



# Sequential *in vitro* enzymatic N-glycoprotein modification reveals site-specific rates of glycoenzyme processing

Received for publication, July 25, 2022, and in revised form, September 1, 2022. Published, Papers in Press, September 9, 2022.  
<https://doi.org/10.1016/j.jbc.2022.102474>

Trevor M. Adams<sup>1</sup>, Peng Zhao, Digantkumar Chapla, Kelley W. Moremen<sup>\*</sup>, and Lance Wells<sup>\*</sup>

From the Department of Biochemistry and Molecular Biology, Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia, USA

Edited by Chris Whitfield

N-glycosylation is an essential eukaryotic posttranslational modification that affects various glycoprotein properties, including folding, solubility, protein–protein interactions, and half-life. N-glycans are processed in the secretory pathway to form varied ensembles of structures, and diversity at a single site on a glycoprotein is termed ‘microheterogeneity’. To understand the factors that influence glycan microheterogeneity, we hypothesized that local steric and electrostatic factors surrounding each site influence glycan availability for enzymatic modification. We tested this hypothesis *via* expression of reporter N-linked glycoproteins in N-acetylglucosaminyltransferase *MGAT1*-null HEK293 cells to produce immature  $\text{Man}_5\text{GlcNAc}_2$  glycoforms (38 glycan sites total). These glycoproteins were then sequentially modified *in vitro* from high mannose to hybrid and on to biantennary, core-fucosylated, complex structures by a panel of N-glycosylation enzymes, and each reaction time course was quantified by LC-MS/MS. Substantial differences in rates of *in vitro* enzymatic modification were observed between glycan sites on the same protein, and differences in modification rates varied depending on the glycoenzyme being evaluated. In comparison, proteolytic digestion of the reporters prior to N-glycan processing eliminated differences in *in vitro* enzymatic modification. Furthermore, comparison of *in vitro* rates of enzymatic modification with the glycan structures found on the mature reporters expressed in WT cells correlated well with the enzymatic bottlenecks observed *in vivo*. These data suggest higher order local structures surrounding each glycosylation site contribute to the efficiency of modification both *in vitro* and *in vivo* to establish the spectrum of microheterogeneity in N-linked glycoproteins.

Glycans are important modulators of protein properties and functions across all clades of life (1, 2). N-glycosylation is a conserved and essential cotranslational and posttranslational modification in higher eukaryotes (3) and plays an important role in protein homeostasis (4). N-glycans are cotranslationally and/or posttranslationally attached *en bloc* by oligosaccharyltransferase (5) to the conserved motif N-X-S/T(C), also known

as a “sequon,” where ‘X’ can be any amino acid except proline (6–8). The initial oligosaccharide is mannose rich, and it is trimmed by a series of glycoside hydrolases in the endoplasmic reticulum (ER), eventually exposing the core structure of N-glycans, which is made up of a chitobiose core (GlcNAc-GlcNAc) with branching mannoses,  $\text{Man}_5\text{GlcNAc}_2$ . N-glycans are generally categorized as belonging to one of three classes based on the extent of their processing: high mannose, hybrid, or complex. High mannose glycans are the least processed and most closely resemble the initial oligosaccharide that is transferred onto proteins, while complex glycans are the most processed and can take a variety of forms. This can include branching, extensions, and core fucosylation (9). However, the efficiency of glycan maturation at a given acceptor site on glycoproteins can often be incomplete, most notably (but not solely) because of steric or electrostatic factors that impact enzyme–substrate recognition (9–11). This often results in heterogeneous ensembles of glycan structures on glycoprotein acceptor (12) and even on individual glycosites on the same glycoprotein (13). This phenomenon is termed ‘microheterogeneity’ and is a hallmark of protein glycosylation that has been a focus of biochemical analysis for several decades (14, 15).

An important branching point in N-glycan processing is the addition of a branching  $\beta$ -2-linked GlcNAc to the  $\alpha$ 3 mannose of the  $\text{Man}_5\text{GlcNAc}_2$  structure by the GT-A family glycosyltransferase *MGAT1* (16, 17). This step marks a class switch from high mannose (*e.g.*,  $\text{Man}_5\text{GlcNAc}_2$ ) to hybrid (*e.g.*,  $\text{GlcNAcMan}_5\text{GlcNAc}_2$ ) glycans, as this branching GlcNAc can be further elaborated by other glycosyltransferases into a variety of structures. The activity of *MGAT1* is also necessary for the subsequent trimming of the two terminal mannoses from the  $\alpha$ 6 mannose by the glycoside hydrolase *MAN2A1* (12), which results in a  $\text{GlcNAcMan}_3\text{GlcNAc}_2$  structure. This  $\text{GlcNAcMan}_3\text{GlcNAc}_2$  structure acts as the substrate for *MGAT2* (13), which marks the point at which hybrid N-glycans transition to complex N-glycans, with two branching GlcNAc moieties that serve as a base for highly elaborated biantennary, triantennary, or tetra-antennary structures.

There is much interest in understanding the underlying criteria that define N-glycan microheterogeneity. N-glycans have been shown to be important modulators in antibody–receptor interactions, both with respect to glycosylation of

<sup>\*</sup> For correspondence: Lance Wells, [lwells@ccrc.uga.edu](mailto:lwells@ccrc.uga.edu); Kelley W. Moremen, [moremen@uga.edu](mailto:moremen@uga.edu).

## Monitoring N-glycan microheterogeneity

the antibody (18–20) and their receptors (21, 22). In particular, the contribution of glycosylation to the properties of therapeutics is of particular interest in the development and manufacturing of biologics (4, 23) and biosimilars (24). Glycosylation of these therapeutics is known to impact their stability and pharmacokinetics (4). Additionally, N-glycans are a vital component of viral glycoprotein properties and are known to impact host immune surveillance (25, 26) and host receptor interactions involved in viral entry (27).

There have been several approaches to studying N-glycan microheterogeneity. Early studies demonstrated that N-glycan microheterogeneity is reproducible on a site-by-site basis (28) and that access by glycosyl hydrolases is predictive of N-glycan processing (29). NMR studies of N-glycan structures suggest that glycan interactions with the protein backbone can alter glycan conformations in ways that can impact N-glycan processing (30–32). There is evidence that changing nearby amino acids can alter N-glycan heterogeneity (10, 33). Some recent studies have taken a systems approach, often by monitoring *in vivo* processing of N-glycans in cell culture systems (10, 11). The involvement of the peptide–glycan interactions in affecting glycan conformation and thus potentially N-glycan processing, has also been supported by molecular dynamics (MD) simulations using yeast protein disulfide isomerase (PDI) as a model reporter glycoprotein (34, 35). Additionally, a meta-analysis of site-specific glycoproteomics papers found that solvent accessibility is related to the extent of branching and core fucosylation (36). A recent study by Mathew *et al.* studied early N-glycan processing steps through MD simulations and *in vitro* processing of N-glycans on the yeast protein disulfide isomerase, a similar approach as this study (37). They found that the shape of the surrounding protein environment and subsequent glycan conformation can influence the rate at which individual sites are processed. They studied early mannose trimming and class switching from high mannose to hybrid glycans using three mannosidases and the GlcNAc transferase MGAT1 on the processing of PDI. Our work here expands on this by using enzymes involved in later N-glycan processing using not only the model yeast protein PDI but also multiple N-linked glycoproteins of interest to human health.

In this study, we report extensive site-specific *in vitro* N-glycan processing data for five multiply N-linked glycosylated proteins, with 38 different sites of N-glycosylation in total. By enriching all sites of all glycoproteins with a common Man<sub>5</sub>-GlcNAc<sub>2</sub> substrate and then monitoring N-glycan processing through time-course reactions, we were able to identify key

bottlenecks that prevent specific sites on glycoproteins from being converted from high mannose to complex N-glycans. These bottlenecks appear to persist *in vivo* upon microheterogeneity analysis of each site of the reporter proteins when expressed in WT cells. Additionally, we found that removing the tertiary structure of the protein abolished all site specificity of N-glycan processing, highlighting the importance of protein tertiary structure in defining N-glycan microheterogeneity.

