymes from rat brain are in the database, and a galactofuranose transferase from *Leishmania donovani* has been reported. Only one galactosyltransferase sequence has been obtained from the genome sequencing programme to date (Z29095, from *C.elegans*).

N-Acetylglucosaminyltransferases

This group of enzymes add GlcNAc from UDP-GlcNAc in a number of linkages; most common are β1-4 (the core) and β1-2 (outer arm), but other linkages, e.g. β1-6, are also common amongst the more elaborate *N*-glycans and *O*-glycans. The enzyme responsible for the initiation of synthesis of the lipid-linked *N*-glycan precursor has been cloned from six different organisms. This enzyme is responsible for the sensitivity of *N*-glycan biosynthesis to tunicamycin, and is of interest as a key point in *N*-glycan biogenesis. Four of the *N*-acetylglucosaminyltransferases have been cloned by genome sequence projects. Owing to the importance of these enzymes in some of the later control points of *N*-glycan elaboration, this area should be active for some time, and one of these (GlcNAc transferase I) has been the subject of a recent gene knockout study unequivocally demonstrating the importance of glycan processing to the multicellular state (142).

N-Acetylgalactosaminyltransferases

These enzymes include two alleles of the blood group A: a ganglioside-specific transferase, and the activity responsible for the initiation of *O*-glycan chains on serine and threonine residues of polypeptides.

Mannosyltransferases

The vast majority of mannosyltransferases use GDP-man as a donor, but in the synthesis of the lipid-linked *N*-glycan precursor several enzymes use Dol-P-man as the mannosyl source (which is itself formed via GDP-man). Mannose can be linked in α1-2, α1-3, α1-4 and α1-6, and is also found in β-linkage (1-4) in the core region of *N*-glycans. The majority of mannosyltransferases have been cloned from the yeast *Saccharomyces cerevisiae*, including the mannose O-link initiating activity.

Glucosyltransferases

Only two of these enzymes have been cloned: those responsible for synthesis of Dol-P-Glc and for addition of glucose to the lipid-linked precursor of *N*-glycan biosynthesis.

Concluding remarks

Substantial progress has been made in the molecular biology of the glycosyltransferases since the first galactosyltransferase was cloned in 1986. As a group, these enzymes are becoming of increasing interest to workers outside of the immediate glycobiology field as tools for the study of intracellular compartmentalization, retention and protein processing. It is clear from the studies on several of the enzymes, e.g. the galactosyltransferases and sialyltransferase, that control of their transcription and tissue-specific expression is an important and interesting area, which will undoubtedly contribute to unravelling many of the roles that specific carbohydrate structures play in embryogenesis, cell adhesion and immunity. Particularly important will be an understanding of the developmental control of transcription and the mechanisms by which certain enzymes are retained within subcompartments of the secretory pathway.

It is intended that this database will be updated at least annually, and posted onto the World Wide Web within the near future. It is also hoped that the scope of the database can be expanded to include more information on each entry, and to include glycosylating activities of a more varied nature than those presented here.

Methods

Data retrieval

As sequence entries and flat files are now exchanged on a daily basis between all the major nucleotide databases (i.e. NDBJ, EMBL and GenBank), it is only necessary to search one of them as the data submitted to all others are, in the jargon, 'mirrored'. A second search was also performed on the *S.cerevisiae* genome sequence database using a download of the archive from the site at 'http://genome-www.stanford.edu'. The current database searches were performed in early 1995, and only entries submitted on or before 31 December 1994 are tabulated here.

Initially, GenBank was searched (address 'gopher://ftp.bio.indiana. edu:70/77/.indices/genbank?') via the World Wide Web (W3) using the Netscape browser (V1.1b) (available by ftp or other protocol from Netscape Industries Inc.) via an Internet connection with the following keywords: glycosyltr*, mannosyltr*, glucosyltr*, galactosyltr*, sialyltr*, glucuronyltr*, N-acetylglucosaminyltr*, N-acetylhexosaminyl*, N-acetylgalactosaminyl*, UDP-GalNAc, GDP-Man and GalNAc. Approximately 1000 entries were retrieved, of which ~150 were from eukaryotic sources. Interestingly, glycosyltr* produced <20 hits, and therefore is not a suitable search string. The W3 retrieval provided a rapid method to obtain accession numbers and to screen out undesired files without the need for a full file download as a brief title is also provided as part of the hypertext file name. The yeast genome was searched using glycosyltr*, mannosyltr*, glucosyltr*, galactosyltr*, sialyltr*, glucuronyltr* and transferase. A number of additional entries were also obtained by scanning the literature. Flat files were then retrieved from GenBank using an email program 'Eudora' (Qualcomm Inc.) to NetServ@ebi.ac.uk using the command 'GET NUC:' followed by the accession number (no spaces). The accumulated email messages were then converted to a single text file by extraction of the text string in the 'In' box of the Eudora mailer package, and the individual entries were then examined to abstract the relevant information as presented in Table I and Figure 2.

Acknowledgements

We thank Prof. Anne Dell and Dr Helen Field for critical reading of the manuscript, and Dr Michael P.Rout for bringing the availability of a *S.cerevisiae* genome information download to our attention. The authors should like to apologize for any errors or omissions in advance, and we greatly welcome corrections, suggestions and updates, preferably by email to m.field@ic.ac.uk. Work in the authors' laboratory is supported by the Wellcome Trust and the Medical Research Council of Great Britain.

