

RESEARCH ARTICLE

Galactosyltransferase 4 is a major control point for glycan branching in N-linked glycosylation

Andrew G. McDonald^{1,2,*}, Jerrard M. Hayes^{1,2,*}, Tania Bezak^{1,2}, Sonia A. Głuchowska¹, Eoin F. J. Cosgrave², Weston B. Struwe², Corné J. M. Stroop³, Han Kok³, Teun van de Laar³, Pauline M. Rudd², Keith F. Tipton¹ and Gavin P. Davey^{1,†}

ABSTRACT

Protein N-glycosylation is a common post-translational modification that produces a complex array of branched glycan structures. The levels of branching, or antennarity, give rise to differential biological activities for single glycoproteins. However, the precise mechanism controlling the glycan branching and glycosylation network is unknown. Here, we constructed quantitative mathematical models of N-linked glycosylation that predicted new control points for glycan branching. Galactosyltransferase, which acts on *N*-acetylglucosamine residues, was unexpectedly found to control metabolic flux through the glycosylation pathway and the level of final antennarity of nascent protein produced in the Golgi network. To further investigate the biological consequences of glycan branching in nascent proteins, we glycoengineered a series of mammalian cells overexpressing human chorionic gonadotropin (hCG). We identified a mechanism in which galactosyltransferase 4 isoform regulated N-glycan branching on the nascent protein, subsequently controlling biological activity in an *in vivo* model of hCG activity. We found that galactosyltransferase 4 is a major control point for glycan branching decisions taken in the Golgi of the cell, which might ultimately control the biological activity of nascent glycoprotein.

KEY WORDS: Glycosylation, Golgi, branching, Galactosyltransferase, Chorionic gonadotropin, hCG

INTRODUCTION

N-linked glycosylation is a protein modification system present in all domains of life that is characterised by extreme complexity and a seemingly random control process. Over 60% of proteins, typically cell surface receptors and transporters, hormones, cytokines and antibodies are post-translationally modified by the addition of glycan structures as they are processed through the endoplasmic reticulum and Golgi. As the N-glycans are trimmed and re-modelled in transit through the Golgi, multiple glycoforms

of a glycoprotein are generated by a process that is partially understood. The *N*-acetylglucosaminyltransferase enzymes (GnTs; GnTI, GnTII, GnTIV and GnTV) utilise UDP-*N*-acetylglucosamine (UDP-GlcNAc) as a donor substrate and transfer the GlcNAc onto the glycoform at different branching points. Flux through the hexosamine and sugar-nucleotide pathways are known to influence this process as they modulate the levels of UDP-GlcNAc substrate available for branching (Sasai et al., 2002; Schachter, 1986). Further modification of the branched glycoforms involves galactosylation by the β -1,4-galactosyltransferases (GalTs), elongation with poly-*N*-acetylglucosamine by β -1,3-*N*-acetylglucosaminyltransferase (GnTE) and capping with sialic acid by α -2,3-sialyltransferase (SiaT). Although it is known that hexosamine flux regulates the degree of branching and number of complex N-glycans, which in turn cooperate to regulate cell proliferation and differentiation (Lau et al., 2007), the complex nature of interactions between the glycosyltransferases that construct the glycan repertoire on glycoproteins is still unresolved (Lowe and Marth, 2003).

The importance of post-translational glycosylation for the effectiveness of biotherapeutic proteins has become increasingly recognised for its role in bio-activity, serum half-life, stability and immunogenicity (Dalziel et al., 2014; Jefferis, 2009; Perlman et al., 2003; Sinclair and Elliott, 2005; Solá and Griebenow, 2009). However, controlling the glycosylation of recombinant proteins to obtain favourable glycoforms is a substantial challenge. Glycoengineering mammalian cells has led to the development of monoclonal antibodies lacking core fucosylation, resulting in a 50–100-fold higher affinity for Fc γ receptors and a concomitant increase in antibody-dependent cellular cytotoxicity (ADCC) (Ferrara et al., 2011; Ferrara et al., 2006; Iida et al., 2006; Shields et al., 2002; Shinkawa et al., 2003; Umaña et al., 1999). Increased glycan antennarity and sialic acid capping of glycoproteins prevents their premature removal from circulation by the asialoglycoprotein receptor on hepatocytes (Ashwell and Morell, 1974). In addition, it has been shown that glycoengineering of erythropoietin to produce hyper-glycosylated and sialylated forms of the protein substantially increased bioactivity and serum half-life (Egrie et al., 2003). Further understanding of the complex process controlling glycan antennarity, galactosylation and sialylation in the Golgi is required.

In this article, we describe computer-generated models and simulations of N-linked glycosylation in mammalian cells in which we identify new control points where glycosylation can be modified. These predictions were then validated in a series of host cell lines expressing a heavily glycosylated glycoprotein hormone, human chorionic gonadotropin (hCG). Our experiments identify a new feature of the regulation of kinetic flux through the

¹School of Biochemistry and Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin 2, Ireland. ²National Institute for Bioprocessing Research and Training (NIBRT), Fosters Avenue, Dublin 4, Ireland. ³Merck, Sharp & Dohme, 5340 BH Oss, The Netherlands.

*These authors contributed equally to this work

†Author for correspondence (gdavey@tcd.ie)

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

glycosylation system; specifically, that galactosyltransferase activity controls N-glycan branching and identifies a means through which variant glycoforms of a single protein give rise to different effects on biological activity.

RESULTS

In silico modelling of glycan post-translational modifications

We constructed a mathematical model of N-linked glycosylation in Chinese hamster ovary (CHO) cells based on the actions of eleven enzymes that are known to be present in the CHO Golgi (see Table 1): two mannosidases, ManI and ManII; the (core) fucosyltransferase, FucT; six *N*-acetylglucosaminyltransferases, GnTI, GnTII, GnTIII, GnTIV, GnTV and GnTE; a galactosyltransferase, GalT; and a sialyltransferase, SiaT. The core network of reactions is shown in Fig. 1A. The details of the mathematical models, which were based on an earlier model by Umaña and Bailey (Umaña and Bailey, 1997) and its extension by Krambeck and Betenbaugh (Krambeck and Betenbaugh, 2005), are given in the Materials and Methods. To elucidate glycan complexity, we developed two software programs, GlycoForm and Glycologue (McDonald et al., 2010), for the display of N-glycans.

Our objective was to determine the key enzymes controlling the production of tetra-antennary glycoforms. To this end, the theoretical models were used to test, *in silico*, the under- and over-expression (compared with their mean *in vivo* concentration) of each enzyme. A preliminary series of *in silico* experiments, in which each of the enzymes of CHO N-glycosylation was eliminated in turn, was used to determine the effects of individual enzyme knockouts on antennarity. The results, which are summarised in Fig. 1B–G, show that the enzymes GalT and GnTE are responsible for much of the glycoform heterogeneity.

We developed methods for the construction of mathematical models of glycosylation in thermodynamically open or closed systems, with or without the consideration of the levels of donor sugar-nucleotides, for reactions occurring within either a homogeneous single Golgi compartment or one compartmentalised into four spatial regions (or temporal phases), termed cis, medial, endo and trans-Golgi network (TGN). For simplicity, we considered a theoretical protein with a single N-glycosylation site. For open-system models it was assumed that protein substrate was introduced at a constant rate, and we assumed a mean residence time of 20 min within the whole Golgi. The results we present here are for a single-compartment and single-substrate open system, although qualitatively similar results were obtained when these restrictions were removed. Further assumptions of the model were: (i) the essential irreversibility of the reaction network; (ii) a constant supply of the donor sugar-nucleotides

CMP-Neu5Ac, GDP-Fuc, UDP-Gal and UDP-GlcNAc; and (iii) constant cell and organelle volumes. Sugar-nucleotides are transferred to the Golgi by means of dedicated transporters (Hirschberg et al., 1998), which also act as antiporters, moving monophosphorylated nucleotides from the Golgi to the cytoplasm. Justification for the irreversibility criterion is provided by the existence of a nucleoside-diphosphatase (NDPase) activity within the Golgi (Hirschberg et al., 1998), which maintains the secondary products of the transferase reactions at a low level, thereby reducing the effects of product inhibition. GnTIII was omitted from the kinetic model, because it is not expressed in wild-type CHO cells (Schriebl et al., 2006; Stanley et al., 2005), and the modelling results indicated that its absence would not adversely affect the production of tetra-antennary structures.

Given that published data are not available for the K_m values of each enzyme for the large variety of acceptors, we used averaged values obtained from the literature (Krambeck and Betenbaugh, 2005), and substituted our own experimentally determined K_m values for the donors (nucleotide sugars), where available. In the absence of GnTIII, the number of unique N-glycans predicted by the theoretical model was 3677 in 11,541 separate reactions. Rather than follow the production of individual N-glycans, we found it convenient to create special model variables that expressed the concentrations of mono-, bi-, tri- and tetra-antennary glycoforms as molar percentages (Fig. 2). We computed the steady-state intracellular concentrations of each variable as a function of local enzyme concentration. The effects of knocking down selected enzymes individually are shown in Fig. 2A–F. Our model predicted only modest increases in tetra-antennary N-glycans when GnTs IV and V were overexpressed separately (Fig. 2E,F) or together (Fig. 2G). The most surprising, yet consistent finding, was that lowering the level of GalT was a major factor influencing the production of N-glycans of high antennarity (Fig. 2G).

A comparison of values taken from the literature indicates that the enzyme GalT (EC 2.4.1.38) has an apparent k_{cat}/K_m value that is considerably higher than the other enzymes involved in glycosylation. Although the K_m values will depend on the nature of the acceptor (Pâquet et al., 1984), and likewise those of k_{cat} , the quoted k_{cat} value for GalT is nonetheless at least four times higher than that of any other enzyme in the network. In Fig. 2G, GalT is knocked down, while the enzymes that have the greatest influence on tetra-antennarity, namely GnTIV and GnTV are overexpressed by 25-fold. This simulation revealed a trend whereby tetra-antennary production increased by ~10-fold when the local concentration of GalT was reduced below 10% of its assumed wild-type value (0.44 μ M), when all other enzyme concentrations

Table 1. Enzymes involved in N-glycosylation in CHO cells.

Abbreviation	EC number	Enzyme name	Donor
GalT	EC 2.4.1.38	β - <i>N</i> -acetylglucosaminylglycopeptide β -1,4-galactosyltransferase	UDP-Gal
FucT	EC 2.4.1.68	glycoprotein 6- α -L-fucosyltransferase	GDP-Fuc
GnTI	EC 2.4.1.101	α -1,3-mannosyl-glycoprotein 2- β - <i>N</i> -acetylglucosaminyltransferase	UDP-GlcNAc
GnTII	EC 2.4.1.143	α -1,6-mannosyl-glycoprotein 2- β - <i>N</i> -acetylglucosaminyltransferase	UDP-GlcNAc
GnTIII	EC 2.4.1.144	β -1,4-mannosyl-glycoprotein 4- β - <i>N</i> -acetylglucosaminyltransferase	UDP-GlcNAc
GnTIV	EC 2.4.1.145	α -1,3-mannosyl-glycoprotein 4- β - <i>N</i> -acetylglucosaminyltransferase	UDP-GlcNAc
GnTE	EC 2.4.1.149	<i>N</i> -acetylglucosaminide β -1,3- <i>N</i> -acetylglucosaminyltransferase	UDP-GlcNAc
GnTV	EC 2.4.1.155	α -1,6-mannosyl-glycoprotein 6- β - <i>N</i> -acetylglucosaminyltransferase	UDP-GlcNAc
SiaT	EC 2.4.99.6	<i>N</i> -acetylglucosaminide α -2,3-sialyltransferase	CMP-NeuAc
ManI	EC 3.2.1.113	mannosyl-oligosaccharide 1,2- α -mannosidase	–
ManII	EC 3.2.1.114	mannosyl-oligosaccharide 1,3-1,6- α -mannosidase	–