## Results

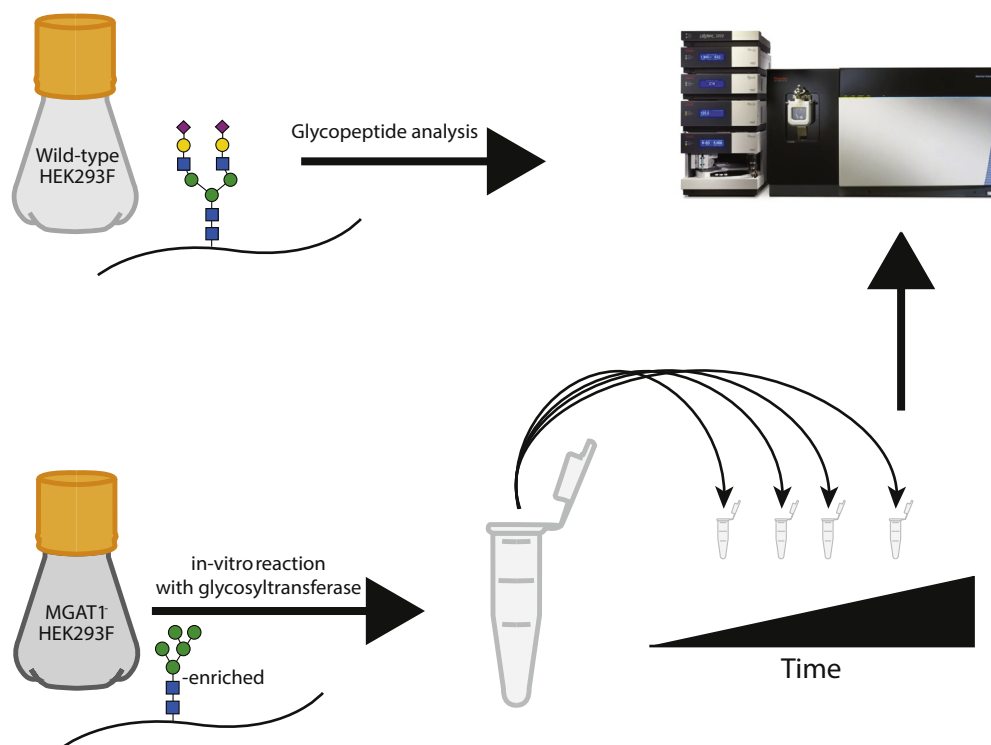
### Expression of reporter proteins in WT-HEK293F cells

In order to probe individual steps of N-glycan processing, we first established a set of reporter proteins to be used as case studies (Table 1, Fig. S1). These proteins were selected based on their various applications in biology, virology, and use as therapeutics as well as their diversity in displayed glycans and the availability of quality crystal structures. CD16a (Fc $\gamma$  receptor IIIa) is an IgG receptor that is known to have differential affinities to antibodies depending on its glycan presentation, which impacts downstream signaling (22, 33, 38). PDI is a resident ER glycoprotein that has been used as a model protein for studying N-glycan processing due to its ease of expression, analysis, and well-defined site-specific glycan heterogeneity (10, 11, 34, 35, 37). Etanercept is a bioengineered therapeutic fusion protein of a TNF $\alpha$  receptor and an IgG1 Fc domain commonly used to help treat autoimmune disorders (39). Erythropoietin is a therapeutic glycoprotein that stimulates red blood cell growth, and its glycosylation is known to impact its pharmacokinetics (40–42). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike glycoprotein is a highly glycosylated trimer that is responsible for the viral entry of the associated coronavirus SARS-CoV-2 *via* binding to the human receptor ACE2 (27, 43).

The reporter proteins were first transiently expressed in high yields in WT HEK293F cells and then harvested from supernatant and purified with Ni-NTA chromatography (Fig. S1, B–F). Glycopeptide analysis using LC-MS/MS was performed on these purified proteins in order to determine their glycan occupancy and diversity when expressed in a “WT” background (Fig. 1). With LC-MS/MS techniques, we were able to obtain a detailed characterization of the N-glycan profile at each site on the reporter proteins, some of which contained dozens of different glycan moieties with a variety of terminal structures including sialylation, as well as core fucosylation (Fig. 2, A–C).

**Table 1**  
Reporter proteins used as models for studying N-glycan processing

| Reporter                        | Glycoprotein reporters |                 |              |   |           |
|---------------------------------|------------------------|-----------------|--------------|---|-----------|
|                                 | Uniprot ID             | Amino acids     | Glycan sites | Glycan structures   | PDB       |
| Etanercept TNFR-IgG Fc fusion   | P20333 + P01857        | 1–235 + 236–467 | 3            | Varied per site   | 3ALQ 3AVE |
| Erythropoietin                  | P01588                 | 28–193          | 3            | Triantennary and tetra-antennary                              | 1EER      |
| SARS-CoV2 Spike                 | P0DTC2                 | 1–1208          | 22           | Varied high man to complex                                    | 6VSB      |
| CD16a Fc $\gamma$ receptor IIIa | P08637                 | 19–192          | 5            | Mostly complex biantennary, triantennary, and tetra-antennary | 5BW7      |
| Pdi1p (yeast)                   | P17967                 | 29–522          | 5            | Complex, varied   | 2B5E      |



**Figure 1. Graphical representation of approach.** Reporter proteins were expressed in HEK293F WT and MGAT1-cells, analyzed via LC-MS/MS, and then processed by purified glycosyltransferases and hydrolases *in vitro*.

All classes of glycans were observed at most sites (Figs. 3, S5–S42), with SARS-CoV-2 spike glycoprotein pictured separately due to its large number of sites (Fig. S2A). Of particular interest are the sites on reporter glycoproteins that greatly differ from other sites on the same protein: sequons 2 and 4 on CD16a (Fig. 3A) and sequon 4 on PDI (Fig. 3B) are predominantly less processed high mannose and hybrid structures, while the other sites of N-glycosylation on the same proteins are mostly highly processed complex structures. This is in contrast to etanercept (Fig. 3C) and erythropoietin (Fig. 3D), which have more homogenous N-glycan presentations. The SARS-CoV-2 spike glycoprotein had a diversity of N-glycan presentations on its 22 sites, with most sites enriched with complex N-glycans and certain sites mostly presenting high mannose N-glycans (Fig. S2). Interestingly, we noted that Man<sub>5</sub>GlcNAc<sub>2</sub> was always the most abundant high mannose structure on 36 of our 38 sites with two exceptions being N0234 and N0717 of SARS-CoV-2 that both contain less than 15% complex structures (Figs. S2, S27, S35).