Abbreviations

Chr, chromosome; Dol, dolichol; Dol-P-Glc, dolichol-P-glucose; EST, expressed sequence tag; FT, fucosyltransferase; fuc, fucose; gal, galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcNAc, *N*-acetylglucosamine; GPI, glycosylphosphatidylinositol; LPG, lipophosphoglycan; Man, mannose; -P-, phosphate; PCR, polymerase chain reaction; (S)LeX, (sialyl)Lewis X; W3, World Wide Web. All monosaccharides referred to in this publication are in the D-pyranose form, except fucose and in other cases where indicated.

References

1. Zettlmeissl,G., Grundmann,U., Becker,A. and Hermentin,P. (18-MAR-1992) Patent number EP0475354-A/3.
2. Chang,M.L. and Lau,J.T.Y. (1993) Unpublished.
3. Grundmann,U.G., Nerlich,C., Rein,T. and Zettlmeissl,G. (1990) Complete cDNA sequence encoding human β-galactoside α-2,6-sialyltransferase. *Nucleic Acids Res.*, **18**, 667.
4. Bast,B.J., Zhou,L., Freeman,G.J., Colley,K.J., Ernst,T.J., Munro,J. and Tedder,T.F. (1992) The HB-6, CDw75, and CD76 differentiation antigens are unique cell-surface carbohydrate determinants generated by the β-galactoside α-2,6-sialyltransferase. *J. Cell Biol.*, **116**, 423–435.
5. Sasaki,K., Watanabe,E., Kawashima,K., Sekine,S., Dohi,T., Oshima,M., Hanai,N., Nishi,T. and Hasegawa,M. (1993) Expression cloning of a novel Gal β-(1-3/1-4) GlcNAc α-2,3-sialyltransferase using lectin resistance selection. *J. Biol. Chem.*, **268**, 22782–22787.
6. Kitagawa,H. and Paulson,J.C. (1993) Cloning and expression of human Gal β-1,3(4)GlcNAc α-2,3-sialyltransferase. *Biochem. Biophys. Res. Commun.*, **194**, 375–382.
7. Kitagawa,H. and Paulson,J.C. (1994) Cloning of a novel α-2,3-sialyltransferase that sialylates glycoprotein and glycolipid carbohydrate groups. *J. Biol. Chem.*, **269**, 1394–1401.
8. Gillespie,W.M., Kelm,S. and Paulson,J. (1992) Cloning and expression of the Gal β-1,3GalNAc α-2,3-sialyltransferase. *J. Biol. Chem.*, **267**, 21004–21010.
9. Sotiropoulou,G., Anisowicz,A. and Sager,R. (1994) Unpublished.
10. Hamamoto,T., Kawasaki,M., Kurosawa,N., Nakaoka,T., Lee,Y.C. and Tsuji,S. (1993) Two step single primer mediated polymerase chain reaction. Application to cloning of putative mouse, β-galactoside α-2,6-sialyltransferase cDNA. *Bioorg. Med. Chem.*, **1**, 141–145.
11. Lee,Y.C., Kurosawa,N., Hamamoto,T., Nakaoka,T. and Tsuji,S. (1993) Molecular cloning and expression of Gal β-1,3GalNAc α-2,3-sialyltransferase from mouse brain. *Eur. J. Biochem.*, **216**, 377–385.
12. Wang,X., Vertino,A., Eddy,R.L., Byers,M.G., Jani-Sait,S.N., Shows,T.B. and Lau,J.T. (1993) Chromosome mapping and organization of the human β-galactoside α-2,6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J. Biol. Chem.*, **268**, 4355–4361.
13. Wen,D.X., Livingston,B.D., Medzihradszky,K.F., Kelm,S., Burlingame, A.L. and Paulson,J.C. (1992) Primary structure of Gal β-1,3(4)GlcNAc