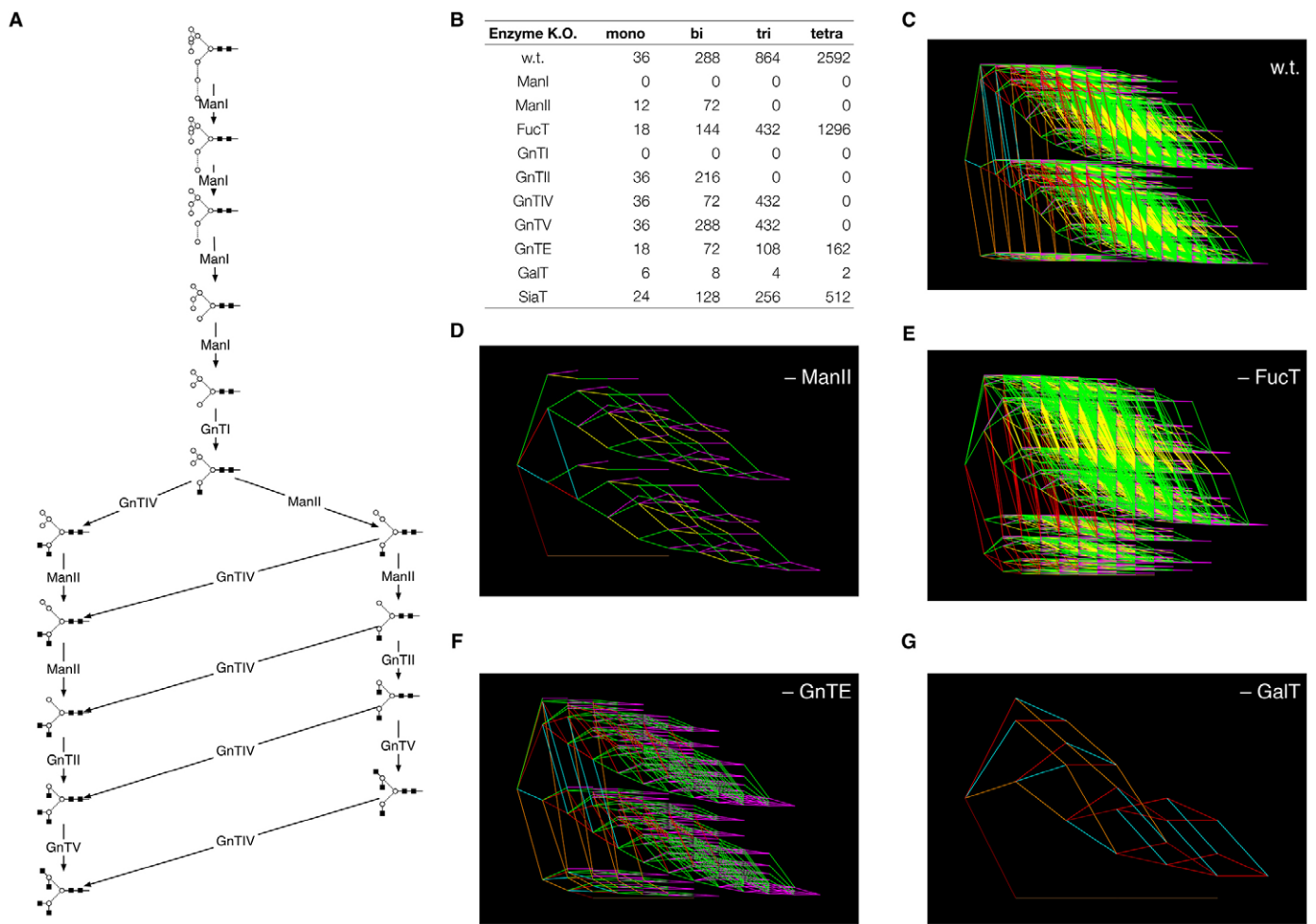


Fig. 1. *In silico* knockouts of enzymes of N-glycosylation and their effects upon reaction network complexity. (A) Core network of glycosylation, showing the initial stages of N-glycosylation starting with Man₉ (M9) and further catalysed by the enzymes described in Table 1. For symbol key, see Fig. 4 inset. (B) Total numbers of mono-, bi-, tri- and tetra-antennary N-glycan structures predicted by the model for wild-type (w.t.) and individual enzyme knockouts. (C–G) Reaction network diagrams corresponding to five of the knockouts shown in B. Colour key: GnTs I, II, IV and V, red; GnTE, yellow; GalT, green; ManI, brown; ManII, orange; FucT, cyan; SiaT, magenta.

were held constant. The results indicate that there is a flux-based competition between GalT and the other enzymes in the network. These results suggest that the overexpression of GnTIV and GnTV, in conjunction with a 90% reduction in Golgi concentrations of GalT, generates a distribution of N-glycans that is comprised of 56% tetra-antennary structures, with a high degree of sialylation (>80%). It was observed in simulation timecourses that a decrease in GalT level led to a delayed appearance of tetra-antennary forms. The results of Fig. 2G should therefore be considered the theoretical maximum. A theoretical study of N-glycosylation predicted that increases in galactosyltransferase activity can influence N-glycan microheterogeneity (Hossler et al., 2007). Our results are the first description of the effect of a decrease in such activity upon antennarity.

Glycan branching in hCG

We next verified the model predictions in a Chinese hamster ovary (CHO) cell line engineered to express recombinant hCG (CHO-HCG). Three derivative stable cell lines were engineered to overexpress the enzymes GnTIV (CHO-GnT4), GnTV (CHO-GnT5) and GnTIV and GnTV together (CHO-GnT4,5). Overexpression of GnTIV and GnTV was detected by

immunoblotting, and translocation to the Golgi was assessed by immunofluorescence microscopy. The relative amounts of GnT and GalT mRNA of each transformed cell line were determined by quantitative real-time PCR. GnTIV mRNA levels increased 15-fold and GnTV mRNA levels were increased 11-fold over that of the parent cell line. Small hairpin RNA (shRNA) constructs were used to knock down the galactosyltransferase genes GalT1 through GalT6 (Table 2).

The N-glycans of immunopurified hCG were removed by PNGase F digestion, fluorescently labelled with 2-aminobenzamide and analysed by hydrophilic interaction liquid chromatography (HILIC) and weak-anion exchange high-pressure liquid chromatography (WAX-HPLC), together with exoglycosidase digestions, according to Royle et al. (Royle et al., 2008). Preliminary structural assignments were made by comparison of the retention times of the N-glycans expressed in GU (glucose unit) values with those reported by the online software tool GlycoBase (<http://glycobase.nibrt.ie/>). Structures identified by HPLC were verified by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry following permethylation.

Initial HILIC profiles identified seven major 2-aminobenzamide (2-AB)-labelled N-glycan peaks from hCG

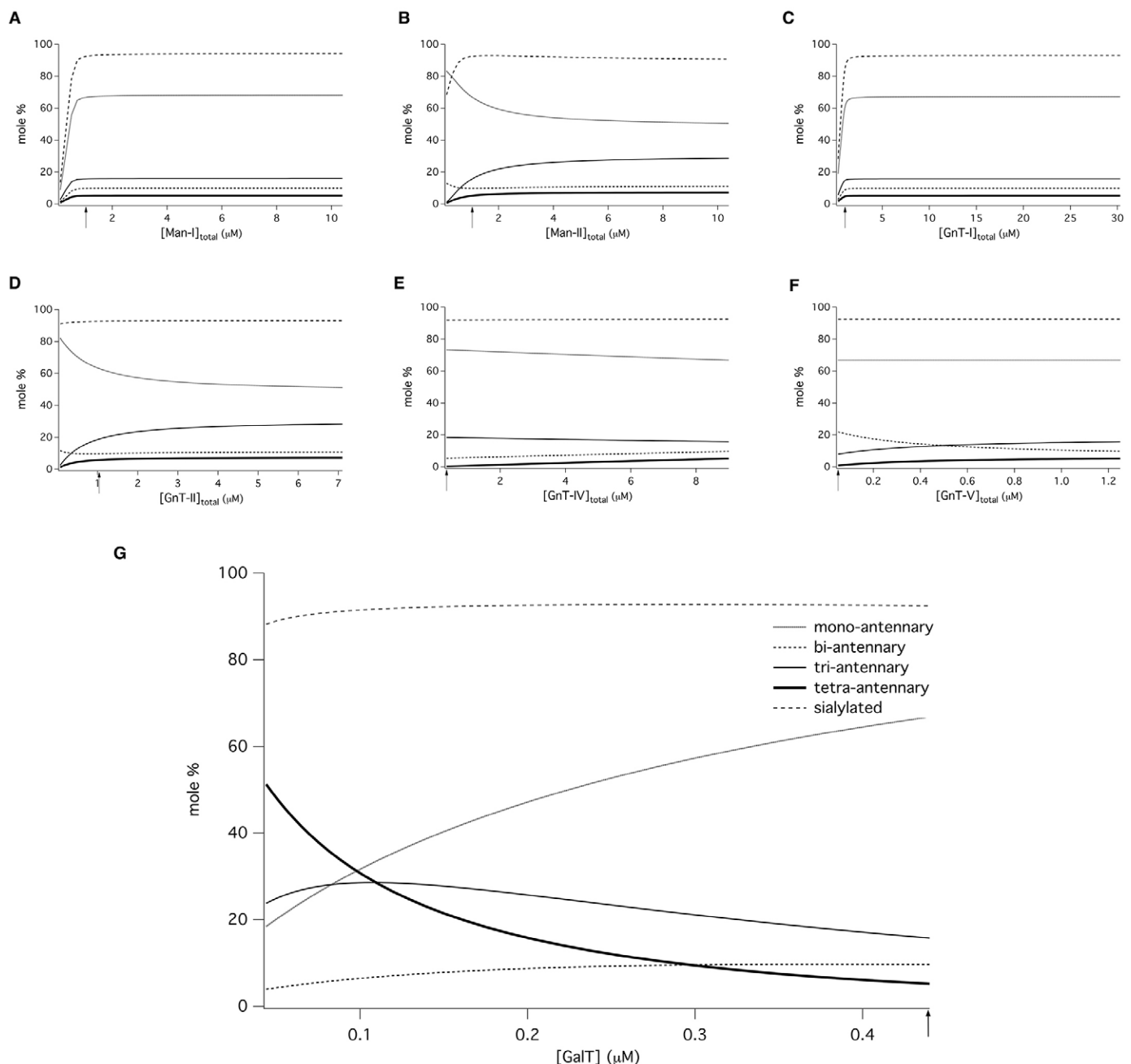


Fig. 2. Computational model predicts galactosyltransferase controls glycan antennarity. *In silico* variation of local enzyme concentrations: (A) ManI; (B) ManII; (C) GnTI; (D) GnTII; (E) GnTIV; (F) GnTV; (G) GalT. Final concentrations of *N*-glycan structures of different antennarities predicted by the kinetic model are plotted as a function of the total concentration of GalT, when combined with a 25-fold overexpression of GnTIV and GnTV. The position of the arrow on each abscissa indicates the default enzyme concentration. Values on each ordinate axis are expressed as a molar percentage of the total concentration of all *N*-glycan forms.

(Fig. 3A). No significant changes were observed in the *N*-glycan profiles from cells overexpressing GnTIV or GnTV alone (Fig. 3Aii,iii) or GnTIV and GnTV together or from cells with only GalT knockdowns (data not shown). Overexpression of GnTIV or GnTV in combination with knockdown of galactosyltransferase (GalT4) caused significant changes to the *N*-glycan profiles of hCG from these cell lines, particularly in the region of higher GU values (9–11) (Fig. 3Avii,viii). These glycosylation changes were not present for other galactosyltransferase knockdowns (Fig. 3Aiv, v,vi,x,xi). Dual overexpression of GnTIV and GnTV together with

GalT4 knockdown also resulted in substantial glycosylation changes to the substrate glycoprotein. Interestingly, a major increase was observed for a glycan structure at approximately GU 6.2 (see Fig. 3Aix).

WAX-HPLC was used to analyse the degree of *N*-glycan sialylation of hCG and revealed a substantial increase in tri-sialylated structures in cell lines overexpressing GnTIV or GnTV alone or GnTIV and GnTV together with GalT4 knockdown (Fig. 3Bvi,vii,viii). Bovine fetuin, which contains neutral, mono, bi, tri and tetra-sialylated glycans, was used as a standard for

Table 2. Quantitative RT-PCR results of CHO cell galactosyltransferase (GalT) knockdowns

Cell line	Gene	Percentage knockdown	Cell line	Gene	Percentage knockdown
CHO-HCG	GalT1	35	CHO-GnT5	GalT1	3
CHO-HCG	GalT2	69	CHO-GnT5	GalT2	56
CHO-HCG	GalT3	25	CHO-GnT5	GalT3	20
CHO-HCG	GalT4	62	CHO-GnT5	GalT4	63
CHO-HCG	GalT5	34	CHO-GnT5	GalT5	12
CHO-HCG	GalT6	84	CHO-GnT5	GalT6	39
CHO-GnT4	GalT1	60	CHO-GnT4,5	GalT1	80
CHO-GnT4	GalT2	38	CHO-GnT4,5	GalT2	56
CHO-GnT4	GalT3	55	CHO-GnT4,5	GalT3	27
CHO-GnT4	GalT4	60	CHO-GnT4,5	GalT4	79
CHO-GnT4	GalT5	56	CHO-GnT4,5	GalT5	50
CHO-GnT4	GalT6	29	CHO-GnT4,5	GalT6	77

GenBank identifiers for each gene are provided in the Materials and Methods.