#### Expression of reporter proteins in Lec1-HEK293F cells

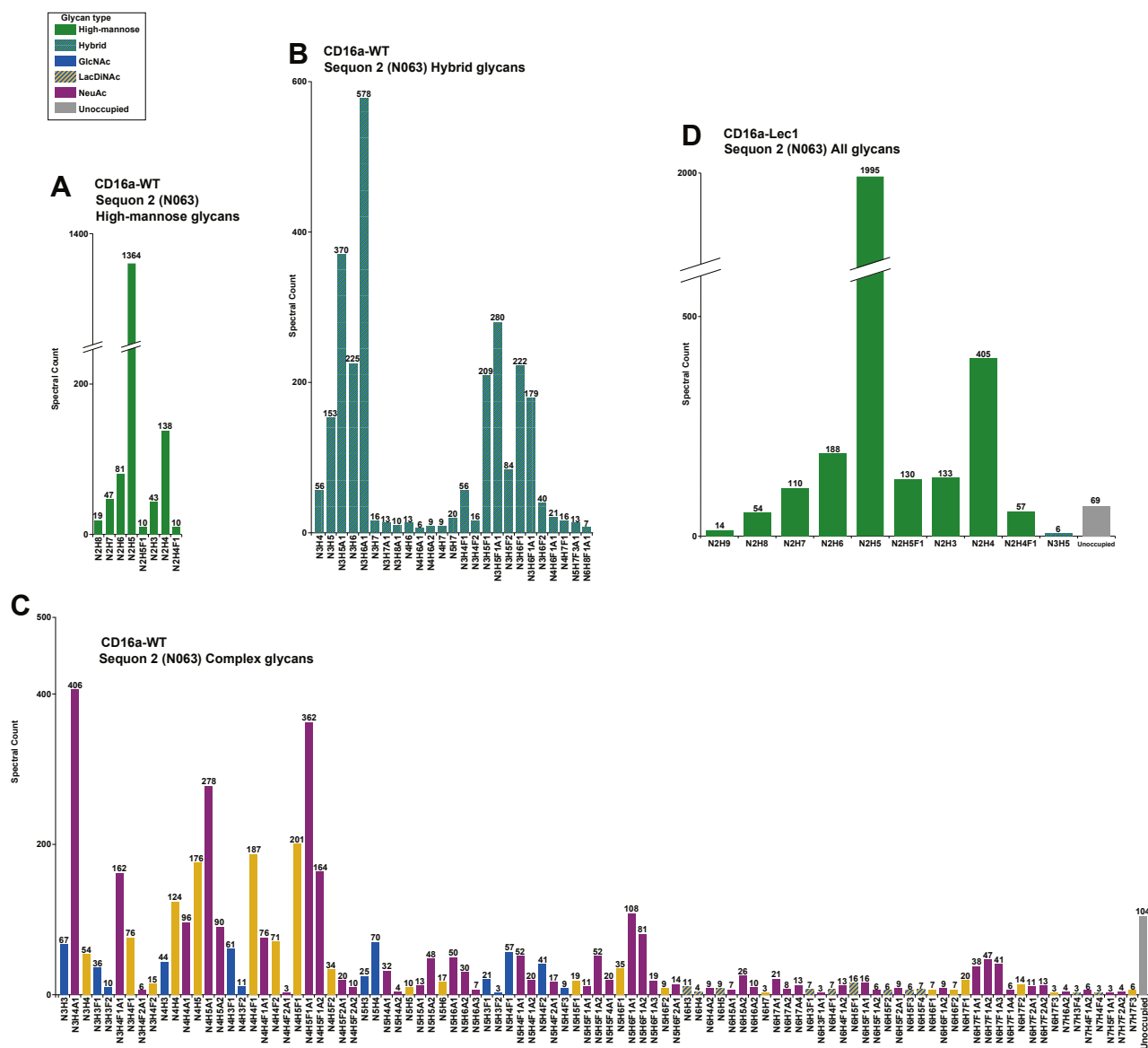
Next, these same proteins were transiently overexpressed in HEK293S GnTI- (MGAT1 null) cells. The activity of MGAT1 is necessary for the formation of both hybrid and complex N-glycans, as the addition of GlcNAc to the nonreducing end  $\alpha$ -3-linked mannose is needed for further elaboration and capping by downstream enzymes. KO of MGAT1 substantially reduces the diversity of glycans at all sites of N-glycosylation and causes a significant enrichment of Man<sub>5</sub>GlcNAc<sub>2</sub> structures N-glycans on expressed glycoproteins, as shown on sequon 2

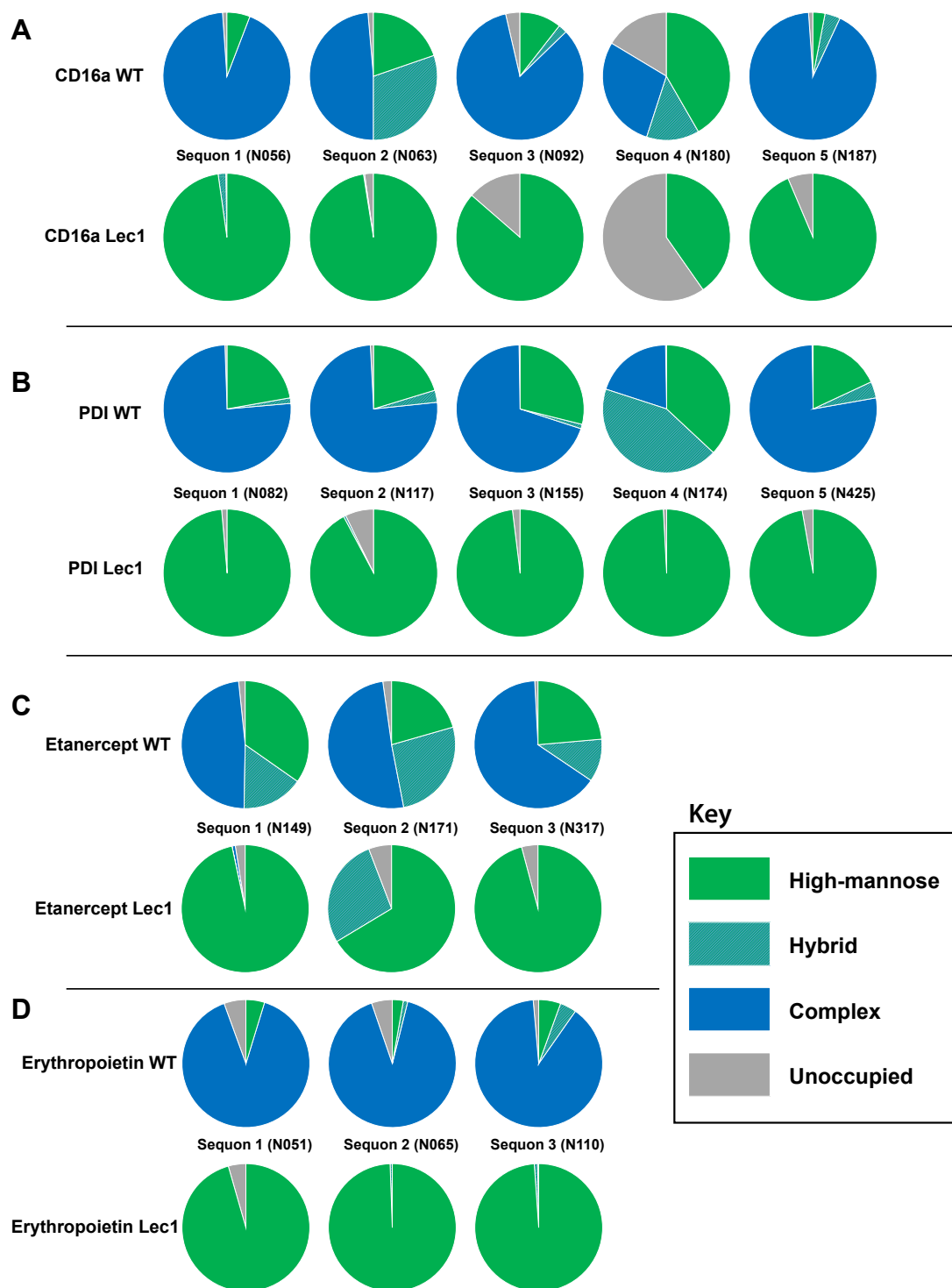
of CD16a (Fig. 2D), as well as the other reporter sites (Figs. S5–S42). This is useful because it allows for the *in vitro* processing of all N-glycans on a glycoprotein to begin from a common substrate. This enrichment was successful for most sites on all reporter proteins (Figs. 3, S2B). However, there is an exception at Sequon 2 of etanercept, which contained a significant population of apparent hybrid N-glycans by an unknown processing event that we are currently exploring. These appear to be true hybrid structures with an attached GlcNAc on the  $\alpha$ -3-linked mannose based on MS2 fragmentation data (Fig. S16B).

#### Conversion of high mannose glycans to hybrid glycans

In order to probe the effects of tertiary structure on N-glycan processing, we first monitored the conversion of Man<sub>5</sub>GlcNAc<sub>2</sub> N-glycans to GlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub> N-glycans via the addition of GlcNAc by the glycosyltransferase MGAT1 on intact reporter proteins expressed and purified from MGAT1-deficient cells. We did this through a series of time-course reactions using purified protein and MGAT1 in the presence of the nucleotide sugar donor UDP-GlcNAc followed by analysis and quantitation via LC-MS/MS (Fig. 4, example TIC can be found in Fig. S43). The ratio of enzymes to molarity of reporter N-glycan sites was kept constant so that we could compare the interprotein as well as intraprotein rates of N-glycan processing. We observed site-specific rates of GlcNAc addition across the range of our 38 sites of N-glycosylation. Generally, these rates corresponded well to the distributions of N-glycans that were found on the respective sites when the reporters were expressed in WT HEK293F cells