α-2,3-sialyltransferase determined by mass spectrometry sequence analysis and molecular cloning. Evidence for a protein motif in the sialyltransferase gene family. *J. Biol. Chem.*, **267**, 21011–21019.
14. Wen,D.X., Svensson,E.C. and Paulson,J.C. (1992) Tissue-specific alternative splicing of the β-galactoside α-2,6-sialyltransferase gene. *J. Biol. Chem.*, **267**, 2512–2518.
15. O'Hanlon,T.P., Lau,K.M., Wang,X. and Lau,J.T.Y. (1989) Tissue-specific expression of β-galactoside α-2,6-sialyltransferase. Transcript heterogeneity predicts a divergent polypeptide. *J. Biol. Chem.*, **264**, 17389–17394.
16. Wang,X., O'Hanlon,T.P., Young,R.F. and Lau,J.T.Y. (1990) Rat β-galactoside α-2,6-sialyltransferase genomic organization: alternate promoters direct the synthesis of liver and kidney transcripts. *Glycobiology*, **1**, 25–31.
17. Wang,X., Smith,T.J. and Lau,J.T.Y. (1991) Transcriptional regulation of the liver β-galactoside α-2,6-sialyltransferase by glucocorticoids. *J. Biol. Chem.*, **265**, 17849–17853.
18. Svensson,E.C., Soreghan,B. and Paulson,J.C. (1990) Organization of the β-galactoside α-2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.*, **265**, 20863–20868.
19. Weinstein,J., Lee,E.U., McEntee,K., Lai,P.H. and Paulson,J.C. (1987) Primary structure of β-galactoside α-2,6-sialyltransferase. Conversion of membrane-bound enzyme to soluble forms by cleavage of the NH2-terminal signal anchor. *J. Biol. Chem.*, **262**, 17735–17743.
20. Livingston,B.D. and Paulson,J.C. (1993) Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family. *J. Biol. Chem.*, **268**, 11504–11507.
21. Lee,Y.C., Kojima,N., Wada,E., Kurosawa,N., Nakaoka,T., Hamamoto,T. and Tsuji,S. (1994) Cloning and expression of cDNA for a new type of Gal β-1,3GalNAc α-2,3-sialyltransferase. *J. Biol. Chem.*, **269**, 10028–10033.
22. Kurosawa,N., Kojima,N., Inoue,M., Hamamoto,T. and Tsuji,S. (1994) Cloning and expression of Gal β-1,3GalNAc-specific GalNAc α-2,6-sialyltransferase. *J. Biol. Chem.*, **269**, 19048–19053.
23. Kurosawa,N., Hamamoto,T., Lee,Y.C., Nakaoka,T., Kojima,N. and Tsuji,S. (1994) Molecular cloning and expression of GalNAc α-2,6-sialyltransferase. *J. Biol. Chem.*, **269**, 1402–1409.
24. Kurosawa,N., Kawasaki,M., Hamamoto,T., Nakaoka,T., Lee,Y.C., Arita, M. and Tsuji,S. (1994) Molecular cloning and expression of chick embryo Gal β-1,4GlcNAc α-2,6-sialyltransferase. Comparison with the mammalian enzyme. *Eur. J. Biochem.*, **219**, 375–381.
25. Sasaki,K., Watanabe,E., Kawashima,K., Hanai,N., Nishi,T. and Hasegawa, M. (1994) Unpublished.
26. Stamenkovic,I., Asheim,H.C., Deggerdal,A., Blomhoff,H.K., Smeland, E.B. and Funderud,S. (1990) The B cell antigen CD75 is a cell surface sialytransferase. *J. Exp. Med.*, **172**, 641–643.
27. Tsuji,S. and Nishi,T. (1994) Expression cloning of a GM3-specific α-2,8-sialyltransferase (GD3 synthase). *J. Biol. Chem.*, **269**, 15950–15956.
28. Nara,K., Watanabe,Y., Maruyama,K., Kasahara,K., Nagai,Y. and Sanai,Y. (1994) Expression cloning of a CMP-NeuAc:NeuAc α-3Gal β-1-4Glc β-1-1′Cer α-2,8-sialyltransferase (GD3 synthase) from human melanoma cells. *Proc. Natl Acad. Sci. USA*, **91**, 7952–7956.
29. Aasheim,H.C., Aas-Eng,A., Deggerdal,A., Blomhoff,H., Funderud,S. and Smeland,E. (1993) Cell-specific expression of human β-galactoside α-2,6-sialyltransferase transcripts differing in the 5′ untranslated region. *Eur. J. Biochem.*, **213**, 467–475.
30. Weston,B.W., Nair,R.P., Larsen,R.D. and Lowe,J.B. (1992) Isolation of a novel human α-(1,3)fucosyltransferase gene and molecular comparison to the human Lewis blood group α-(1,3/1,4)fucosyltransferase gene. Syntenic, homologous, nonallelic genes encoding enzymes with distinct acceptor substrate specificities. *J. Biol. Chem.*, **267**, 4152–4160.
31. Kukowska-Latallo,J.F., Larsen,R.D., Nair,R.P. and Lowe,J.B. (1990) A cloned human cDNA determines expression of a mouse stage-specific embryonic antigen and the Lewis blood group α-(1,3/1,4)fucosyltransferase. *Genes Dev.*, **4**, 1288–1303.
32. Natsuka,S., Gersten,K.M., Zenita,K., Kannagi,R. and Lowe,J.B. (1994) Molecular cloning of a cDNA encoding a novel human leukocyte α-1,3-fucosyltransferase capable of synthesizing the sialyl Lewis x determinant. *J. Biol. Chem.*, **269**, 16789–16794; erratum: *J. Biol. Chem.*, **269**, 20806.
33. Goelz,S.E., Hession,C., Goff,D., Griffiths,B., Tizard,R., Newman,B., Chi-Rosso,G. and Lobb,R. (1990) ELFT: a gene that directs the expression of an ELAM-1 ligand. *Cell*, **63**, 1349–1356.
34. Koszdin,K.L. and Bowen,B.R. (1992) The cloning and expression of a human α-1,3 fucosyltransferase capable of forming the E-selectin ligand. *Biochem. Biophys. Res. Commun.*, **187**, 152–157.
35. Lowe,J.B., Kukowska-Latallo,J.F., Nair,R.P., Larsen,R.D., Marks,R.M., Macher,B.A., Kelly,R.J. and Ernst,L.K. (1991) Molecular cloning of a human fucosyltransferase gene that determines expression of the Lewis x and VIM-2 epitopes but not ELAM-1-dependent cell adhesion. *J. Biol. Chem.*, **266**, 17467–17477.