WAX and enabled the identification of each sialic acid pool. Following WAX-HPLC analysis the total N-glycan pool was fractionated and neutral, mono-, bi-, tri- and tetra-sialylated structures were collected and further analysed by HILIC (supplementary material Fig. S1). This revealed a similar increase in sialylated structures in these cell lines, with the majority of N-glycans having one or two sialic acids with GU values of 7–9 in the parent CHO cell line and with two or three

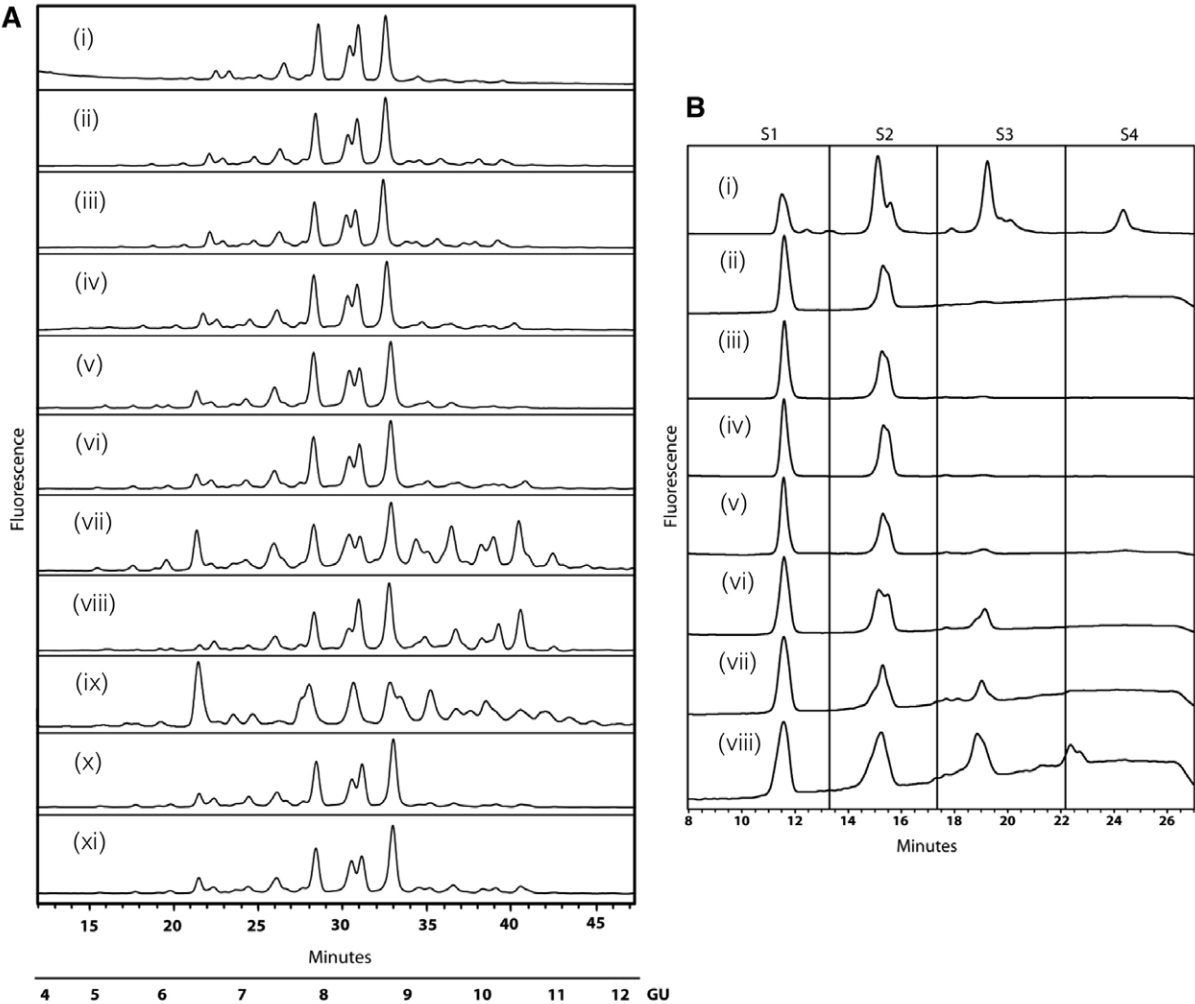


Fig. 3. GalT4 controls level of glycan antennarity and sialylation in N-linked glycosylation. (A) HILIC analysis of hCG immuno-purified from engineered cell lines shows that GalT4 knockdown causes substantial changes to the glycan profiles of the substrate glycoprotein (vii, viii, ix) suggesting that GalT4 is a major flux control point for glycan branching. (i) HCG, (ii) CHO-GnT4, (iii) CHO-GnT5, (iv) CHO-GnT4-GalT1, (v) CHO-GnT4-GalT2, (vi) CHO-GnT4-GalT3, (vii) CHO-GnT4-GalT4, (viii) CHO-GnT5-GalT4, (ix) CHO-GnT4,5-GalT4, (x) CHO-GnT4-GalT5, (xi) CHO-GnT4-GalT6. (B) GalT4 knockdown increases tri-sialylated glycan structures (S3) in CHO cells. WAX-HPLC of HCG from engineered cell lines and revealed a substantial increase in tri-sialylated structures in cell lines with GalT4 knockdown (v, vi, vii). (i) Bovine fetuin standard, (ii) HCG, (iii) CHO-GnT4, (iv) CHO-GnT5, (v) CHO-GnT4,5, (vi) CHO-GnT4-GalT4, (vii) CHO-GnT5-GalT4, (viii) CHO-GnT4,5-GalT4.

sialic acids (GU 8–10) in the engineered cell lines. In all cases, the sialic acids were in an α 2,3 linkage to galactose, as shown by digestion with *Streptococcus pneumoniae* (Nan1) sialidase, a recombinant exoglycosidase specific for α -2,3-linked sialic acids.

The HILIC profile of the undigested pool of N-glycans from the CHO-GnT4-GalT4 cell line showed a large number of overlapping peaks with high GU values, owing to the large number of differently sialylated species present (Fig. 4B). This was not seen for hCG from the parent cell line (Fig. 4A). Exoglycosidase digestions with enzymes used alone or in combinations followed by HILIC were used to elucidate the structures of the constituent monosaccharides. This was performed on the total N-glycan pool (supplementary material Fig. S2) and on the WAX fractionated samples. This is a two-dimensional glycan analysis using WAX in the first dimension and HILIC in the second dimension as a way to deconvolute complex glycan pools for structural assignments. The subsequent analysis revealed that the majority of N-glycans of hCG expressed in CHO cells were predominantly complex bi-antennary glycans, A2G(4)2, F(6)A2G(4)2, A2G(4)2S(3)1 and

F(6)A2G(4)1S(3)1 [as presented in Oxford notation, for more information, see <http://glycobase.nibrt.ie/glycobase/documents/abbreviations.pdf>], comprising ~8.5%, 21.5%, 26% and 23.5% of the total N-glycans, respectively (Fig. 4A). Large decreases in these bi-antennary glycans with concomitant increases in tri-antennary structures, particularly F(6)A3G(4)3, F(6)A3G(4)2S(3)2, F(6)A3G(4)3S(3)1, A3G3S2 were present in the CHO-GnT4-GalT4 cell line (Fig. 4B). Results of the N-glycan analysis of hCG are summarised in Table 3 and the proposed structures are shown in supplementary material Table S1. Interestingly, substantial increases were also observed in the N-glycan Man₅ [Man₅ refers to the core N-glycan plus an additional two mannose residues, see <http://glycobase.nibrt.ie/glycobase/documents/abbreviations.pdf>], which increased ~10-fold, from 0.4% to 4%, in this cell line. Substantial decreases in bi-antennary structures and increases in tri-antennary structures were also observed in the CHO-GnT5-GalT4 cell line. A similar effect was observed when both GnTIV and GnTV were co-overexpressed with GalT4 knockdown and a further substantial increase was seen in Man₅, which increased to ~15% of the total N-glycans of hCG from this cell line. Tetra-antennary structures in

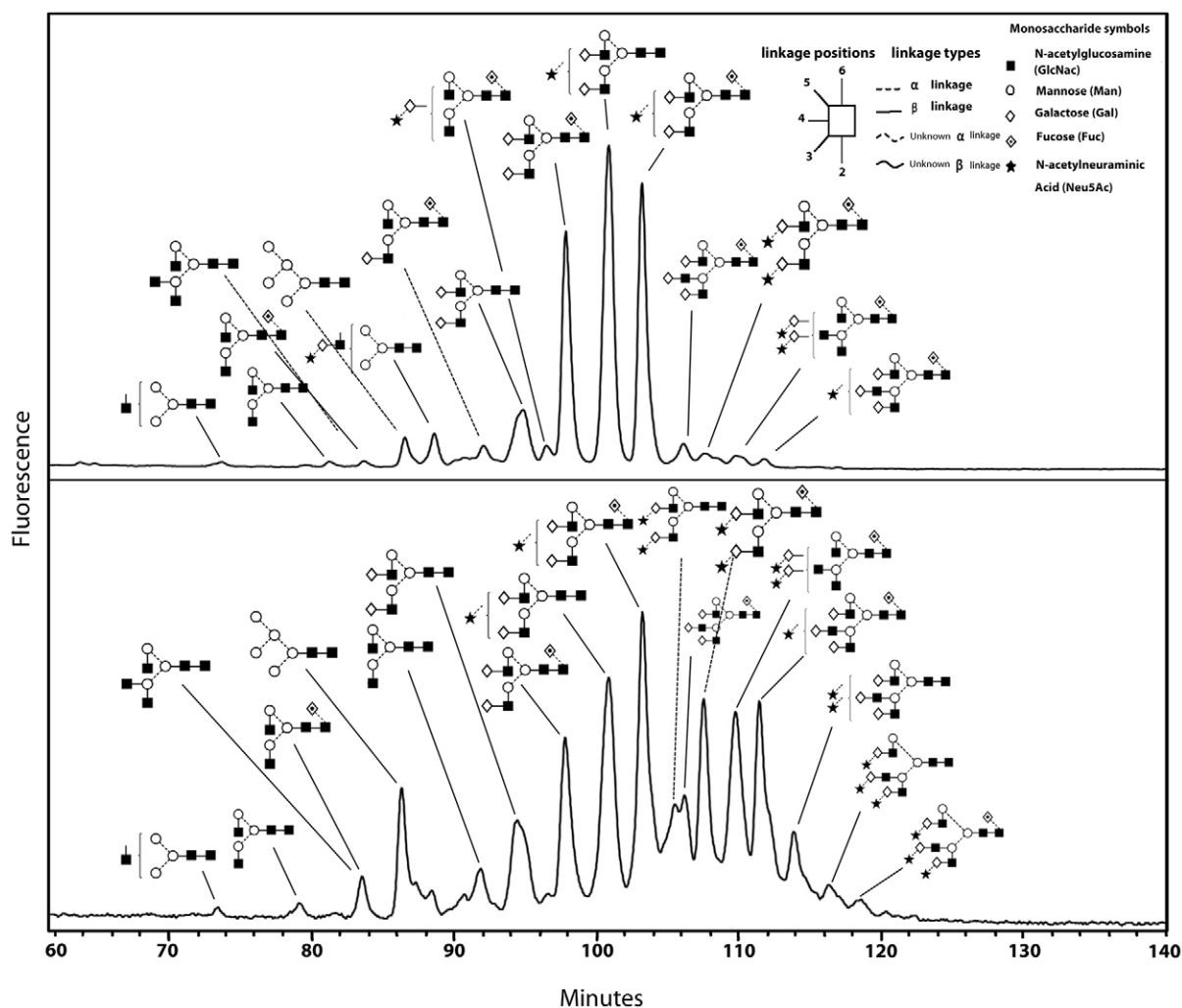


Fig. 4. GalT4 regulates glycan complexity in substrate glycoproteins. Substantial increases in tri-antennary and tri-sialylated structures were observed upon GalT4 knockdown. Glycan profiles were measured by HILIC and glycan structures were determined by a combination of exoglycosidase digestions and WAX-HPLC. (A) hCG immuno-purified from CHO cells. (B) hCG immuno-purified from CHO-GnT4-GalT4 cells. Nomenclature of N-glycans is shown in the inset together with a guide to N-glycan monosaccharide symbols, linkage types and positions. For further information on the Oxford notation and nomenclature see <http://glycobase.nibrt.ie/glycobase/documents/abbreviations.pdf>

Table 3. Results of N-glycan analysis of hCG and summary of the effects of overexpression of GnTIV, GnTV and knockdown of GalT4 on the glycosylation of the substrate hCG

	hCG	GnT-4GalT-4	GnT-5GalT-4	GnT-4,5,GalT-4
A1	0.8	1.2	0.8	2.7
A2	76.2	53.2	48	60.7
A3	17.3	36.6	49	20.3
A4	0.1	1.4	1.2	1
Hybrids and high mannose	5.5	7.7	1.1	15.3
Galactosylation				
mono	4.2	4.1	2.1	2.2
di	26.3	13.1	14.4	23.2
tri	2.7	4.4	6.5	2.5
tetra	0.1	1.4	1.2	1
Sialylation				
mono	55.2	35.7	54.2	31.3
di	1.7	26.8	14.7	16.6
tri	0	3.2	2.1	1.3
tetra	0	0	0	0
Antennary galactosylation				
A1G1	0.4	0.9	0.1	1.8
A2G1	5.3	4.3	2.0	2.2
A2G2	47.5	38.6	45.8	57.7
A3G1	1.6	9.8	12.7	4.7
A3G2	3.2	9.7	3.4	4.8
A3G3	3.5	16.8	30	10.4

A1, mono-antennary structure; A2, bi-antennary structure; A3, tri-antennary structure; A4, tetra-antennary structure. G1, G2, G3, mono-, bi-, tri-galactosylated structures. Results are presented as the relative percentages of each type of glycan structure.

the hCG cell line were 0.1% of the total glycans; however, they increased ~10–14 fold in the engineered cell lines, displaying similarity to the trend observed in the modelling experiments (Fig. 2G). Tandem mass spectrometry (MS/MS) was used to verify the structures identified by HPLC (see supplementary material Fig. S3). In addition to major changes in antennarity, antennary galactosylation and sialylation were increased substantially in the engineered cell lines. Approximately 70% of structures were found to be core fucosylated. All galactose residues were found to be in β -1,4 linkages to GlcNAc residues, as assessed by digestion with *Streptococcus pneumoniae* β -galactosidase, an enzyme that specifically cleaves galactose linked to *N*-acetylglucosamine in a β (1,4) position.