## Monitoring N-glycan microheterogeneity





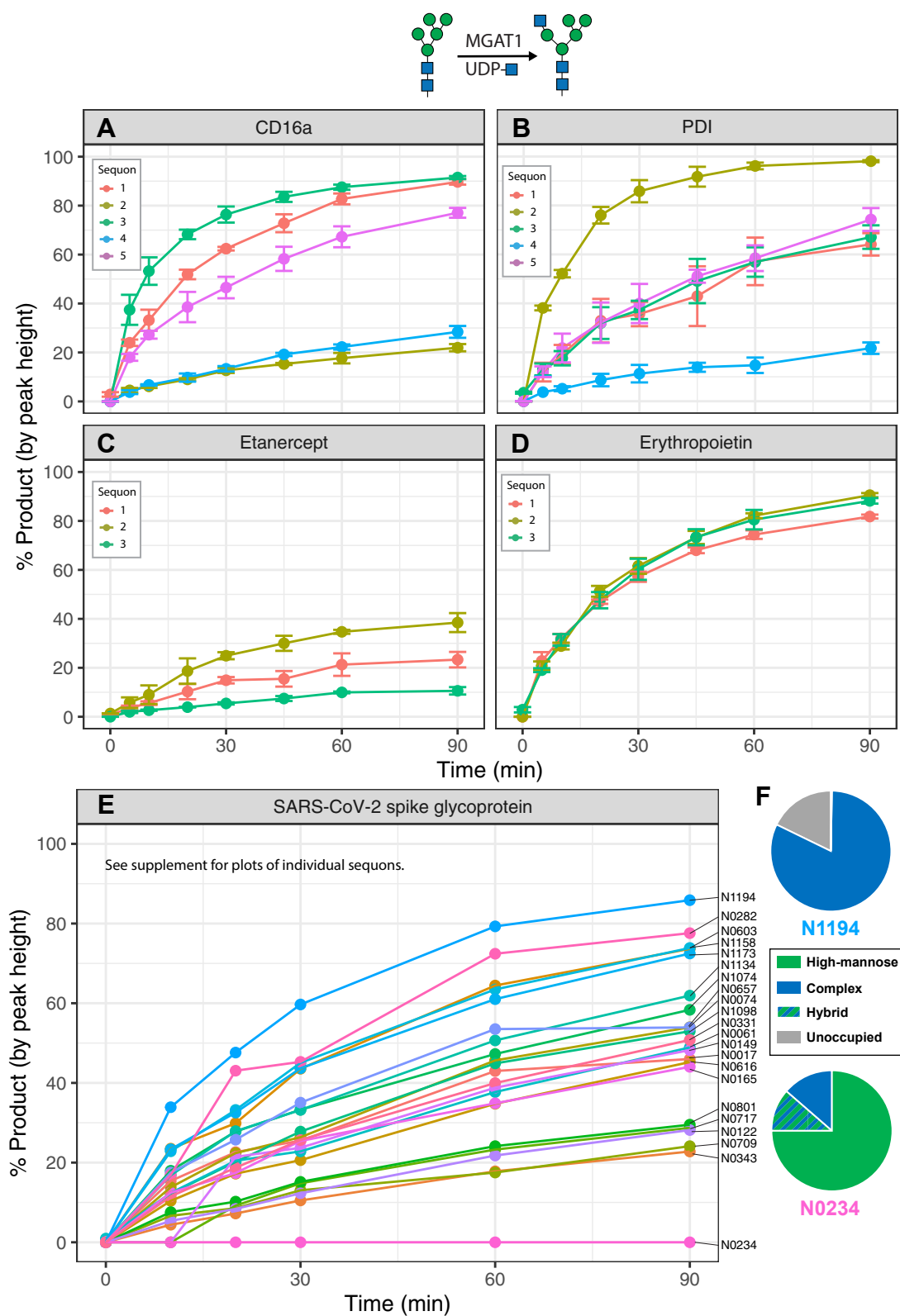
**Figure 3. Site occupancy of reporter proteins expressed in WT-HEK293F and Lec1-HEK293F cells.** Relative proportion of glycan classes at each site on reporter proteins when expressed in a WT or Lec1 (*MGAT1*<sup>-</sup>) background. A, CD16a. B, PDI. C, etanercept. D, erythropoietin. Relative populations were ascertained with glycopeptide analysis and quantified with spectral counts.

examined conversion of the GlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub> product to GlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub> following digestion with MAN2A1 (Fig. 5). Similar patterns of site-specific rates were seen as with MGAT1, with sequons 2 and 4 of CD16a (Fig. 5A) and sequon 4 of PDI (Fig. 5B) all having much lower levels of activity compared to the other sites on the same protein. Again, all sites on etanercept (Fig. 5C) were processed much more slowly

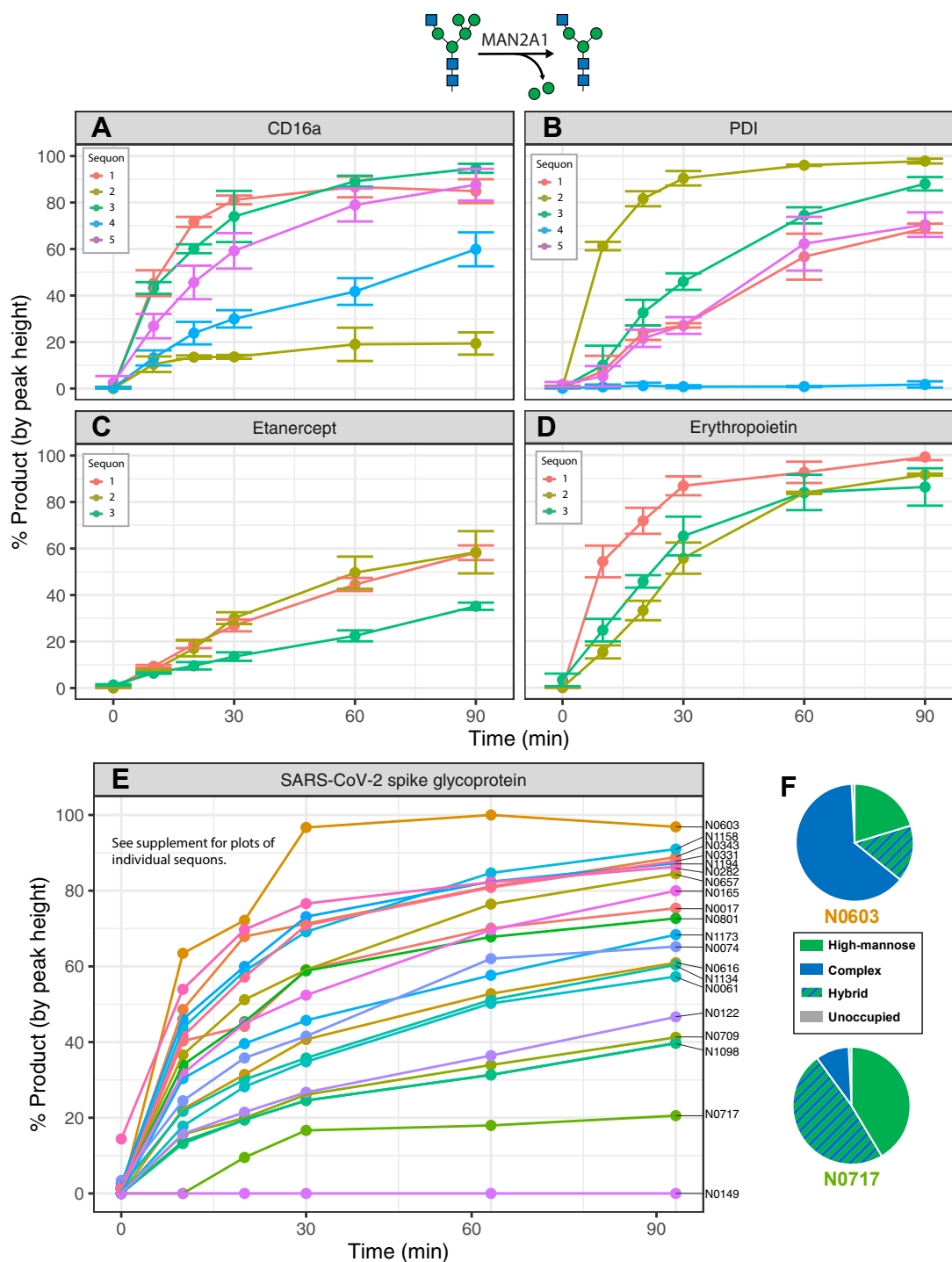
than those on erythropoietin (Fig. 5D). Notably, no cleavage products were observed at sequon 5 (N0149) on SARS-CoV-2 spike glycoprotein (Fig. S3). Inspection of the GlcNAcMan<sub>4</sub>GlcNAc<sub>2</sub> intermediate in MAN2A1 processing revealed that at this site, only one mannose was able to be removed (Fig. S4). Despite this *in vitro* observation, when expressed in a WT background, this site produces an abundance of complex-



## Monitoring N-glycan microheterogeneity



**Figure 4. Site-specific monitoring of MGAT1 activity.** Time-course reaction of GlcNAc addition to reporter proteins with recombinant MGAT1. A, CD16a. B, Protein disulfide isomerase. C, Etanercept. D, Erythropoietin. E, SARS-CoV-2 spike glycoprotein. F, WT glycopeptide profiles for the sites on SARS-CoV-2 spike glycoprotein with the fastest (N1194) and slowest (N0234) rates of MGAT1 activity. Error bars and legend omitted for SARS-CoV-2 spike glycoprotein due to large number of sites. Reaction progress calculated as proportion of the sum of monoisotopic peak heights of product ( $\text{Man}_5\text{GlcNAc}_3$ ) versus the sum of product and reactant ( $\text{Man}_5\text{GlcNAc}_2$ ) peak heights. Experiments performed in triplicate, error bars represent SD. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.