36. Larsen,R.D., Ernst,L.K., Nair,R.P. and Lowe,J.B. (1990) Molecular cloning, sequence and expression of a human GDP-L-fucose:β-D-galactoside 2-α-L-fucosyltransferase cDNA that can form the H blood group antigen. *Proc. Natl Acad. Sci. USA*, **87**, 6674–6678.
37. Nishihara,S., Nakazato,M., Kudo,T., Kimura,H., Ando,T. and Narimatsu,H. (1993) Human α-1,3 fucosyltransferase (FucT-VI) gene is located at only 13 kb 3′ to the Lewis type fucosyltransferase (FucT-III) gene on chromosome 19. *Biochem. Biophys. Res. Commun.*, **190**, 4246.
38. Kumar,R., Potvin,B., Muller,W.A. and Stanley,P. (1991) Cloning of a human α-(1,3)-fucosyltransferase gene that encodes ELFT but does not confer ELAM-1 recognition on Chinese hamster ovary cell transfectants. *J. Biol. Chem.*, **266**, 21777–21783.
39. Piau,J.P., Labarriere,N., Dabouis,G. and Denis,M.G. (1994) Evidence for two distinct α-(1,2)-fucosyltransferase genes differentially expressed throughout the rat colon. *Biochem. J.*, **300**, 623–626.
40. Weston,B.W., Smith,P.L., Kelly,R.J. and Lowe,J.B. (1992) Molecular cloning of a fourth member of a human α-(1,3)fucosyltransferase gene family. Multiple homologous sequences that determine expression of the Lewis x, sialyl Lewis x, and difucosyl sialyl Lewis x epitopes. *J. Biol. Chem.*, **267**, 24575–24584; erratum: *J. Biol. Chem.*, **268**, 18398, 1993.
41. Sasaki,K., Kurata,K., Funayama,K., Nagata,M., Watanabe,E., Ohta,S., Hanai,N. and Nishi,T. (1994) Expression cloning of a novel α 1,3-fucosyltransferase that is involved in biosynthesis of the sialyl Lewis x carbohydrate determinants in leukocytes. *J. Biol. Chem.*, **269**, 14730–14737.
42. Russo,R.N., Shaper,N.L. and Shaper,J.H. (1990) Bovine β-1-4-galactosyltransferase: two sets of mRNA transcripts encode two forms of the protein with different amino-terminal domains. In vitro translation experiments demonstrate that both the short and the long forms of the enzyme are type II membrane-bound glycoproteins. *J. Biol. Chem.*, **265**, 3324–3331.
43. Shaper,N.L., Shaper,J.H., Meuth,J.L., Fox,J.L., Chang,H., Kirsch,I.R. and Hollis,G.F (1986) Bovine galactosyltransferase: identification of a clone by direct immunological screening of a cDNA expression library. *Proc. Natl Acad. Sci. USA*, **83**, 1573–1577.
44. Masibay,A.S. and Qasba,P.K (1989) Expression of bovine β-1,4-galactosyltransferase cDNA in COS-7 cells. *Proc. Natl Acad. Sci. USA*, **86**, 5733–5737.
45. Joziasse,D.H., Shaper,J.H., van den Eijnden,D.H., van Tunen,A.J. and Shaper,N.L (1989) Bovine α-1-3-galactosyltransferase: isolation and characterization of a cDNA clone. Identification of homologous sequences in human genomic DNA. *J. Biol. Chem.*, **264**, 14290–14297.
46. Narimatsu,H., Sinha,S., Brew,K., Okayama,H. and Qasba,P.K. (1986) Cloning and sequencing of cDNA of bovine N-acetylglucosamine. (β-1-4) galactosyltransferase. *Proc. Natl Acad. Sci. USA*, **83**, 4720–4724.
47. D'Agostaro,G., Bendiak,B. and Tropak,M. (1989) Cloning of cDNA encoding the membrane-bound form of bovine β-1,4-galactosyltransferase. *Eur. J. Biochem.*, **183**, 211–217.
48. Appert,H.E., Rutherford,T.J., Tarr,G.E., Wiest,J.S., Thomford,N.R. and McCorquodale,D.J. (1986) Isolation of a cDNA coding for human galactosyltransferase. *Biochem. Biophys. Res. Commun.*, **139**, 163–168.
49. Masri,K.A., Appert,H.E. and Fukuda,M.N. (1988) Identification of the full-length coding sequence for human galactosyltransferase (β-N-acetylglucosaminide: β-1,4-galactosyltransferase). *Biochem. Biophys. Res. Commun.*, **157**, 657–663.
50. Narimatsu,H. (1989) Unpublished.
51. Watzele,G. and Berger,E.G. (1990) Near identity of HeLa cell galactosyltransferase with the human placental enzyme. *Nucleic Acids Res.*, **18**, 7174.
52. Uejima,T., Uemura,M., Nozawa,S. and Narimatsu,H. (1992) Complementary DNA cloning for galactosyltransferase associated with tumor and determination of antigenic epitopes recognized by specific monoclonal antibodies. *Cancer Res.*, **52**, 6158–6163.
53. Chatterjee,S.K., Mukerjee,S. and Tripathi,P.K. (1994) Unpublished.
54. Humphreys-Beher,M.G., Bunnell,B., van Tuinen,P., Ledbetter,D.H. and Kidd,V.J. (1986) Molecular cloning and chromosomal localization of human 4-β-galactosyltransferase. *Proc. Natl Acad. Sci. USA*, **83**, 8918–8922; erratum: *Proc. Natl Acad. Sci. USA*, **86**, 8747, 1989.
55. Larsen,R.D., Rivera-Marrero,C.A., Ernst,L.K., Cummings,R.D. and Lowe,J.B. (1990) Frameshift 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.*, **265**, 7055–7061.
56. Mengle-Gaw,L., McCoy-Haman,M.F. and Tiemeier,D.C. (1991) Genomic structure and expression of human beta-1,4-galactosyltransferase. *Biochem. Biophys. Res. Commun.*, **176**, 1269–1276.
57. Galili,U. and Swanson,K. Genomic structure and expression of human beta-1,4-galactosyltransferase. *Proc. Natl Acad. Sci. USA*, **88**, 7401–7404.
58. Yamamoto,F.i., McNeill,P.D. and Hakomori,S.I. (1991) Identification in human genomic DNA of the sequence homologous but not identical to