In vivo hCG activities

The biological activity of hCG was measured using a seminal vesicle weight assay. The basis of this assay is that the hormone increases the weight of the seminal vesicles of immature rats which are then excised and weighed and values are expressed as a measure of hormone activity. Engineered hCG from parent CHO and two stable engineered cell lines was injected subcutaneously into immature male rats. The potency of hCG expressed in parent CHO cells was found to be 16,200 IU/ml. The potency of hCG from GalT4-knockdown cells overexpressing GnT4 was 11,300 IU/ml, whereas hCG from GalT4-knockdown cells overexpressing both GnTIV and GnTV had a potency of 11,400 IU/ml, reflecting a reduction in efficacy of ~30% compared to control.

DISCUSSION

The aims of this study were to comprehend and model the complex N-glycosylation system in mammalian cells, and to identify unique control points governing glycan branching and subsequent biological efficacy. Our study demonstrates a novel means by which the antennarity of N-glycans on protein

expressed in CHO cell lines can be modified and increased. Our approach was guided by computer simulations of N-glycosylation in CHO cells and was based upon a principle of kinetic competition for glycoprotein substrates. The simulations were, in turn, refined by the subsequent experimental results. A major finding of these simulations was that translational suppression of galactosyltransferase activity enhances the effects of the N-glycosyltransferases that are responsible for the formation of high-antennary complex N-glycans. The results obtained from the mathematical model depend on the relative values of the parameters, rather than their absolute values. Initially, the *in silico* knockouts (Fig. 1B–G) showed that tri- and tetra-antennary structures could be formed in the absence of β -1,4-galactosyltransferase activity, despite their lack of galactose or sialic acid residues. This prompted an investigation into the regulation of GalT activity. Subsequently, our *in-vitro* experiments were based on the computer simulation results, which showed a correlation between downregulation of GalT activity and an increase in both tri- and tetra-antennary structures (Fig. 2G). *In silico* variation of the concentrations of GnTE, FucT and SiaT at 10-fold above and below their default values had little or no effect upon antennarity in the steady state.

Our experimental approach was to test the computer-simulated results in a CHO cell line engineered to overexpress the glycoprotein hormone hCG. Sugar chains of naturally occurring hCG, and hCG purified from CHO cells are mainly composed of bi-antennary complex N-glycans (Kessler et al., 1979; Weisshaar et al., 1991). Our experiments demonstrate that a decrease in galactosyltransferase activity is required, possibly through a competitive effect with GnTIV and GnTV for access to the antennary GlcNAc residues. Interestingly, it is primarily the reduction in GalT4 activity that results in substantial 2–3-fold and 10–14-fold increases in the relative levels of tri- and tetra-antennary N-glycan structures, respectively (Table 3). This glycoengineering approach delivers hCG protein in which the

majority of glycans are of the tri-antennary structure. The 10–14-fold increase in tetra-antennary glycans in the engineered cell lines reflects a similar 10-fold increase observed in the mathematical modelling experiments. Large differences in sialylation were also observed, although we found that the higher antennary glycans were not fully sialylated, possibly owing to the insufficient α -2,3-sialyltransferase activity in CHO cells (Baranski et al., 1990; Fukuta et al., 2000). Substantial glycosylation changes also occurred when GnTIV and GnTV were co-overexpressed when GalT4 was knocked down (Fig. 3Aix), with observed increases in tri- and tetra-antennary structures, although the effect was not as marked as when each GnT was individually overexpressed. Although overexpression of GnTV might have been expected to give more substantial increases in tetra-antennary structures, their absence is likely due to the protein backbone structure of hCG, as this is considered to be the determining factor for the addition of highly branched sugar chains to glycoproteins (Lau et al., 2007; Lis and Sharon, 1993; Stanley, 2007; Tulsiani et al., 1982; Ujita et al., 1999). Intriguingly, the dual-enzyme overexpression system gave an increase in oligomannose structures that were almost entirely in the form of Man₅, a result that could be attributed to a reduction in the flux through GnTI, as further simulation results were able to confirm. A reduction in ManII activity would not be expected to give the same result because this enzyme has a marked preference for Man₅GlcNAc (Lee et al., 2001). Whereas a reduction in the level of GalT might also be expected to reduce the extent of sialylation, as is the case with the galactosylation-deficient CHO Lec8 mutant cell line (Stanley, 1989), such a decline was only noticeable in the case of dual-overexpression and might be ascribed to the underprocessing of Man₅. An interesting observation was the effect of GalT4 on the galactosylation of N-glycans, as this isoform is usually associated with the galactosylation of mucin core-2 O-glycans

(Lee et al., 2001). This finding may provide indirect confirmation of an earlier report of full galactosylation of tetra-antennary N-glycans in the absence of any GalTI activity (Hossler et al., 2007). Our data suggest that GalT4 is the dominant transferase in galactosylation of N-glycans in CHO cells.

The enzyme GnTE is responsible, in combinatorial fashion, for much of the potential variety of N-glycan structures, owing to its action in conjunction with the other glycosylation enzymes. Although one recent study has provided evidence of complex extended-chain N-glycan structures in CHO cells (North et al., 2010), our analysis did not discover any antennae containing *N*-acetylglucosamine repeating units. This was a prediction of the model under conditions where the levels of nucleotide sugar-donors are saturating because this enzyme has the lowest k_{cat}/K_m of the enzymes of N-glycosylation. Therefore, much higher expression levels would be required before such structures would be detected in fluorescent-labelling experiments.

Recent evidence has shown that several Golgi enzymes are capable of forming homodimers and heterodimers (Hassinen et al., 2010; Schoberer et al., 2013), supporting the ‘kin recognition’ model of the retention of functionally related Golgi enzymes (Nilsson et al., 1993). In conjunction with these findings, our results might be an indication that GalT4 and GnTs IV and V colocalize in the same region of the Golgi. To our knowledge, the effects of cooperativity on the kinetics of glycosylation have not yet been determined. The effects of enzyme knockdown and overexpression on the localisation of the enzymes, as well as their kinetics, is a subject deserving of future study.

The carbohydrates of glycoproteins are crucial in mediating their biological function, serum retention and interaction with receptors and binding partners. A classic example of this is immunoglobulin G. The N-linked glycans attached to the canonical asparagine 297 residues in both heavy chains of the

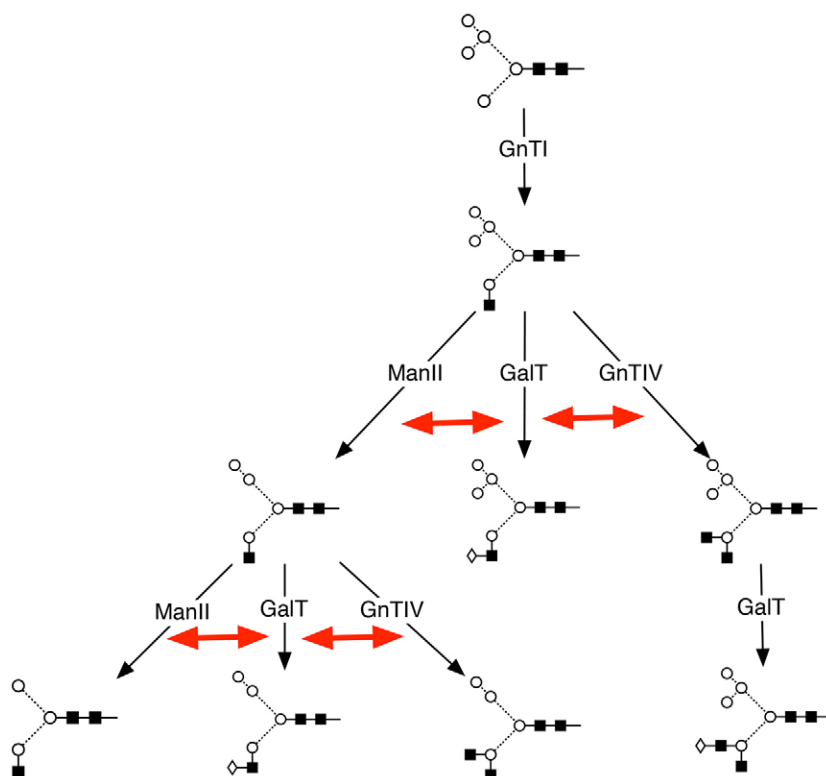


Fig. 5. Galactosyltransferase controls glycan branching by a process of internal kinetic competition among the enzymes of N-glycosylation. Starting with the Man₅, and following the action of the enzyme GnTI, the enzymes GnTIV, ManII and GalT act in parallel. At the steps indicated by the double-headed red arrows, GalT acts in competition with those enzymes leading to oligosaccharide structures of higher antennarity (branching level). Similar competition exists between the subsequent activities of GnTII, GnTV and GalT (not shown).

antibody molecule determine the binding to Fc γ receptors and therefore dictate the immunological response. In addition, individual monosaccharide residues of IgG specify the particular immune response induced by IgG: for example, IgG lacking galactose (IgG G0) residues can interact with mannose-binding lectin (MBL) to activate complement (Malhotra et al., 1995); IgG galactosylation is increased in pregnant women and is involved in placental transport (Kibe et al., 1996); sialylation of IVIG is involved in anti-inflammatory effects (Kaneko et al., 2006); and loss of the $\alpha(1-6)$ core fucose results in increased binding to Fc γ RIIIa and a concomitant increase in ADCC (Okazaki et al., 2004). What we observed, in the case of hCG with increasing the glycan antennarity and sialylation of the hormone, was a consistent decrease in the biological effect *in vivo*, as measured in seminal vesicle weight assays. The molecular basis for this decrease in biological activity might be due to the larger and bulkier glycans found on glycoengineered hCG and decreased binding of the hormone to the luteinising hormone/choriogonadotropin receptor (LHCGR) of the ovary and testis.

In summary, our findings demonstrate the predictive value of the computer model and the identification of galactosyltransferase as a major control point for N-glycan branching (Fig. 5). The results show that cells have in-built mechanisms for glycoengineering proteins that subsequently affect therapeutic efficacy. The combination of computer modelling and bioengineering with recent advances in sequencing of the CHO genome (Xu et al., 2011) will be a powerful toolset for elucidating further aspects of the N-linked glycosylation network and for improving the development of biotherapeutics.

MATERIALS AND METHODS

Chemicals, enzymes and reagents

Restriction enzymes were purchased from Roche Diagnostics Ltd (West Sussex, UK). PNGaseF and exoglycosidase enzymes were from Prozyme (Hayward, Ca. 94545, USA). Nuclease-free water and TRI reagent were from Ambion Inc. (Austin, TX, USA). CD-CHO chemically defined medium, L-glutamine, HT supplement, Dulbecco's modified Eagle's medium (DMEM) containing 4500 mg/l glucose, MEM non-essential amino acids, the SuperScript III First-Strand Synthesis System for RT-PCR, pCR4Blunt-TOPO vector, One Shot chemically competent TOP10 *E. coli*, ViraPower™ Lentiviral Packaging Mix, Lipofectamine™ 2000 Transfection Reagent, heat inactivated fetal bovine serum, trypsin and Opti-MEM I reduced serum medium were obtained from Invitrogen (Dun Laoghaire, Ireland). *Pfu* DNA polymerase, 10 mM dNTPs, GoTaq Flexi polymerase were from Promega Ltd (Southampton, UK). The Qiagen Plasmid Mini Preparation kit, the EndoFree Plasmid Purification kit, the QIAquick PCR purification kit, the QIAquick gel extraction kit, the RNeasy Mini Kit and the DNase set were bought from Qiagen Inc. (Dorking, Surrey, UK). Oligonucleotides and sequencing were provided by Eurofins MWG Operon (Ebersberg, Germany). The pRNAT-U6.2/Lenti constructs were purchased from GenScript USA Inc. (Piscataway, NJ, USA). The AffinityScript QPCR cDNA Synthesis Kit and the Brilliant® II SYBR® Green QPCR Master Mix were from Agilent Technologies (Cork, Ireland).