**Figure 5. Site-specific monitoring of MAN2A1 activity.** Time-course reaction of GlcNAc addition to reporter proteins with recombinant MGAT1. A, CD16a. B, Protein disulfide isomerase. C, etanercept. D, erythropoietin. E, SARS-CoV-2 spike glycoprotein. F, WT glycopeptide profiles for the sites on SARS-CoV-2 spike glycoprotein with the fastest (N0603) and slowest (N0717) rates of MAN2A1 activity for a site that was able to form product. Legend omitted for SARS-CoV-2 spike glycoprotein due to large number of sites. Reaction progress calculated as proportion of the sum of monoisotopic peak heights of product (Man<sub>3</sub>GlcNAc<sub>3</sub>) versus the sum of product and reactant (Man<sub>3</sub>GlcNAc<sub>2</sub>) peak heights. Experiments performed in triplicate; error bars represent SD. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

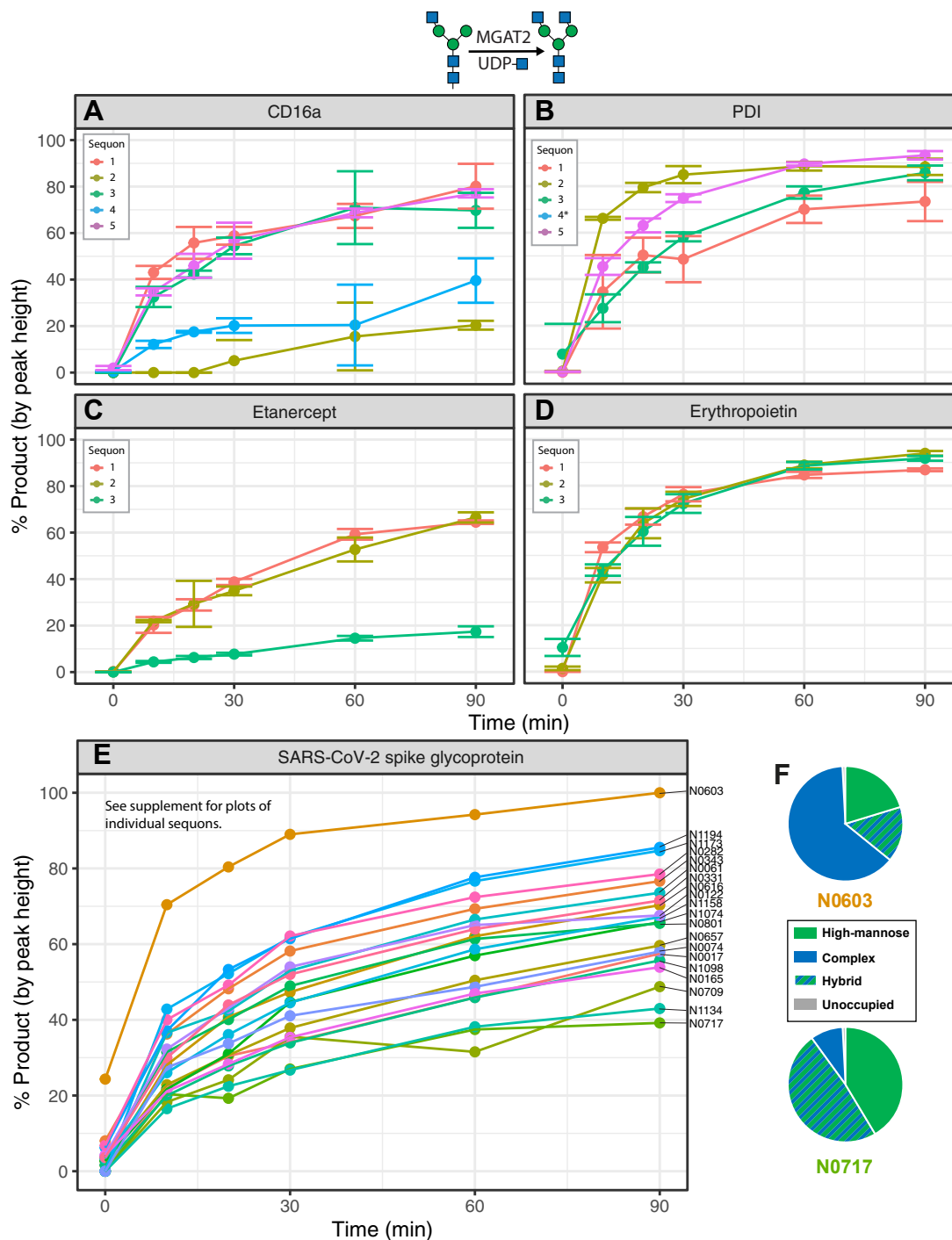
type N-glycans (Figs. S2, S25). This is in contrast to sites N0234 and N0717 on the spike glycoprotein, which exhibit both slow transfer rates and an enrichment of high mannose N-glycans (Figs. 5E, S2 and S3) but still form the GlcNAc-Man<sub>3</sub>GlcNAc<sub>2</sub> product.

The next step in the formation of complex glycans is the addition of a  $\beta$ -2-linked GlcNAc to the  $\alpha$ 6-mannose of the GlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub> moiety. We reacted the reporter proteins expressed in the MGAT1-null cells with an excess of MGAT1, MAN2A1, and the UDP-GlcNAc donor in order to

## Monitoring N-glycan microheterogeneity

enrich the GlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub> substrate, then performed another set of time-course reactions with MGAT2. Sites which could not be efficiently converted to GlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub> structures by MGAT1 and MAN2A1 treatment (e.g., sequon 4 on PDI (Fig. 5B)) were excluded from further analyses. Similar patterns of modification were observed in the MGAT2

reactions as were seen in the MGAT1 and MAN2A1 experiments, with lower levels of activity observed at sequons 2 and 4 on CD16a compared to other sites on the same protein (Fig. 6A). Additionally, all sites on etanercept (Fig. 6C) were processed more slowly than those on erythropoietin (Fig. 6D). A broad range of processing rates was observed on SARS-



**Figure 6. Site-specific monitoring of MGAT2 activity.** Time-course reaction of GlcNAc addition to reporter proteins with recombinant MGAT1. A, CD16a. B, protein disulfide isomerase. C, etanercept. D, erythropoietin. E, SARS-CoV-2 spike glycoprotein. F, WT glycopeptide profiles for the sites on SARS-CoV-2 spike glycoprotein with the fastest (N0603) and slowest (N0717) rates of MGAT2 activity. Error bars and legend omitted for SARS-CoV-2 spike glycoprotein due to large number of sites. Asterisks on site legend indicate that not enough substrate was generated from previous N-glycan processing steps to monitor reaction progress. Reaction progress calculated as proportion of the sum of monoisotopic peak heights of product (Man<sub>3</sub>GlcNAc<sub>4</sub>) versus the sum of product and reactant (Man<sub>3</sub>GlcNAc<sub>3</sub>) peak heights. Experiments performed in triplicate; error bars represent SD. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.



CoV-2 spike glycoprotein (Fig. 6E). Sites that were modified fastest in our *in vitro* modification studies were also enriched in complex N-glycans when expressed in WT HEK293 cells, and the sites that were the slowest for *in vitro* modification corresponded to sites that were relatively enriched with high mannose N-glycans when generated in WT cells (Figs. 6F, S3).