either the histo-blood group ABH genes or α1-3 galactosyltransferase pseudogene. *Biochem. Biophys. Res. Commun.*, **175**, 986–994.
59. Joziasse,D.H., Shaper,J.H., Wang,J.E. and Shaper,N.L. (1991) Characterisation of an α-1-3 galactosyltransferase homologue on human chromosome 12 that is organised as a processed pseudogene. *J. Biol. Chem.*, **266**, 6991–6998.
60. Shaper,N.L., Hollis,G.F., Douglas,J.G., Kirsch,I.R. and Sharper,J.H. (1988) Characterization of the full length cDNA for murine beta-1,4-galactosyltransferase. Novel features at the 5′-end predict two translational start sites at two in-frame AUGs. Cloning and sequencing of a full-length cDNA of mouse N-acetylglucosamine (beta 1-4)galactosyltransferase. *J. Biol. Chem.*, **263**, 10420–10428.
61. Harduin-Lepers,A., Shaper,N.L., Mahoney,J.V. and Shaper,J.H. (1992) Murine beta 1,4-galactosyltransferase: round spermatid transcripts are characterized by an extended 5′-untranslated region. *Glycobiology*, **2**, 361–368.
62. Harduin-Lepers,A., Shaper,J.H. and Shaper,N.L. (1993) Characterization of two cis-regulatory regions in the murine beta 1,4-galactosyltransferase gene. Evidence for a negative regulatory element that controls initiation at the proximal site. *J. Biol. Chem.*, **268**, 14348–14359.
63. Nakazawa,K., Ando,T., Kimura,T. and Narimatsu,H. (1988) Cloning and sequencing of a full-length cDNA of mouse N-acetylglucosamine (β1-4) galactosyltransferase. *J. Biochem.*, **104**, 165–168.
64. Schulte,S. and Stoffel,W. (1993) Ceramide UDP galactosyltransferase from myelinating rat brain: purification, cloning and expression. *Proc. Natl Acad. Sci. USA*, **90**, 10265–10269.
65. Stahl,N., Jurevics,H., Morell,P., Suzuki,K. and Popko,B. (1994) Isolation, characterization, and expression of cDNA clones that encode rat UDP-galactose: ceramide galactosyltransferase. *J. Neurosci. Res.*, **38**, 234–242.
66. Hollis,G.F., Douglas,J.G., Shaper,N.L., Shaper,J.H., Stafford-Hollis,J.M., Evans,R.J. and Kirsch,I.R. (1989) Genomic structure of murine β-1,4-galactosyltransferase. *Biochem. Biophys. Res. Commun.*, **162**, 1069–1075.
67. Shaper,J.H., Hollis,G.F. and Shaper,N.L. (1988) Genomic structure of murine beta-1,4-galactosyltransferase. *Biochimie*, **70**, 1683–1688.
68. Ghosh,S., Basu,S.S. and Basu,S. (1992) Isolation of a cDNA clone for beta 1-4 galactosyltransferase from embryonic chicken brain and comparison to its mammalian homologs. *Biochem. Biophys. Res. Commun.*, **189**, 1215–1222.
69. Modak,S.P. (1989) Unpublished.
70. Larsen,R.D., Rivera-Marrero,C.A., Ernst,L.K., Cummings,R.D. and Lowe,J.B. (1990) Frameshift and nonsense mutations in a human genomic sequence homologous to a murine UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-galactosyltransferase cDNA. *J. Biol. Chem.*, **265**, 7055–7061.
71. Shaper,N.L., Hollis,G.F., Douglas,J.G., Kirsch,I.R. and Shaper,J.H. (1988) Characterization of the full length cDNA for murine beta-1,4-galactosyltransferase. Novel features at the 5′-end predict two translational start sites at two in-frame AUGs. *J. Biol. Chem.*, **263**, 10420–10428.
72. Shaper,N.L., Wright,W.W. and Shaper,J.H. (1990) Murine beta 1,4-galactosyltransferase: both the amounts and structure of the mRNA are regulated during spermatogenesis. *Proc. Natl Acad. Sci. USA*, **87**, 791–795.
73. Larsen,R.D., Rajan,V.P., Ruff,M.M., Kukowska-Latallo,J., Cummings, R.D. and Lowe,J.B. (1989) 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*, **86**, 8227–8231.
74. Sulston,J. *et al.* (1992) The *C. elegans* genome sequencing project: a beginning [see comments]. *Nature*, **356**, 37–41.
75. Chappell,T.G., Hajibagheri,N.M., Ayscough,K., Pierce,M. and Warren,G. (1994) Localization of an alpha 1,2 galactosyltransferase activity to the Golgi apparatus of *Schizosaccharomyces pombe*. *Mol. Biol. Cell*, **5**, 519–528.
76. Joziasse,D.H., Shaper,N.L., Kim,D., van den Eijnden,D.H. and Shaper,J.H. (1992) Murine α-1,3-galactosyltransferase. A single gene locus specifies four isoforms of the enzyme by alternative splicing. *J. Biol. Chem.*, **267**, 5534–5541.
77. Uehara,K. and Muramatsu,T. (1994) Unpublished.
78. Ryan,K.A., Garraway,L.A., Descoteaux,A., Turco,S.J. and Beverley,S.M. (1993) Isolation of virulence genes directing surface glycosyl-phosphatidylinositol synthesis by functional complementation of Leishmania. *Proc. Natl Acad. Sci. USA*, **90**, 8609–8613.
79. Warren,C.E., Smookler,D.S. and Dennis,J.W. (1994) Unpublished.
80. Shoreibah,M., Perng,G.S., Adler,B., Weinstein,J., Basu,R., Cupples,R., Wen,D., Browne,J.K., Buckhaults,P., Fregien,N. and Pierce,M. (1993) Isolation, characterization, and expression of a cDNA encoding N-acetylglucosaminyltransferase V. *J. Biol. Chem.*, **268**, 15381–15385.