Mathematical models of glycosylation

Models of the N-glycosylation reactions of CHO cells were developed as systems of ordinary differential equations. The fullest expression of the model is thermodynamically open and includes compartmentation of the Golgi into four regions, cis, medial, endo and trans-Golgi network. As with the earlier model of Umaña and Bailey (Umaña and Bailey, 1997), and its subsequent extension (Krambeck and Betenbaugh, 2005), an

initial glycoform is assumed to enter the system at a constant rate, represented by the zero-order rate constant, q_p . Glycosylated proteins leave the system at a rate proportional to their concentrations, with a first-order rate constant k_G , defined as the reciprocal of the mean residence time in the Golgi.

Letting x_i denote the concentration of the i -th N-glycan, X_i , and f_i the net rate of change of x_i through the action of the enzymes of N-glycosylation (see Table 1), the equations of the model take the general form:

$$\begin{aligned}\frac{dx_{1,1}}{dt} &= q_p + f_1(x_{1,1}) - k_G x_{1,1} \\ \frac{dx_{i,1}}{dt} &= f_i(x_{i,1}, y_{i,1}, \dots, x_{M,1}) - k_G x_{i,1} \quad i=2, \dots, N \\ \frac{dx_{i,j}}{dt} &= k_G x_{i,j-1} + f_i(x_{i,j}, y_{i,j}, \dots, x_{M,j}) - k_G x_{i,j} \quad i=1, \dots, N; j=2, \dots, 4,\end{aligned}\quad (1)$$

where $x_{i,j}$ denotes the concentration of N-glycan X_i in Golgi compartment number j ; f_i is the overall rate law describing the net rate of change of X_i through enzyme action; N is the total number of structures predicted by the model; M is the number of N-glycans Y_i leading directly to the formation of X_i .

The five primary assumptions of model are: (1) rates of influx (q_p) and efflux (k_G) are constant; (2) cell and organelle (Golgi) volumes are constant; (3) concentrations x_i are those of a glycoprotein with a single N-linked glycosylation site; (4) all reactions are irreversible; and (5) compartments are spatially homogeneous.

The functions f_i were sums of enzyme-kinetic rate laws. A single-substrate Michaelis–Menten function was used to describe the kinetics of the mannosidases ManI and ManII:

$$v = \frac{V_{\max}[S]}{K_m + [S]}. \quad (2)$$

For reactions catalysed by the glycosyltransferases and GalT, a random-order mechanism was used, which assumed irreversibility according to assumption number 4. For the reactions of the form $Ax + B \rightarrow A + Bx$, where Ax is the donor molecule and B is the glycoprotein acceptor, the rate of formation of products is given by:

$$v = \frac{V_{\max}[Ax][B]}{[Ax][B] + K_m^{Ax}[B] + K_m^B[Ax] + K_s^{Ax}K_m^B}. \quad (3)$$

K_m values for each donor were obtained from values published in the literature. Given that published data on substrate–enzyme dissociation constants (K_s values) are scarce, it was assumed that $K_s^{Ax} = K_m^{Ax}$ for all Ax . The irreversibility criterion is justified by the presence of nucleoside-diphosphatase activity (EC 3.6.1.6) in Golgi, which would tend to keep the concentrations of UDP and other nucleoside diphosphates at low levels and would additionally minimise the effects of product inhibition. The overall rates of change of sugar-nucleotide donor Ax (UDP-GlcNAc, UDP-Gal, GDP-Fuc and CMP-NeuAc) were incorporated into the model according to:

$$\frac{d[Ax]}{dt} = - \sum_i v_i, \quad (4)$$

for each reaction i in which a molecule of Ax is converted into A , and where each of the v_i is of the form of equation (2) or (3).

As a first approximation, the enzymes of N-glycosylation were distributed according to data obtained from a HeLa cell line (Rabouille et al., 1995). In the absence of specific localisation data, the other enzymes, FucT, GnTs II–V and GnTE, were distributed equally between compartments. In simulations, the percentages were normalised relative to the compartment with the highest value; the normalised values were then applied as factors to the $V_{\max}$ in each rate equation of the model.

The differential equations were integrated in Berkeley Madonna (<http://www.berkeleymadonna.com/>). Steady-state concentrations were computed independently – first in Madonna, by integrating for 100 min of simulation time and recording the final concentrations, and, second, in Mathematica (Wolfram Research, Inc.) by means of the FindRoot[] function.

Cell culture

A CHO cell line (CHO-HCG) overexpressing recombinant hCG was provided by Organon/Merck (Oss, The Netherlands). CHO-HCG cells were cultured between at 0.2×10^6 and 1×10^6 cells/ml in 250 ml sterile Erlenmeyer shake flasks (Corning Life Sciences) on an orbital shaker at 200 rpm in a humidified atmosphere at 37°C with 5% CO₂. Cell culture supernatant from CHO-HCG cells was used as a source of recombinant hCG for glycan analysis and was prepared by harvesting cells at 130 g for 10 min at 4°C.

Immunopurification of hCG

Cell-free supernatants from stable cell lines were prepared by centrifugation at 300 g for 10 min and 4°C and then filtered through a 500 ml Filtropur V 25 filtration unit of pore size 0.2 µm (Sarstedt) and diluted 1:1 with PBS-methionine buffer (4.1 mM Na₂HPO₄·2H₂O, 1.76 mM KH₂PO₄, 0.13 M NaCl, 0.1 mg/ml L-methionine, pH 7.2). Columns were loaded with 1 ml of NHS-Sepharose 4 FF resin conjugated to the monoclonal antibody against β-hCG (number 119A) which was a gift from NV Organon (Oss, The Netherlands). All purification steps were carried out at 4°C under gravity flow. Columns were equilibrated with 50 ml of PBS-methionine buffer. Diluted cell-free supernatants were loaded onto the immune-affinity columns at a flow rate of 1 ml/min. Columns were washed with 50 ml of wash buffer (4.1 mM Na₂HPO₄·2H₂O, 1.76 mM KH₂PO₄, 2 M NaCl, 0.1 mg/ml L-methionine, pH 7.2), followed by 50 ml of PBS-methionine buffer at a flow rate of 2 ml/min. Purified hCG was eluted with 45 ml of elution buffer (10 mM ammonium acetate, 0.1 mg/ml L-methionine, pH 2.7) at a flow rate of 0.5 ml/min. The eluted fractions were immediately neutralised to a pH between 7 and 8 with 25% ammonia and concentrated to ~1 mg/ml using 10-kDa Amicon Ultra-15 Millipore columns and then desalted using PD-10 columns (Amersham Biosciences). The protein concentrations of the hCG protein samples were determined by the method of Bradford (Bradford, 1976) and standardised with BSA. The hCG samples were snap frozen in liquid nitrogen in 60 µg aliquots and freeze dried.

SDS-PAGE analysis of hCG

SDS-PAGE was performed as described by Laemmli (Laemmli, 1970) using a 5% (w/v) polyacrylamide stacking gel and 12% (w/v) polyacrylamide resolving gel. The hCG samples were diluted with 5 µl of 4× sample loading buffer, as previously described (Laemmli, 1970), and denatured at 95°C for 3 min and samples (10 µg) per well were loaded. Electrophoresis was performed at 200 V for ~1 h at room temperature in a Mini-Protein II Cell apparatus (BioRad) filled with running buffer (2.5 g Tris base, 14.4 g glycine and 1 g SDS dissolved in 500 ml of double-distilled H₂O, pH 8.3). Proteins resolved by SDS-PAGE were visualised by staining with Coomassie Blue stain (500 ml methanol, 500 ml H₂O, 100 ml acetic acid, and 2 g Coomassie Brilliant Blue R-250) for at least 30 min on an orbital shaker at room temperature. The gel was destained for ~3 h in 10% methanol 10% acetic acid, 80% MiliQ H₂O). For each sample, the three hCG subunits were excised from the SDS-PAGE gel and stored at –80°C.

Sequence information

The mRNA sequences of the *Cricetulus griseus* (Chinese hamster) GnT and GalT genes were obtained from the National Centre for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov>). The GenBank accession numbers of GnTI, GnTIII, GnTV and GalTs 1, 2, 3, 4, 5 and 6 are U65791, AY598727, U62587, AF318896, AY117536, AY117537, AY117538, AY117539 and AF506839, respectively.

Sequence alignments of Chinese hamster, *Homo sapiens* (human), *Mus musculus* (mouse) and *Rattus norvegicus* (rat) revealed that Chinese hamster GnTs are highly homologous to mouse and rat GnTs (89 to 92% identical). Furthermore, the percentage sequence identities of GnT genes within each species were less than 7%. Thus, specific RT-PCR primers were designed based on mouse and rat GnTII and GnTIV gene sequences, and used to clone GnTII and GnTIV, respectively, from total RNA extracted from CHO cells. It should be noted that these experiments were performed prior to release of the CHO-K1 genomic sequence (Xu et al., 2011).

Generation and sequencing of CHO GnTII, GnTIV and GnTV cDNAs

RNA was isolated from CHO cells using TRI reagent (Ambion) according to the manufacturer's instructions. CHO RNA (4 µg) was reverse transcribed into first strand cDNA using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's protocol. An oligo(dT) primer was used to generate GnTII and GnTV first-strand cDNA. A gene specific GnTIV primer was used to generate the first strand cDNA for GnTIV. RT-PCR primer sequences used for the amplification of each gene were: GnTII forward: 5'-TGGAGACCATGAGGTTCCGCATCTAC-3'; GnTII reverse: 5'-TCACTGCAGTCTTCTATAAC-3'; GnTIV forward: 5'-ATGAGGCTCGCAATGGCACCTTCC-3'; GnTIV reverse: 5'-CTAGTCGGCCTTTTTCAGAAAGATC-3'; GnTV forward: 5'-ATGGCTTCTTTACTCCCTGGAAGTTGTCC-3'; GnTV reverse: 5'-CTATAGGCAGTCTTTGCATAGGGCCACTTG-3'. PCR cycling conditions were as follows: GnTII cDNA was denatured at 95°C for 2 min. This was followed by 30 cycles of denaturation at 95°C for 1 min, annealing at 49°C for 30 s and extension at 72°C for 3 min. A final extension at 72°C for 7 min completed the PCR. GnTIV cDNA was amplified using the same PCR conditions as GnTII, except the annealing temperature was 45°C. GnTV was denatured at 95°C for 2 min, underwent 35 cycles of 95°C for 1 min, 62°C for 30 s, 72°C for 4.5 min and a final cycle of 72°C for 7 min. The proofreading Pfu DNA polymerase (Promega) and 1 µM each of forward and reverse primer were used to amplify 2 µl of first strand cDNA. Following amplification GnT II, IV or V, cDNAs were resolved on a 1% (w/v) agarose gel, as described previously (Sambrook et al., 1989). A QIAquick gel extraction kit (Qiagen) was used to extract and purify the cDNAs from the agarose gel, followed by cloning into the pCR4Blunt-TOPO vector (Invitrogen) according to the manufacturer's instructions. The resulting constructs were transformed into and propagated in One Shot chemically competent *E. coli* TOP10 cells by heat-shock. Direct colony PCR using GoTaq Flexi DNA polymerase (Promega) was used to screen for positive transformants using 2 µl of a 200-µl culture inoculated with single colonies as template for PCR. The same primer sequences (described above) (each 0.4 µM final) were used to screen for positive GnT transformants. PCR products were analysed by agarose gel electrophoresis and positive transformants were selected, cultured and purified using a Qiagen Plasmid Mini Preparation kit (Qiagen). DNA inserts were sequenced by Eurofins MWG Operon using the sequencing primers T3, T7, M13 uni (–21) and M13 rev (–29).