### Core fucosylation of N-glycans by FUT8

Following the aforementioned experiments, we wanted to see if similar patterns of site-specific N-glycan processing rates would apply to core fucosylation. Core fucosylation is the attachment of an  $\alpha$ 1,6-linked fucose to the GlcNAc that is directly attached to the asparagine at the core of N-linked glycans, a reaction which is catalyzed by the fucosyltransferase FUT8. This reaction is generally specific to complex N-glycans (44), and thus, we generated GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> glycans on our collection of reporter proteins by reacting with an excess of MGAT1, MAN2A1, and MGAT2 in the presence of the UDP-GlcNAc sugar donor. We then examined the rates of modification of the respective glycans with FUT8 (Fig. 7). Interestingly, at many sites we found substantial core fucosylation prior to *in vitro* processing despite the reporter proteins being expressed in an MGAT1-null cell line and thus lacking complex N-glycans (Fig. 7, A, B, D and E, S5–S42). Otherwise, we observed a diversity of fucosylation rates among the different sites. Sequon 2 on CD16a (Fig. 7A) and sequon 1 and 5 on PDI were markedly slow (Fig. 7B), as well as sequon 3 on etanercept (Fig. 7C). All sites on erythropoietin were fucosylated rapidly, which possibly reflects their high initial levels of fucosylation even before the FUT8 reaction (Fig. 7D). Most sites on SARS-CoV-2 spike glycoprotein were efficiently fucosylated (Fig. 7E), with a few exceptions (N0122, N0801, N1074, and N1098) (Fig. S3). Glycopeptide analysis data reflects that sites modified more rapidly by FUT8 *in vitro* had higher levels of fucosylation *in vivo* (Fig. 7F).

### Impact of tertiary structure on site specificity

In order to determine whether these site-specific differences in glycosyltransferase rates were due to tertiary structure, we repeated the transfer of GlcNAc onto PDI-Man<sub>5</sub>GlcNAc<sub>2</sub> glycans using MGAT1 but first digested the protein with trypsin to cleave the fully folded protein substrate into glycopeptides. Without the reporter protein tertiary structure, all site specificity of GlcNAc transfer rate was lost and the overall rate of transfer was reduced (Fig. 8A compared to Fig. 3B). Since FUT8 activity requires access to the core GlcNAc linked to the Asn residue of the peptide backbone, we were curious to see if the glycoprotein being cleaved to glycopeptides would eliminate the site specificity observed on intact protein. Similar to MGAT1, all site-specific modification by FUT8 was lost following cleavage of CD16a to glycopeptides, with all sites (including the slow site 2 (Fig. 7A)) exhibiting similar rates of modification (Fig. 8B). CD16a was used instead of PDI for the FUT8 experiment to more fully illustrate that local tertiary structure was important for more than just one of our reporter

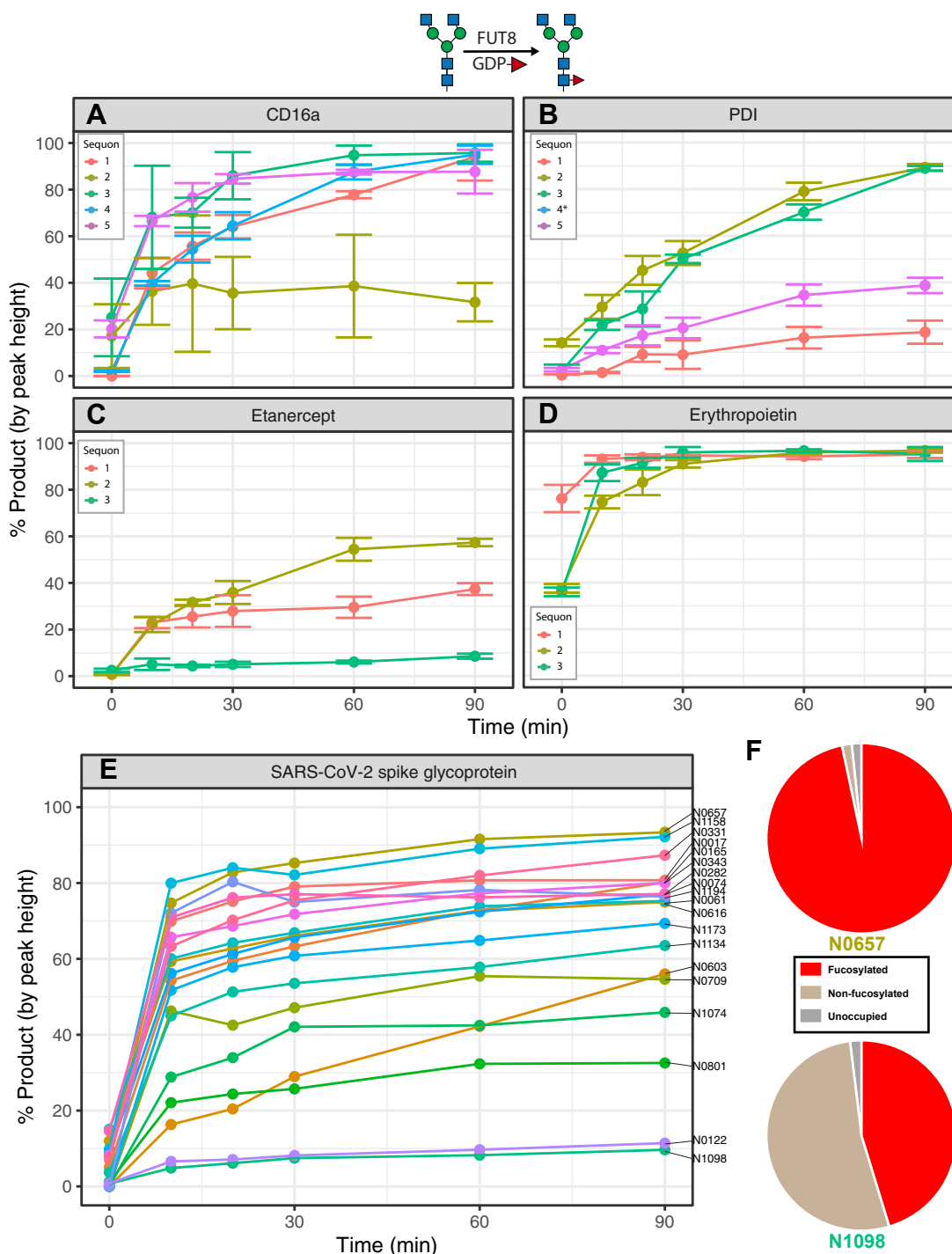
proteins and because site 4 on PDI cannot adequately form complex N-glycans due to low MAN2A1 activity (Fig. 5B).

### Discussion

While N-glycans are a crucial component in the production of membrane bound and secreted glycoproteins, the determinants that define the diversity of N-glycan structures at any given site are not well understood. Factors that have been suggested to influence N-glycan microheterogeneity include the expression of glycosyltransferases and glycoside hydrolases (45, 46), the availability of nucleotide sugar donors (47), secretory pathway trafficking (11, 34, 46), and the accessibility of the acceptor site (10, 37, 48). The impact of enzyme availability in cells has mostly been probed through genetic engineering approaches (45). However, availability of enzymes and sugar nucleotides cannot sufficiently explain site-specific differences on the same polypeptide nor can protein trafficking in the secretory pathway. Thus, we hypothesize that it is the impact of acceptor site accessibility that allows for site-specific differences on the same protein. This could occur by multiple mechanisms including the substrate glycan interacting with the substrate protein backbone at a specific site and thus hampering engagement of the glycoenzyme with the substrate glycan. An example of this is site 4 on PDI that was elegantly demonstrated by Aebi *et al.* (10). Additionally, the glycosylation on the Fc fragment of etanercept (site 3, N317) differs from what is normally seen on IgG as there is an abundance of high mannose glycans accompanying the expected biantennary complex structures (49), which may indicate altered accessibility to the N-glycan at that site in this non-naturally occurring fusion protein. Another possibility is local secondary and tertiary structure of the substrate protein at individual sites of modification and the glycoenzyme active site resulting in steric or electrostatic clashes that prohibit optimal binding for catalysis. Interestingly, our analysis revealed that proteolytically digesting substrate proteins before transfer reactions abolished site-specific rate differences (Fig. 8) and in the case of MGAT1 reduced the overall rate of reaction. This strongly supports higher order structure of the substrate playing an important role in transfer rates and in agreement with proposed mechanisms for microheterogeneity.