81. Pownall,S., Kozak,C.A., Schappert,K., Sarkar,M., Hull,E., Schachter,H. and Marth,J.D. (1992) Molecular cloning and characterization of the mouse UDP-N-acetylglucosamine:alpha-3-D-mannoside β-1,2-N-acetylglucosaminyltransferase I gene. *Genomics*, **12**, 699–704.
82. Kumar,R., Yang,J., Eddy,R.L., Byers,M.G., Shows,T.B. and Stanley,P. (1992) Cloning and expression of the murine gene and chromosomal location of the human gene encoding N-acetylglucosaminyltransferase I. *Glycobiology*, **2**, 383–393.
83. Kumar,R., Yang,J., Larsen,R.D. and Stanley,P. (1990) Cloning and expression of N-acetylglucosaminyltransferase I, the medial Golgi transferase that initiates complex N-linked carbohydrate formation. *Proc. Natl Acad. Sci. USA*, **87**, 9948–9952.
84. Bierhuizen,M.F. and Fukuda,M. (1992) Expression cloning of a cDNA encoding UDP-GlcNAc:Gal β-1-3-GalNAc-R (GlcNAc to GalNAc) β-1-6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen. *Proc. Natl Acad. Sci. USA*, **89**, 9326–9330.
85. Rajput,B., Ma,J. and Vijay,I K. (1993) Unpublished.
86. Deleted in proof.
87. Lehrman,M.A., Zhu,X. and Khounlo,S. (1988) Amplification and molecular cloning of the hamster tunicamycin-sensitive N-acetylglucosamine-1-phosphate transferase gene. The hamster and yeast enzymes share a common peptide sequence. *J. Biol. Chem.*, **263**, 19796–19803.
88. Zhu,X. and Lehrman,M.A. (1990) Cloning, sequence, and expression of a cDNA encoding hamster UDP-GlcNAc:dolichol phosphate N-acetylglucosamine-1-phosphate transferase. *J. Biol. Chem.*, **265**, 14250–14255.
89. Rajput,B., Ma,J., Muniappa,N., Schantz,L., Naylor,S.L., Lalley,P.A. and Vijay,I.K. (1992) Mouse UDP-GlcNAc: dolichyl-phosphate N-acetylglucosaminephosphotransferase. Molecular cloning of the cDNA, generation of anti-peptide antibodies and chromosomal localization. *Biochem. J.*, **285**, 985–992.
90. Nishikawa,A., Saito,H., Gu,J., Ihara,Y., Soejima,H., Wada,Y., Sekiya,C., Kangawa,K., Niikawa,N. and Taniguchi,N. (1994) cDNA cloning and chromosomal mapping of human N-acetylglucosaminyltransferase V+. *Biochem. Biophys. Res. Commun.*, **198**, 318–327.
91. Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C. (1993) Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nature Genet.*, **4**, 373–380.
92. Wilson,R. *et al.* (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature*, **368**, 32–38.
93. Sarkar,M., Hull,E., Nishikawa,Y., Simpson,R.J., Moritz,R.L., Dunn,R. and Schachter,H. (1991) Molecular cloning and expression of cDNA encoding the enzyme that controls conversion of high-mannose to hybrid and complex N-glycans: UDP-N-acetylglucosamine: α-3-D-mannoside β-1,2-N-acetylglucosaminyltransferase I. *Proc. Natl Acad. Sci. USA*, **88**, 234–238.
94. Bierhuizen,M.F.A., Mattei,M.G. and Fukuda,M. (1993) 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.*, **7**, 468–478.
95. Bierhuizen,M.F. and Fukuda,M. (1993) Unpublished.
96. Ihara,Y., Nishikawa,A., Toma,T., Soejima,H., Niikawa,N. and Taniguchi, N. (1993) cDNA cloning, expression and chromosomal localisation of human N-acetylglucosaminyltransferase III (GnT-III). *J. Biochem.*, **113**, 692–698.
97. Hull,E., Sarkar,M., Spruijt,M.P., Hoppener,J.W., Dunn,R. and Schachter, H. (1991) Organization and localization to chromosome 5 of the human UDP-N-acetylglucosamine:alpha-3-D-mannoside β-1,2-N-acetylglucosaminyltransferase I gene. *Biochem. Biophys. Res. Commun.*, **176**, 608–615.
98. Fukada,T., Iida,K., Kioka,N., Sakai,H. and Komano,T. (1994) Cloning of cDNA encoding N-acetylglucosyltransferase I from rat liver and analysis of its expression in rat tissues. *Biosci. Biotechnol. Biochem.*, **58**, 200–201.
99. Nishikawa,A., Ihara,Y., Hatakeyama,M., Kangawa,K. and Taniguchi,N. (1992) Purification, cDNA cloning, and expression of UDP-N-acetylglucosamine: β-D-mannoside β-1,4N-acetylglucosaminyltransferase III from rat kidney. *J. Biol. Chem.*, **267**, 18199–18204.
100. Bakker,H., Agterberg,M., Van Tetering,A., Koeleman,C.A.M., Van den Eijnden,D.H. and Van Die,I. (1994) A *Lymnaea stagnalis* gene, with sequence similarity to that of mammalian β1-4 galactosyltransferases, encodes a novel UDP-GlcNAc:GlcNAc β-R β1-4N-acetylglucosaminyltransferase. *J. Biol. Chem.*, **269**, 30326–30333.
101. Liu,X. and Chang,K.P. (1992) The 63 kilobase circular amplicon of a tunicamycin-resistant *Leishmania amazonensis* contains a functional gene that can be used as a dominant selectable marker in transfection. *Mol. Cell. Biol.*, **12**, 4112–4122.
102. Hartog,K.O. and Bishop,B. (1987) Genomic sequence coding for tunicamycin resistance in yeast. *Nucleic Acids Res.*, **15**, 3627.