Generation of stable CHO cell lines overexpressing GnTIV and GnTV

CHO GnTIV and GnTV were reverse transcribed from mRNA as described above and cloned into pBluescript SK+ (Stratagene) followed by sub-cloning into the mammalian expression vector pcDNA6/V5-His (Invitrogen) to create the plasmids pcDNA6GnT4 and pcDNA6GnT5 for high level, stable expression in CHO-HCG cells. Both genes were cloned in-frame with a C-terminal V5 epitope. The sequences of the forward primers used were: 5'-CATAAGCTTGCCACCATGAGGCTCCGCAATGGC-3' and 5'-CTAAGCTTGCCACCATGGCTTTCTTTACTC-3' for GnTIV and GnTV respectively. The sequences of the reverse primers used were: 5'-CATGCGGCCGCTGTCCGCCCTTTTTCAGAAAGATC-3' and 5'-GTCCGCGGATATAGGCAGTCTTTCATAG-3' for GnTIV and GnTV respectively. Prior to transfection, the pcDNA6GnT4 and pcDNA6GnT5 plasmids were linearised by digestion with the restriction

enzyme Sca-1 and purified by phenol-chloroform extraction. CHO-HCG cells (1×10^6) were transfected with the linearised plasmids using Amaxa cell line Nucleofector Kit V according to the manufacturer's instructions. Transfected cells were cultured in CD-CHO medium plus supplements for 24 h and then selected for GnT-IV/V overexpression using the antibiotic blasticidin (5 $\mu\text{g}/\text{ml}$) in 96-well plates with ~ 1000 cells seeded per well. Overexpression of GnTIV or GnTV was detected by immunoblotting using mouse monoclonal antibodies directed against the C-terminal V5 epitope. Briefly, following 3–4 weeks of blasticidin selection, cells (1×10^4) were harvested by centrifugation at 0.8 g for 5 min at 4°C, washed with PBS and resuspended in RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% HILIC-40, 1% sodium deoxycholate, 0.1% SDS). Cells were gently agitated at 4°C for 30 min and cell debris and unlysed cells were removed by centrifugation at 120 g for 5 min at 4°C. 10 μg of total protein was subjected to SDS-PAGE as described by Laemmli (Laemmli, 1970) and following separation, proteins were transferred to PVDF membrane (Millipore) and probed with mouse monoclonal anti-V5 antibodies (1:5000 dilution). A secondary goat anti-mouse horseradish peroxidase (HRP)-conjugated antibody (1:20,000) was used to detect the primary anti-V5 antibody and proteins were visualised by enzyme-linked chemiluminescence using Millipore ECL reagents.

Immunofluorescence microscopy

The intracellular location of overexpressed GnTIV and GnTV was determined by immuno-fluorescence microscopy using mouse monoclonal antibodies directed against the C-terminal V5 epitope. CHO cell samples were prepared by centrifugation at 800 g for 5 min at 18°C and washed with PBS. Cells were resuspended and fixed with paraformaldehyde (6%) in PBS (1:1). Following washing in PBS the cell suspension (20 μl) was placed on a poly-L-lysine coated slide and incubated at room temperature for 15 min. Cells were incubated with the primary V5 antibody (1:1000) in PBS containing Tween-20 (0.01%), BSA (5%) and sodium azide (0.1%) for 1 h at room temperature and washed in PBS. The cells were then incubated with 1:1000 dilution of a secondary goat Alexa-Fluor-546-conjugated anti-mouse-IgG antibody (Invitrogen). Slides were mounted with Vectashield containing DAPI stain (Vector Laboratories Ltd) and viewed using a Zeiss Axio fluorescent microscope. The Golgi marker antibody GM130 (Abcam) was used to locate the Golgi membrane within CHO cells.

Generation of stable shRNA knockdown cell lines

Recombinant lentivirus was produced in HEK293FT cells using 9 μg of ViraPower™ Lentiviral Packaging Mix (Invitrogen), 3 μg of endotoxin-free lentiviral construct and 36 μl of Lipofectamine™ 2000 Transfection Reagent (Invitrogen) according to the manufacturer's instructions. Viral supernatants were filtered through a Millex-HV 0.45 μm PVDF filter (Millipore) and stored as 0.5 ml aliquots at -80°C . CHO cells were transduced with a 0.5 ml aliquot of viral supernatant in the presence of 6 $\mu\text{g}/\text{ml}$ polybrene (Millipore) according to the manufacturer's instructions. Cells stably expressing the shRNA knockdown constructs (Genscript, sequences are available upon request) were selected by fluorescent activated cell sorting (FACS) based on the co-expression of coral green fluorescent protein (cGFP), contained within the PRNAT.U6 lentiviral vector to achieve 98–100% GFP positive cells. The stable cell lines were expanded and supernatants were collected and stored at -80°C .

Real-time quantitative PCR

RNA was extracted from the stable cell lines and purified using an RNeasy mini kit (Qiagen) and RNase free DNase Set (Qiagen). Each RNA pellet was resuspended in 40 μl RNase-free $d\text{H}_2\text{O}$ (Ambion) and stored at -80°C . The cDNAs were generated from the RNA templates using an AffinityScript QPCR cDNA Synthesis Kit (Agilent Technologies). Real-time quantitative PCR was carried out using a Brilliant® II SYBR® Green QPCR kit (Agilent Technologies) and a Stratagene MxPro-Mx3000P QPCR machine (Agilent Technologies). The cycling parameters were as follows. Segment 1: 1 cycle of 95°C for 10 min. Segment 2 (association curve): 40 cycles of 95°C for 30 s, 50°C

for 1 min and 72°C for 30 s. Segment 3 (dissociation curve): 1 cycle of 95°C for 1 min, 55°C for 30 s and 95°C for 30 s. Real-time quantitative PCR primer sets were designed using Beacon Designer 7 software (Premier Biosoft International) and synthesised by Eurofins MWG Operon. The optimum forward and reverse primer concentrations were determined using 200 ng of wild type CHO7J RNA as template and all nine possible mixes of 200, 400 and 600 nM forward primer with 200, 400 and 600 nM reverse primer. The optimum RNA template amount and percentage efficiency of each primer set were determined from a standard curve using cDNA amounts equivalent to 999 ng, 333 ng, 111 ng, 37 ng, 12.3 ng, 4.1 ng and 1.37 ng of iHILICut RNA template. The optimum forward and reverse primer concentrations were used in the real-time quantitative PCR reactions and reactions were set up in triplicate. The percentage gene knockdowns were determined using a β -actin primer set as normalizer, wild-type template as calibrator and knockdown samples as unknown. The equivalent of 90 ng of iHILICut RNA was used as template. Reactions were set up in triplicate and the program comparative quantitation (calibrator) was run. When the program was complete, Rel. Quant. to Cal. (dR) was chosen from the results menu to view the percentage of mRNA remaining.

Glycan analysis of hCG

N-glycans were released from hCG and analysed by the method described by Royle et al. (Royle et al., 2006) with a number of modifications. Briefly, gel slices containing hCG were washed with 20 mM NaHCO_3 , pH 7, dried in a vacuum centrifuge and rehydrated in 20 μl NaHCO_3 (30 mM), pH 7 containing 100 mU protein-N-glycosidase F (PNGaseF). The gel slices were incubated at 37°C overnight. The PNGaseF-released glycans were eluted with three alternating washes of miliQ water and acetonitrile and evaporated to dryness in a vacuum centrifuge. Dried glycans were reduced and alkylated by incubation in formic acid (1% v/v) at 37°C for 1 h and dried as above. Glycans were then fluorescently labelled with 2-aminobenzamide (2-AB) using a labelling kit (Ludger Ltd). Excess 2-AB label was removed using Phytip normal-phase resin columns (Phynexus Inc.). The purified glycans were dried and resuspended in MiliQ water (20 μl). Initial glycan profiles were obtained by normal-phase HPLC using a TSK-Gel Amide-80 (4.6 $\times$ 250 mm) column run on a Waters 2695 separations module and detected with a Waters 2475 fluorescent detector. The excitation wavelength was set at 320 nm and emission wavelength at 420 nm. Solvent A was 50 mM ammonium formate, pH 4.4. Solvent B was acetonitrile (100%). For initial profiles on HILIC-HPLC, glycans were injected in 80% acetonitrile and run over a 1-h gradient. The system was calibrated using a 2-AB-labelled dextran ladder. Exoglycosidase digestions and WAX-HPLC-fractionated samples were run over a 3-h gradient. Enzymes used were 1 U/ml *Streptococcus pneumoniae* sialidase (Nan1), 1–2 U/ml *Arthrobacter ureafaciens* sialidase (ABS), 1 U/ml *Streptococcus pneumoniae* β -galactosidase, 1 U/ml bovine kidney fucosidase (BKF), 0.1 U/ml *S. pneumoniae* N-acetylglucosaminidase (GUH), 1 U/ml Jack Bean mannosidase (JBM). WAX-HPLC was performed using a Vydac 301VHP575 (7.5 $\times$ 50 mm) column (Anachem Ltd, Luton, Bedfordshire, UK). Solvent A was 0.5 M formic acid, pH 9 and solvent B was methanol (10% v/v). 2-AB-labelled glycans were injected in MiliQ water and run over a 30-min gradient. Glycans were separated into differently charged fractions by comparison of their elution times with standard N-linked bovine fetuin carbohydrates. Glycan structures were determined by comparison of retention times, expressed in GU values, with those reported in GlycoBase, in combination with WAX-HPLC, WAX-fractionation and exoglycosidase digestions.

Mass spectrometry analysis of N-glycans

N-glycans of hCG were analysed by MALDI mass spectrometry following permethylation according to Ciucanu and Kerek (Ciucanu and Kerek, 1984). Dried glycans were dissolved in 500 μl DMSO (HPLC grade, Sigma-Aldrich). Powdered sodium hydroxide (99.999%, Sigma-Aldrich) and 100 ml iodo-methane (99.9%, Sigma-Aldrich) were added. Samples were vortexed for 1 h. Samples were placed on ice and 1 ml of water was added followed by 1 ml of HPLC-grade dichloromethane.

Sample mixtures were vortexed and the organic phase (containing N-glycans) was removed and saved. The dichloromethane addition was repeated twice and the extracts were pooled. The organic layer was washed with 5 × 1 ml HPLC grade H₂O. Samples were dried prior to MALDI analysis. Permethyated samples were diluted in 75 µl methanol and 25 µl HPLC-grade H₂O. 1 µl sample was mixed with 1 µl of 2,5-dihydroxybenzoic acid (10 mg/ml in H₂O:ACN, 1:1 v/v) on a stainless steel MALDI target plate. Samples were dried and re-crystallised with 0.5 µl ethanol. Permethyated N-glycans were analysed using a Waters MALDI Micro MX mass spectrometer (Manchester, UK) with an ultraviolet 337 nm wavelength nitrogen laser. Instrument settings were as follows: pulse voltage of 2100, suppression voltage of 780 and a laser firing rate of 5 Hz. Spectra processing was performed with Waters MassLynx v4.1 software. Ions were observed as [M+Na]⁺ adducts.

In vivo seminal vesicle weight assay for hCG

The potency of recombinant hCG relative to the International Standard of the hormone or a reference preparation calibrated in international units (IU) was determined under given conditions by measuring the increase in the mass of seminal vesicles (or the prostate gland) of immature rats. The international unit is the activity associated with a stated amount of the International Standard, which consists of a mixture of a freeze-dried extract of hCG from the urine of pregnant women and lactose. The equivalent in international units of the International Standard was established by the World Health Organisation.

Immature male rats of weaning age and of the same strain were used in the study. Animals differed in age by not more than 3 days and had a maximum range in body mass of 10 g. Rats were randomly placed into six groups of at least five animals each, and using one of six littermates per group. Three doses (4 IU, 8 IU and 16 IU) were chosen for the comparison between the reference hCG and recombinant hCG, such that the smallest dose was sufficient to produce a positive response in some of the rats and that the largest dose did not produce a maximal response in any of the rats.

Each rat was injected subcutaneously with a given dose according to its allocated study group on 4 consecutive days at the same time each day. On the fifth day, 24 h after the last injection, rats were killed and the seminal vesicles were harvested. Seminal vesicles were weighed immediately after removing any extraneous fluid and tissue. The precision of the assay was improved by a suitable correction of the organ mass with reference to the total body mass of the animal from which it was taken; an analysis of covariance was used. All animal experiments were performed according to approved guidelines.

Acknowledgements

We thank Martha Motherway Gildea (School of Biochemistry and Immunology, Trinity College, Dublin) for technical assistance.

Competing interests

C.J.M.S., H.K. and T.V.d.L. are employees of Merck, Sharp & Dohme. No other authors have any competing interests.

Author contributions

A.G.M. ran the modelling experiments. J.M.H., T.B. and S.A.G. performed the molecular biology experiments. J.M.H., E.F.J.C. and W.B.S. performed the glycan analysis experiments. C.J.M.S., H.K., T.V.d.L., P.M.R., K.F.T. and G.P.D. conceived, planned and oversaw the experiments. A.G.M., J.M.H., T.B. and G.P.D. wrote the manuscript.

Funding

This work was supported by the National Institute for Bioprocessing Research and Training (NIBRT) in Ireland and the Industry Development Authority of Ireland (IDA). Deposited in PMC for immediate release.