MD studies and Markov state modeling by Mathew *et al.* demonstrated that the relative amount of time a glycan spends extended away from the protein and exposed to solvent correlates with site specificity of glycan-processing rates on the yeast model protein disulfide isomerase (37). Additionally, they monitored *in vitro* N-glycan processing rates with ER mannosidases as well as MGAT1 and MAN2A1, and the results with overlapping enzymes in our present study agree well. Site 4 of PDI was identified as a “slow” site (Figs. 4B and 5B), and their Michaelis–Menten analysis of PDI processing kinetics are also consistent with our observations (Figs. 4B and 5B). Additionally, their studies on the kinetics of earlier glycan-processing steps involving ER mannosidase I and Golgi mannosidase IB show that this site specificity is conserved in earlier steps of N-glycan processing. However, their approach to

## Monitoring N-glycan microheterogeneity

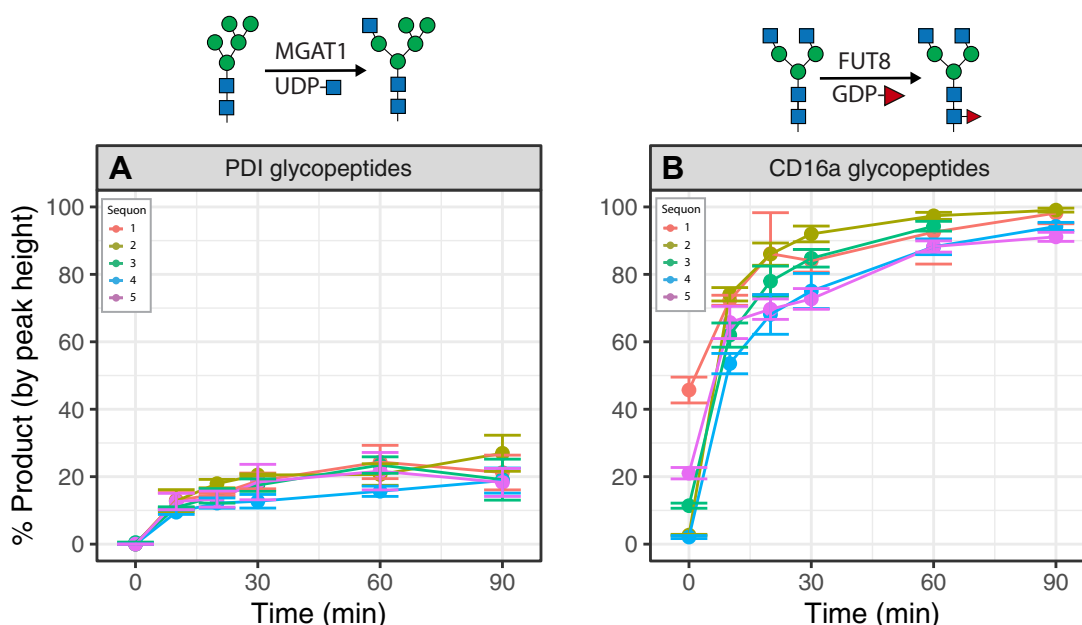


**Figure 7. Site-specific monitoring of FUT8 activity.** Time-course reaction of fucose addition to reporter proteins with recombinant FUT8. *A*, CD16a. *B*, protein disulfide isomerase. *C*, etanercept. *D*, erythropoietin. *E*, SARS-CoV-2 spike glycoprotein. *F*, WT glycopeptide profiles for the sites on SARS-CoV-2 spike glycoprotein with the fastest (N0657) and slowest (N1098) rates of FUT8 activity. Error bars and legend omitted for SARS-CoV-2 spike glycoprotein due to large number of sites. Asterisks on site legend indicate that not enough substrate was generated from previous N-glycan processing steps to monitor reaction progress. Reaction progress calculated as proportion of the sum of monoisotopic peak heights of product ( $\text{Man}_3\text{GlcNAc}_4$ ) versus the sum of product and reactant ( $\text{Man}_3\text{GlcNAc}_4\text{Fuc}_1$ ) peak heights. Experiments performed in triplicate; error bars represent SD. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

reduce tertiary structure through reduction and alkylation prior to *in vitro* modification led to results contrasting with our glycopeptide experiments (Fig. 8), and they observed differences in modification rate at different sites. This may be due to some secondary structures of the protein not being

completely disrupted without the use of protease digestion to cleave the model protein utilized or perhaps the denatured protein can still influence site kinetics.

Generally, these results indicate that the tertiary structure specific to an acceptor site can be an important factor in



**Figure 8. Tertiary structure is responsible for site specificity of N-glycan processing rates.** Time-course reaction of N-glycan processing on proteins digested into glycopeptides with proteases. *A*, reaction of MGAT1 with PDI glycopeptides. *B*, reaction of FUT8 with CD16a glycopeptides. Reaction progress calculated as proportion of the sum of monoisotopic peak heights of product *versus* the sum of product and reactant peak heights. Experiments performed in triplicate; error bars represent SD.

defining the types of N-glycans seen at a given site. In particular, the efficiency (or lack thereof) of MGAT1 and MAN2A1 appears to be highly predictive of high mannose-type glycans at a sequon. *In vivo*, it is likely that MGAT1 is rate limiting as the most common high mannose structure at most sites when expressed in a WT background is the MGAT1 substrate  $\text{Man}_5\text{GlcNAc}_2$  (Figs. S5–S42), and its activity is required for downstream processing by enzymes like MAN2A1. The rate-limiting role of MAN2A1 was also observed in our *in vitro* studies, particularly at sites that were also poorly modified with MGAT1. This may be partly due to lower activity of the recombinant enzyme employed in our *in vitro* studies; four times as much MAN2A1 had to be used in assays compared to the glycosyltransferases (1:250 enzyme:substrate molar ratio for MAN2A1 *versus* 1:1000 for glycosyltransferases). MAN2A1 processing has previously been identified as a potential bottleneck in N-glycan processing (37). Potential steric barriers to MAN2A1 action are suggested by the structure of MAN2A1:substrate complex that demonstrates a significant portion of the total N-linked glycan must fit into the active site of the enzyme for efficient binding and catalysis (50). If a site is not processed totally by MAN2A1, it may form hybrid structures but cannot form complex structures due to the necessity of mannose trimming on the  $\alpha 6$  branch of the trimannosyl N-glycan core. By contrast, the active site structure of MGAT1 involved in acceptor recognition has not yet been determined but likely also presents significant steric barriers for access to some poorly modified sites.

While these studies provide a sound starting point for determining what structural features may be important in determining N-glycan destiny, much work remains. We purposely chose reporter proteins and processing enzymes with experimentally determined structures (Table 1), (43, 51–54).

We are currently utilizing MD simulations of glycosylated reporter proteins and site-specific docking of specific glycan modified reporters with glycoenzymes to determine site-specific glycans interacting with the reporter protein as well as clashes between the reporter sites and the glycoenzymes. This will guide future work involving mutagenesis studies to influence the rate at which glycosyltransferases and glycosyl hydrolases are able to modify acceptor sites. There is evidence that this approach can indeed alter the distribution of N-glycans at a specific site, as evidenced through modification of a tyrosine residue proximal to sequon 4 on protein disulfide isomerase (10). Taking a systematic approach that involves site-specific rate monitoring coupled with modeling and mutagenesis should result in common rules that not only will allow prediction of microheterogeneity but will allow us to tune it.