103. Zou,J., Scocca,J.R. and Krag,S.S. (1994) Unpublished.
104. Yamamoto,F.i., Marken,J., Tsuji,T., White,T., Clausen,H. and Hakomori, S.I. (1989) Cloning and characterization of DNA complementary to human UDP-GalNAc: Fuc α1-2Gal α1-3GalNAc transferase (histo-blood group A transferase) mRNA. *J. Biol. Chem.*, **265**, 1146–1151.
105. Nagata,Y., Yamashiro,S., Yodoi,J., Lloyd,K.O., Shiku,H. and Furukawa, K. (1992) Expression cloning of β1-4 N-acetylgalactosaminyltransferase cDNAs that determine the expression of GM2 and GD2 gangliosides. *J. Biol. Chem.*, **267**, 12082–12089; erratum: *J. Biol. Chem.*, **269**, 7045, 1994.
106. Homa,F.L., Hollander,T., Lehman,D.J., Thomsen,D.R. and Elhammer, A.P. (1993) Uupublished.
107. Hagen,F.K., Van Wuyckhuyse,B. and Tabak,L.A. (1993) Purification, cloning, and expression of a bovine UDP-GalNAc: polypeptide N-acetyl-galactosaminyltransferase. *J. Biol. Chem.*, **268**, 18960–18965.
108. Strahl-Bolsinger,S., Immervoll,T., Deutzmann,R. and Tanner,W. (1993) PMT1, the gene for a key enzyme of protein O-glycosylation in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **90**, 8164–8188.
109. Orlean,P., Albright,C.F. and Robbins,P.W. (1988) Cloning and sequencing of the yeast gene for dolichol phosphate mannose synthase, an essential protein. *J. Biol. Chem.*, **263**, 17499–17507.
110. Barrell,B.G. *et al. Saccharomyces cerevisiae* chromosome IX sequencing project, Sanger Centre, Hinxton Hall, Cambridge, UK and MRC Laboratory of Molecular Biology, Cambridge, UK.
111. Connor,R. and Churcher,C. (1994) Unpublished.
112. Albright,C.F. and Robbins,P.W. (1990) The sequence and transcript heterogeneity of the yeast gene ALG1, an essential mannosyltransferase involved in N-glycosylation. *J. Biol. Chem.*, **265**, 7042–7049.
113. Lussier,M., Camirand,A., Sdicu,A.M. and Bussey,H. (1993) KTR2: a new member of the KRE2 mannosyltransferase gene family. *Yeast*, **9**, 1057–1063.
114. Hausler,A. and Robbins,P.W. (1992) Glycosylation in *Saccharomyces cerevisiae*: cloning and characterization of an α-1,2-mannosyltransferase structural gene. *Glycobiology*, **2**, 77–84.
115. Nakayama,K.I., Nagasu,T., Shimma,Y.i., Kuromitsu,J.r. and Jigami,Y. (1992) OCH1 encodes a novel membrane bound mannosyltransferase: outer chain elongation of asparagine-linked oligosaccharides *EMBO J*, **11**, 2511–2519.
116. Waterston,R. *et al.* (1992) A survey of expressed genes in *Caenorhabditis elegans*. *Nature Genet.*, **1**, 114–123.
117. Jackson,B.J., Kukuruzinska,M.A. and Robbins,P. (1993) Biosynthesis of asparagine-linked oligosaccharides in *Saccharomyces cerevisiae*: the alg2 mutation. *Glycobiology*, **3**, 357–364.
118. Stagljar,I., te Heesen,S. and Aebi,M. (1994) New phenotype of mutations deficient in glucosylation of the lipid-linked oligosaccharide: cloning of the ALG8 locus. *Proc. Natl Acad. Sci. USA*, **91**, 5977–5981.
119. te Heesen,S., Lehle,L., Weissman,A. and Aebi,M. (1994) Isolation of the ALG5 locus encoding the UDP-glucose: dodecyl-phosphate glucosyl-transferase from *S.cerevisiae*. *Eur. J. Biochem.*, **224**, 71–79.
120. Schachter,H. (1994) In Fukuda,M. and Hindsgaul,O. (eds), *Molecular Glycobiology*. IRL Press, Oxford.
121. Ferdinand,W. (1976) *The Enzyme Molecule*. J.Wiley and Sons Ltd, UK.
122. Lowe,J.B. (1994) In Fukuda,M. and Hindsgaul,O. (eds), *Molecular Glycobiology*. IRL Press, Oxford.