Supplementary material

Supplementary material available online at <http://jcs.biologists.org/lookup/suppl/doi:10.1242/jcs.151878/-DC1>

References

Ashwell, G. and Morell, A. (1974). The dual role of sialic acid in the hepatic recognition and catabolism of serum glycoproteins. *Biochem. Soc. Symp.* **1974**, 117–124.

- Baranski, T. J., Faust, P. L. and Kornfeld, S. (1990). Generation of a lysosomal enzyme targeting signal in the secretory protein pepsinogen. *Cell* **63**, 281–291.
- Bradford, M. M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72**, 248–254.
- Ciucanu, I. and Kerek, F. (1984). A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.* **131**, 209–217.
- Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N. and Dwek, R. A. (2014). Emerging principles for the therapeutic exploitation of glycosylation. *Science* **343**, 1235681.
- Egrie, J. C., Dwyer, E., Browne, J. K., Hitz, A. and Lykos, M. A. (2003). Darbepoetin alfa has a longer circulating half-life and greater in vivo potency than recombinant human erythropoietin. *Exp. Hematol.* **31**, 290–299.
- Ferrara, C., Stuart, F., Sondermann, P., Brünker, P. and Umaña, P. (2006). The carbohydrate at FcγRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. *J. Biol. Chem.* **281**, 5032–5036.
- Ferrara, C., Grau, S., Jäger, C., Sondermann, P., Brünker, P., Waldhauer, I., Hennig, M., Ruf, A., Rufer, A. C., Stihle, M. et al. (2011). Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcγRIII and antibodies lacking core fucose. *Proc. Natl. Acad. Sci. USA* **108**, 12669–12674.
- Fukuta, K., Abe, R., Yokomatsu, T., Kono, N., Asanagi, M., Omae, F., Minowa, M. T., Takeuchi, M. and Makino, T. (2000). Remodeling of sugar chain structures of human interferon-gamma. *Glycobiology* **10**, 421–430.
- Hassinen, A., Rivinoja, A., Kaupila, A. and Kellokumpu, S. (2010). Golgi N-glycosyltransferases form both homo- and heterodimeric enzyme complexes in live cells. *J. Biol. Chem.* **285**, 17771–17777.
- Hirschberg, C. B., Robbins, P. W. and Abeijon, C. (1998). Transporters of nucleotide sugars, ATP, and nucleotide sulfate in the endoplasmic reticulum and Golgi apparatus. *Annu. Rev. Biochem.* **67**, 49–69.
- Hossler, P., Mulukutla, B. C. and Hu, W. S. (2007). Systems analysis of N-glycan processing in mammalian cells. *PLoS ONE* **2**, e713.
- Iida, S., Misaka, H., Inoue, M., Shibata, M., Nakano, R., Yamane-Ohnuki, N., Wakitani, M., Yano, K., Shitara, K. and Satoh, M. (2006). Nonfucosylated therapeutic IgG1 antibody can evade the inhibitory effect of serum immunoglobulin G on antibody-dependent cellular cytotoxicity through its high binding to FcγRIIIa. *Clin. Cancer Res.* **12**, 2879–2887.
- Jefferis, R. (2009). Glycosylation as a strategy to improve antibody-based therapeutics. *Nat. Rev. Drug Discov.* **8**, 226–234.
- Kaneko, Y., Nimmerjahn, F. and Ravetch, J. V. (2006). Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **313**, 670–673.
- Kessler, M. J., Reddy, M. S., Shah, R. H. and Bahl, O. P. (1979). Structures of N-glycosidic carbohydrate units of human chorionic gonadotropin. *J. Biol. Chem.* **254**, 7901–7908.
- Kibe, T., Fujiimoto, S., Ishida, C., Togari, H., Wada, Y., Okada, S., Nakagawa, H., Tsukamoto, Y. and Takahashi, N. (1996). Glycosylation and placental transport of immunoglobulin G. *J. Clin. Biochem. Nutr.* **21**, 57–63.
- Krambeck, F. J. and Betenbaugh, M. J. (2005). A mathematical model of N-linked glycosylation. *Biotechnol. Bioeng.* **92**, 711–728.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Lau, K. S., Partridge, E. A., Grigorian, A., Silvescu, C. I., Reinhold, V. N., Demetriou, M. and Dennis, J. W. (2007). Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* **129**, 123–134.
- Lee, J., Sundaram, S., Shaper, N. L., Raju, T. S. and Stanley, P. (2001). Chinese hamster ovary (CHO) cells may express six beta 4-galactosyltransferases (beta 4GalTs). Consequences of the loss of functional beta 4GalT-1, beta 4GalT-6, or both in CHO glycosylation mutants. *J. Biol. Chem.* **276**, 13924–13934.
- Lis, H. and Sharon, N. (1993). Protein glycosylation. Structural and functional aspects. *Eur. J. Biochem.* **218**, 1–27.
- Lowe, J. B. and Marth, J. D. (2003). A genetic approach to Mammalian glycan function. *Annu. Rev. Biochem.* **72**, 643–691.
- Malhotra, R., Wormald, M. R., Rudd, P. M., Fischer, P. B., Dwek, R. A. and Sim, R. B. (1995). Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. *Nat. Med.* **1**, 237–243.
- McDonald, A. G., Tipton, K. F., Stroop, C. J. and Davey, G. P. (2010). GlycoForm and GlycoLogic: two software applications for the rapid construction and display of N-glycans from mammalian sources. *BMC Res. Notes* **3**, 173.
- Nilsson, T., Slusarewicz, P., Hoe, M. H. and Warren, G. (1993). Kin recognition. A model for the retention of Golgi enzymes. *FEBS Lett.* **330**, 1–4.
- North, S. J., Huang, H. H., Sundaram, S., Jang-Lee, J., Etienne, A. T., Trollope, A., Chalabi, S., Dell, A., Stanley, P. and Haslam, S. M. (2010). Glycomics profiling of Chinese hamster ovary cell glycosylation mutants reveals N-glycans of a novel size and complexity. *J. Biol. Chem.* **285**, 5759–5775.
- Okazaki, A., Shoji-Hosaka, E., Nakamura, K., Wakitani, M., Uchida, K., Kakita, S., Tsumoto, K., Kumagai, I. and Shitara, K. (2004). Fucose depletion from human IgG1 oligosaccharide enhances binding enthalpy and association rate between IgG1 and FcγRIIIa. *J. Mol. Biol.* **336**, 1239–1249.
- Pâquet, M. R., Narasimhan, S., Schachter, H. and Moscarello, M. A. (1984). Branch specificity of purified rat liver Golgi UDP-galactose: N-acetylglucosamine β-1,4-galactosyltransferase. Preferential transfer of galactose on the GlcNAc

- β 1,2-Man α 1,3-branch of a complex biantennary Asn-linked oligosaccharide. *J. Biol. Chem.* **259**, 4716–4721.
- Perlman, S., van den Hazel, B., Christiansen, J., Gram-Nielsen, S., Jeppesen, C. B., Andersen, K. V., Halkier, T., Okkels, S. and Schambye, H. T. (2003). Glycosylation of an N-terminal extension prolongs the half-life and increases the *in vivo* activity of follicle stimulating hormone. *J. Clin. Endocrinol. Metab.* **88**, 3227–3235.
- Rabouille, C., Hui, N., Hunte, F., Kieckbusch, R., Berger, E. G., Warren, G. and Nilsson, T. (1995). Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. *J. Cell Sci.* **108**, 1617–1627.
- Royle, L., Radcliffe, C. M., Dwek, R. A. and Rudd, P. M. (2006). Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions. *Methods Mol. Biol.* **347**, 125–143.
- Royle, L., Campbell, M. P., Radcliffe, C. M., White, D. M., Harvey, D. J., Abrahams, J. L., Kim, Y. G., Henry, G. W., Shadick, N. A., Weinblatt, M. E. et al. (2008). HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. *Anal. Biochem.* **376**, 1–12.
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2 edn. New York, NY: Cold Spring Harbour Laboratory Press.
- Sasai, K., Ikeda, Y., Fujii, T., Tsuda, T. and Taniguchi, N. (2002). UDP-GlcNAc concentration is an important factor in the biosynthesis of beta1,6-branched oligosaccharides: regulation based on the kinetic properties of N-acetylglucosaminyltransferase V. *Glycobiology* **12**, 119–127.
- Schachter, H. (1986). Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides. *Biochem. Cell Biol.* **64**, 163–181.
- Schoberer, J., Liebminger, E., Botchway, S. W., Strasser, R. and Hawes, C. (2013). Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack in planta. *Plant Physiol.* **161**, 1737–1754.
- Schriebl, K., Trummer, E., Lattenmayer, C., Weik, R., Kunert, R., Müller, D., Katinger, H. and Vorauer-Uhl, K. (2006). Biochemical characterization of rhEpo-Fc fusion protein expressed in CHO cells. *Protein Expr. Purif.* **49**, 265–275.
- Shields, R. L., Lai, J., Keck, R., O'Connell, L. Y., Hong, K., Meng, Y. G., Weikert, S. H. and Presta, L. G. (2002). Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc gamma RIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **277**, 26733–26740.
- Shinkawa, T., Nakamura, K., Yamane, N., Shoji-Hosaka, E., Kanda, Y., Sakurada, M., Uchida, K., Anazawa, H., Satoh, M., Yamasaki, M. et al. (2003). The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J. Biol. Chem.* **278**, 3466–3473.
- Sinclair, A. M. and Elliott, S. (2005). Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J. Pharm. Sci.* **94**, 1626–1635.
- Solá, R. J. and Griebenow, K. (2009). Effects of glycosylation on the stability of protein pharmaceuticals. *J. Pharm. Sci.* **98**, 1223–1245.
- Stanley, P. (1989). Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol. Cell. Biol.* **9**, 377–383.
- Stanley, P. (2007). A method to the madness of N-glycan complexity? *Cell* **129**, 27–29.
- Stanley, P., Sundaram, S., Tang, J. and Shi, S. (2005). Molecular analysis of three gain-of-function CHO mutants that add the bisecting GlcNAc to N-glycans. *Glycobiology* **15**, 43–53.
- Tulsiani, D. R., Hubbard, S. C., Robbins, P. W. and Touster, O. (1982). α -D-Mannosidases of rat liver Golgi membranes. Mannosidase II is the GlcNAcMAN5-cleaving enzyme in glycoprotein biosynthesis and mannosidases Ia and IB are the enzymes converting Man9 precursors to Man5 intermediates. *J. Biol. Chem.* **257**, 3660–3668.
- Ujita, M., McAuliffe, J., Suzuki, M., Hindsgaul, O., Clausen, H., Fukuda, M. N. and Fukuda, M. (1999). Regulation of I-branched poly-N-acetylactosamine synthesis. Concerted actions by I-extension enzyme, I-branching enzyme, and beta1,4-galactosyltransferase I. *J. Biol. Chem.* **274**, 9296–9304.
- Umaña, P. and Bailey, J. E. (1997). A mathematical model of N-linked glycoform biosynthesis. *Biotechnol. Bioeng.* **55**, 890–908.
- Umaña, P., Jean-Mairet, J., Moudry, R., Amstutz, H. and Bailey, J. E. (1999). Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat. Biotechnol.* **17**, 176–180.
- Weisshaar, G., Hiyama, J. and Renwick, A. G. C. (1991). Site-specific N-glycosylation of human chorionic gonadotrophin—structural analysis of glycopeptides by one- and two-dimensional ¹H NMR spectroscopy. *Glycobiology* **1**, 393–404.
- Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S. et al. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* **29**, 735–741.