123. Lance,P., Lau,K.M. and Lau,J.T. (1989) Isolation and characterization of a partial cDNA for a human sialyltransferase. *Biochem. Biophys. Res. Commun.*, **164**, 225–232.
124. Yamamoto,F., McNeill,P.D. and Hakomori,S. (1992) Human histo-blood group A2 transferase coded by A2 allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. *Biochem. Biophys. Res. Commun.*, **187**, 366–374.
125. Kornfeld,R. and Kornfeld,S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.*, **54**, 631–664.
126. van Meer,G. (1993) Transport and sorting of membrane lipids. *Curr. Opin. Cell Biol.*, **5**, 661–667.
127. Vidugiriene,J. and Menon,A.K. (1994) The GPI anchor of cell-surface proteins is synthesized on the cytoplasmic face of the endoplasmic reticulum. *J. Cell Biol.*, **127**, 333–341.
128. Kukuruzinska,M.A., Bergh,M.L. and Jackson,B.J. (1987) Protein glycosylation in yeast. *Annu. Rev. Biochem.*, **56**, 915–944.
129. van Meer,G. (1989) Lipid traffic in animal cells. *Annu. Rev. Cell Biol.*, **5**, 247–264.
130. Field,M.C. and Menon,A.K. (1993) In Schlesinger (ed.), *Lipid Modifications of Proteins*. CRC Press, Boca Raton, FL.
131. Englund,P.T. (1993) The structure and biosynthesis of glycosyl phosphatidylinositol protein anchors. *Annu. Rev. Biochem.*, **62**, 121–135.
132. Munro,S. (1991) Sequences within and adjacent to the transmembrane segment of alpha-2,6-sialyltransferase specify Golgi retention. *EMBO J.*, **10**, 3577–3588.
133. Nilsson,T., Lucocq,J.M., Mackay,D. and Warren,G. (1991) The membrane spanning domain of β-1,4-galactosyltransferase specifies trans Golgi localization. *EMBO J.*, **10**, 3567–3575.
134. Schachter,H. (1991) The 'yellow brick road' to branched complex N-glycans. *Glycobiology*, **1**, 453–461.
135. Devereux,J., Haeberli,P. and Smithies,O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.*, **12**, 387–395.
136. Swafford,D.L. (1993) *Phylogenetic Analysis Using Parsimony Version 3.1.1*. Illinois History Survey, Champaign, IL.
137. McConville,M. and Ferguson,M.A.J. (1993) The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. *Biochem. J.*, **294**, 305–324.
138. Yip,C.L., Welch,S.K., Klebl,F., Gilbert,T., Seidel,P., Grant,F.J., O'Hara, P.J. and MacKay,V.L. (1994) Cloning and analysis of the *Saccharomyces cerevisiae* MNN9 and MNN1 genes required for complex glycosylation of secreted proteins. *Proc. Natl Acad. Sci. USA*, **91**, 2723–2727.
139. Hill,K., Boone,C., Goebel,M., Puccia,R., Sdicu,A.M. and Bussey,H. (1992) Yeast KRE2 defines a new gene family encoding probable secretory proteins, and is required for the correct N-glycosylation of proteins. *Genetics*, **130**, 273–283.
140. Oullette,F., Clarke,M.W., Keng,T., Storms,R.K., Zhong,W. *et al.* (1993) Sequencing of chromosome I from *S.cerevisiae*: analysis of a 32 kb region between the LTE1 and SPO7 genes. *Genome*, **36**, 32–42.
141. Natsuka,S. and Lowe,J.B. (1994) Enzymes involved in mammalian glycosylation. *Curr. Opin. Struct. Biol.*, **4**, 683–691.
142. Ioffe,E. and Stanley,P. (1994) Mice lacking N-acetylglucosaminyl-transferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. *Proc. Natl Acad. Sci. USA*, **91**, 728–732.

Received on April 24, 1995; accepted on May 12, 1995

Note Added in Proof

The W3 site at "http://bellatrix.pcl.ox.ac.uk/people/barry_spring95/iains_transferase_guide.html" constructed by Dr Iain Wilson contains hypertext links to GenBank for many of the entries listed in Table I.