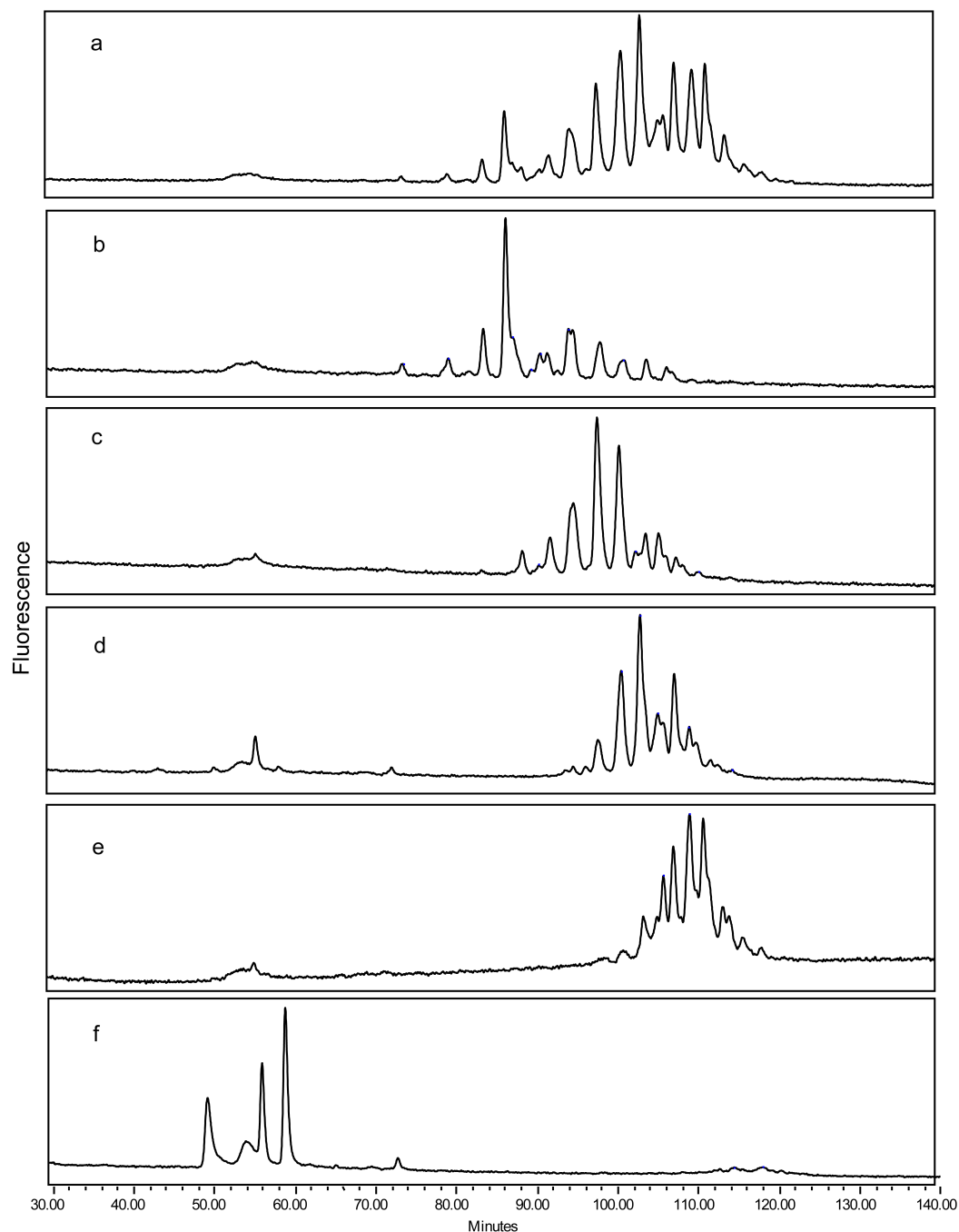


Fig. S1. Tri-sialylated glycan structures are increased in the CHO-GnT4-GalT4 cell line. HILIC analysis of WAX fractionated N-glycans from CHO-GnT4-GalT4 cell line. This is a 2-dimensional glycan analysis using WAX in the first dimension and HILIC in the second dimension as a way to deconvolute complex glycan pools for structural assignments. Major increases in tri-antennary structures were observed in this cell line in comparison to hCG purified from CHO cells **a**. Total N-glycan pool; **b**. Neutral fraction; **c**. Mono-sialylated fraction; **d**. Di-sialylated fraction; **e**. Tri-sialylated fraction; **f**. Tetra-sialylated fraction.

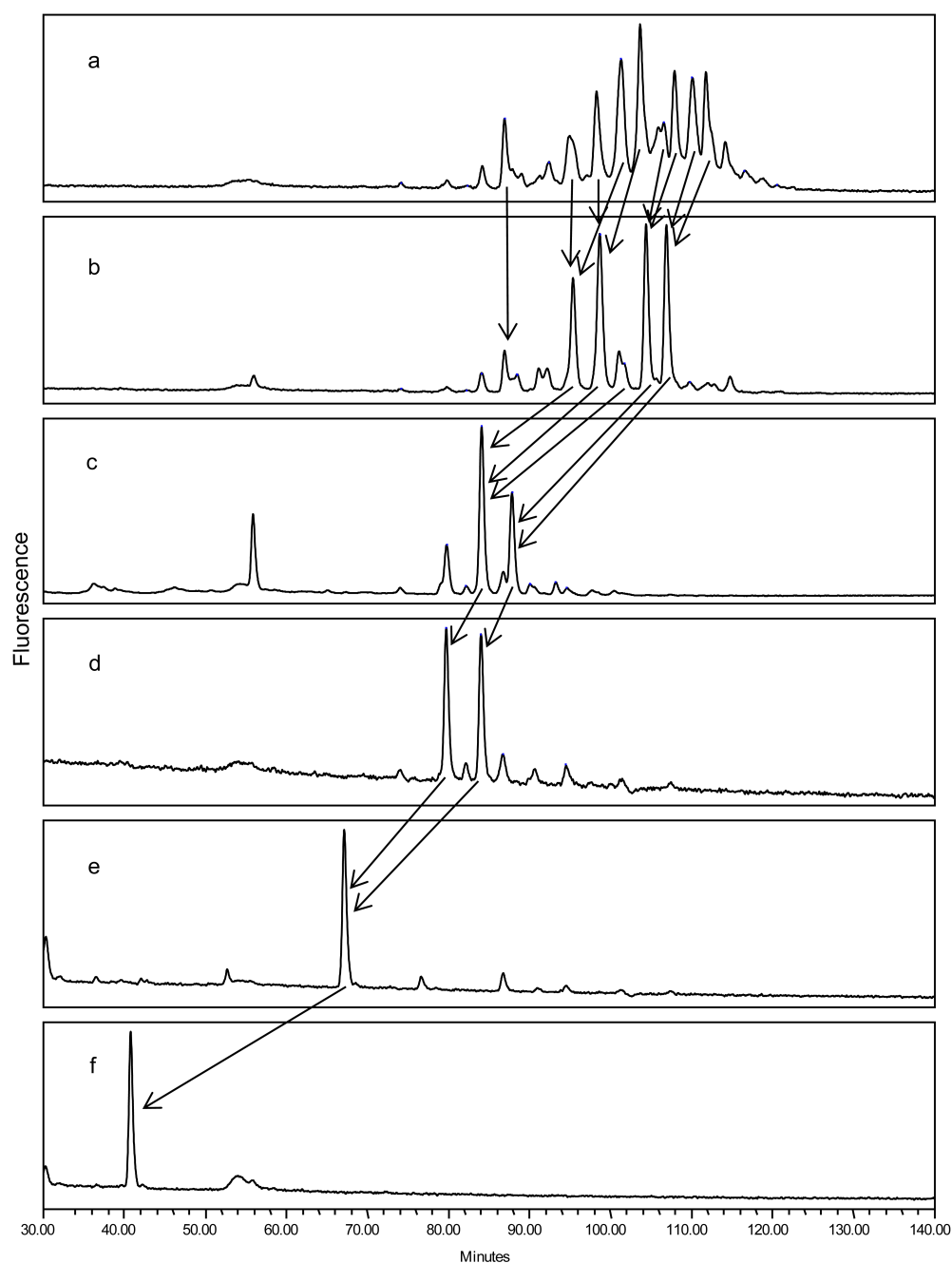


Fig. S2. Exoglycosidase digestions of N-glycans from CHO-GnT4-GalT4. **a.** Undigested N-glycan pool; **b.** *Arthrobacter ureafaciens* sialidase (ABS) digest, ABS is specific for α 2-3/6 linked sialic acids; **c.** ABS + *Streptococcus pneumoniae* (SPG) β -galactosidase digest, SPG is specific for β 1,4-linked galactose residues; **d.** ABS + SPG + bovine kidney fucosidase (BKF) digest, BKF is specific to α 1,6/2 fucose residues with a preference for core fucoses; **e.** ABS+SPG+BKF+ *S. pneumoniae* N-acetylglucosaminidase (GUH), GUH is specific for β 1,2/4/6-linked GlcNac residues; **f.** ABS+SPG+BKF+GUH+Jack Bean Mannosidase (JBM), JBM is specific for α 1-2,6,3 mannose residues. Following digestion HILIC-HPLC was used to analyse the composition of the constituent monosaccharides and to determine individual N-glycan structures.

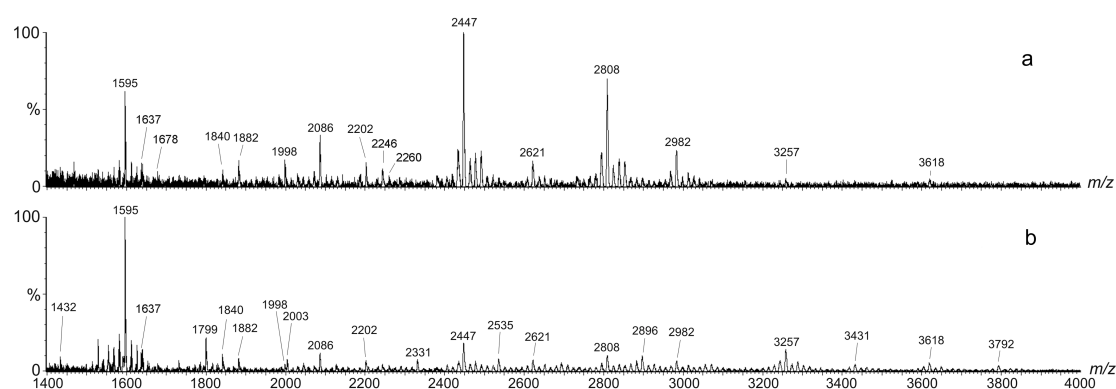


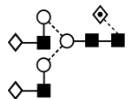
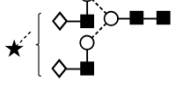
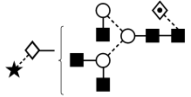
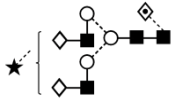
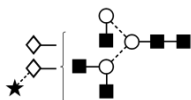
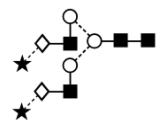
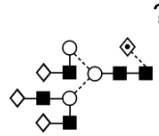
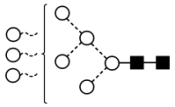
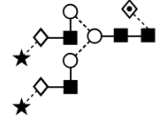
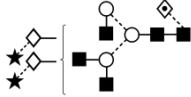
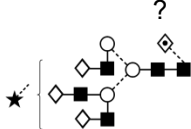
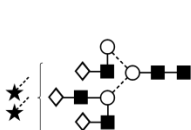
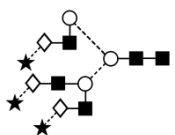
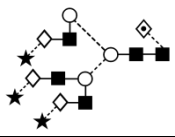
Fig. S3. MALDI-TOF profiles of reduced and permethylated N-glycans from hCG and CHO-GnT4-GalT4

Mass spectrometry was used to verify N-glycan structures identified by HILIC. Permethylated N-glycans were analysed using a Waters MALDI Micro MX mass spectrometer with an ultraviolet 337 nm wavelength nitrogen laser. Spectra processing was performed with Waters MassLynx v4.1 software. Ions were observed as $[M+Na]^+$ adducts. Structures and m/z values are shown in Table SI. **a.** hCG; **b.** CHO-GnT4-GalT4.

Table SI. N-glycan structures identified for hCG purified from CHO-GnT4-GalT4 cell lines.

Structures were determined by a combination of HILIC, WAX-HPLC and exoglycosidase digestions.

Peak	GU Value	Peak Area (%)	Glycan		
			Name	Structure	Observed m/z ($M+Na$) ⁺
1	4.94	0.21	A1		1432.97
2	5.45	0.40	A2		1678.00
		0.15	F(6)A1		ND
3	5.70	0.27	M4A1		1636.85
4	5.89	0.35	F(6)A2		ND
		1.00	A3		ND
5	6.19	3.97	M5		1595.85
6	6.29	0.97	F(6)A3		ND
7	6.42	0.87	A1G(4)1S(3)1		1997.92
8	6.68	1.06	F(6)A2[6]G(4)1		ND
9	6.82	2.49	F(6)A2[3]G(4)1		
10	7.13	0.53	F(6)A3G(4)1		2330.85
		5.50	A2G(4)2		2085.98
11	7.41	0.79	F(6)A2G(4)1S(3)1		ND

12	7.56	7.57	F(6)A2G(4)2		2259.77
13	7.97	9.50	A2G(4)2S(3)1		2246.90
		2.73	F(6)A3G(4)1S(3)1		ND
14	8.30	12.44	F(6)A2G(4)2S(3)1		2620.74
		0.89	A3G(4)2S(3)1		ND
15	8.63	4.00	A2G(4)2S(3)2		2807.48
16	8.73	1.15 4.26	F(6)A3G(4)3		ND
17	8.93		M8		ND
		9.08	F(6)A2G(4)2S(3)2		2981.30
18	9.27	4.40	F(6)A3G(4)2S(3)2		???
19	9.54	4.4 9.30	F(6)A3G(4)3S(3)1		3070.82
20	9.94	1.36 4.89	A3G3S2		3257.77
21	10.35	2.09	A3G3S3		3618.74
22	10.77	1.12	FA3G3S3		3792.26