## Experimental procedures

### Expression and purification of glycoprotein reporters and glycosylation enzymes for *in vitro* modification

Expression constructs encoding the reporter proteins were generated with either  $\text{NH}_2$ -terminal fusion tags (CD16a (low affinity immunoglobulin gamma Fc region receptor III-A, FCGR3A), UniProt P08637, residues 19 to 193; Erythropoietin (EPO), UniProt P01588, residues 28 to 193; Etanercept (TNF receptor-IgG1 fusion), GenBank AKX26891, residues 1–467) or C-terminal fusion tags (yeast PDI1 (protein disulfide-isomerase), UniProt P17967, residues 1–494; SARS-CoV-2 Spike glycoprotein, UniProt P0DTC2, residues 1–1208). The constructs employing N-terminal fusion sequences employed the pGen2 expression vector while the PDI1 construct was generated in the PGEc2 vector as

## Monitoring N-glycan microheterogeneity

previously described (55). For the pGen2 constructs, the fusion protein coding region was comprised of a 25 aa signal sequence, an His<sub>8</sub> tag, AviTag, the “superfolder” GFP coding region, the 7 aa recognition sequence of the tobacco etch virus (TEV) protease followed by the catalytic domain region for reporter proteins (55). Constructs encoding MGAT1, MAN2A1, MGAT2, and FUT8 employed the pGen2 vector and were expressed and purified as previously described (55). For the PDI1 construct, the pGec2 vector was employed and encoded the segment of *Saccharomyces cerevisiae* PDI1 indicated, followed by an SGSG tetrapeptide, the 7 aa TEV recognition sequence, the “superfolder” GFP coding region, and an His<sub>8</sub> tag (55). For SARS-CoV-2 Spike, the construct contained an additional COOH-terminal trimerization sequence and His6 tag as previously described (56). The recombinant reporter proteins were expressed as soluble secreted proteins by transient transfection of suspension culture HEK293F cells (FreeStyle 293-F cells, Thermo Fisher Scientific) for WT glycosylated structures and in HEK293S (GnTI-) cells (ATCC) to generate Man<sub>5</sub>GlcNAc<sub>2</sub>-Asn glycan structures (55, 57). Cultures were maintained at 0.5 to 3.0 × 10<sup>6</sup> cells/ml in a humidified CO<sub>2</sub> platform shaker incubator at 37 °C with 50% humidity. Transient transfection was performed using expression medium comprised of a 9:1 ratio of Freestyle293 expression medium (Thermo Fisher Scientific) and EX-Cell expression medium including Glutmax (Sigma–Aldrich). Transfection was initiated by the addition of plasmid DNA and PEI as transfection reagent (linear 25 kDa PEI, Polysciences, Inc). Twenty-four hours post-transfection, the cell cultures were diluted with an equal volume of fresh media supplemented with valproic acid (2.2 mM final concentration) and protein production was continued for an additional 5 days at 37 °C (3). The cell cultures were harvested, clarified by sequential centrifugation at 1200 rpm for 10 min and 3500 rpm for 15 min at 4 °C, and passed through a 0.8 μM filter (Millipore). The protein preparation was adjusted to contain 20 mM Hepes, 20 mM imidazole, 300 mM NaCl, pH 7.5, and subjected to Ni-NTA Superflow (Qiagen) chromatography using a column preequilibrated with 20 mM Hepes, 300 mM NaCl, 20 mM imidazole, pH 7.5 (Buffer I). Following loading of the sample, the column was washed with 3 column volumes of Buffer I followed by 3 column volumes of Buffer I containing 50 mM imidazole and eluted with Buffer I containing 300 mM imidazole at pH 7.0. The protein was concentrated to approximately 3 mg/ml using an ultrafiltration pressure cell (Millipore) with a 10 kDa molecular mass cutoff membrane and buffer exchanged with 20 mM Hepes, 100 mM NaCl, pH 7.0, 0.05% sodium azide, and 10% glycerol.

### In vitro N-Glycan processing

For the time-course reactions, purified reporter proteins generated in HEK293S (GnTI-) cells were used. Reactions were performed at 37 °C in 1.5 ml Eppendorf tubes in a reaction volume of 150 μl, with 20 mM Hepes (VWR) pH 7.5, and 300 mM NaCl (Fisher). For glycosyltransferases, the corresponding nucleotide sugar, UDP-GlcNAc (Sigma) for MGAT1

and MGAT2 and GDP-Fucose (CarboSynth) for FUT8 was kept in excess at 1 mM. MGAT1 and MGAT2 reactions were supplemented with 1 mM MnCl<sub>2</sub> (Sigma). The concentration of total N-glycans for each reaction was kept at 5 μM; for example, for a reporter protein with 5 sites of N-glycosylation, the concentration of the protein would be 1.25 μM. For MGAT1, MGAT2, and FUT8 reactions, a 1:1000 enzyme-to-glycan ratio was used with the concentration of respective enzyme at 5 nM; for MAN2A1 reactions, a 1:250 enzyme-to-glycan ratio was used with the concentration of MAN2A1 at 20 nM. Prior to adding enzymes, time-course reaction vessels were equilibrated at 37 °C for 15 min. At each time point, 20 μl of samples were taken and reactions were deactivated by heating at 95 °C for 5 min. The samples were then digested by proteases and processed for LC-MS/MS analysis. In order to prepare substrate reporter proteins for N-glycan processing steps downstream of MGAT1 (e.g., MAN2A1, MGAT2, FUT8), reporters were reacted with the appropriate combination of enzymes and sugar nucleotides at a 1:100 enzyme:glycan ratio for 3 h. Reporters were confirmed to have been >80% converted to desired product by LC-MS/MS, detailed later.

### Enzymatic digestion of PDI1, etanercept, EPO, CD16a, and SARS-CoV-2 spike from WT and HEK293S (GnTI-) cells

All proteins were reduced by incubating with 10 mM of DTT (Sigma) at 56 °C and alkylated by 27.5 mM of iodoacetamide (Sigma) at room temperature in dark. For the intact glycopeptide analysis, aliquots of PDI1 proteins were digested respectively using trypsin (Promega), a combination of trypsin and Glu-C (Promega), or a combination of trypsin and AspN (Promega); aliquots of etanercept proteins were digested respectively using trypsin (Promega) or AspN (Promega); aliquots of EPO proteins were digested respectively using a combination of trypsin and Glu-C (Promega) or Glu-C (Promega); aliquots of CD16a proteins were digested respectively using chymotrypsin (Athens Research and Technology), AspN (Promega), or a combination of chymotrypsin (Athens Research and Technology) and Glu-C (Promega); aliquots of S proteins were digested respectively using alpha lytic protease (New England BioLabs), chymotrypsin (Athens Research and Technology), a combination of trypsin and Glu-C (Promega), or a combination of Glu-C and AspN (Promega). For the analysis of deglycosylated glycopeptides, aliquots of PDI1 proteins were digested respectively using trypsin (Promega) or a combination of trypsin and Glu-C (Promega); aliquots of etanercept proteins were digested respectively using trypsin (Promega) or AspN (Promega); aliquots of EPO proteins were digested respectively using a combination of trypsin and Glu-C (Promega) or Glu-C (Promega); aliquots of CD16a proteins were digested respectively using chymotrypsin (Athens Research and Technology) or AspN (Promega); aliquots of S proteins were digested respectively using chymotrypsin (Athens Research and Technology), a combination of trypsin and Glu-C (Promega), or AspN (Promega). Following digestion, the proteins were deglycosylated by Endo-H (Promega)



followed by PNGaseF (Promega) treatment in the presence of  $^{18}\text{O}$  water (Cambridge Isotope Laboratories).

#### LC-MS/MS analysis of glycopeptides of PDI1, etanercept, EPO, CD16a, and SARS-CoV-2 spike from WT and HEK293S (GnTI-) cells

The resulting peptides from respective enzymatic digestion of each protein were separated on an Acclaim PepMap RSLC C18 column (75  $\mu\text{m}$   $\times$  15 cm) and eluted into the nano-electrospray ion source of an Orbitrap Fusion Lumos Tribrid or an Orbitrap Eclipse Tribrid mass spectrometer (Thermo Fisher Scientific) at a flow rate of 200 nl/min. The elution gradient for PDI1, etanercept, EPO, and CD16a proteins consists of 1% to 40% acetonitrile in 0.1% formic acid over 220 min followed by 10 min of 80% acetonitrile in 0.1% formic acid. The elution gradient for S protein consists of 1% to 40% acetonitrile in 0.1% formic acid over 370 min followed by 10 min of 80% acetonitrile in 0.1% formic acid. The spray voltage was set to 2.2 kV and the temperature of the heated capillary was set to 275  $^{\circ}\text{C}$ . For the intact glycopeptide analysis, full MS scans were acquired from  $m/z$  200 to 2000 at 60k resolution, and MS/MS scans following higher energy collisional dissociation with stepped collision energy (15%, 25%, 35%) were collected in the orbitrap at 15k resolution. For the deglycosylated glycopeptide analysis, full MS scans were acquired from  $m/z$  200 to 2000 at 60k resolution, and MS/MS scans following collision-induced dissociation (CID) at 38% collision energy were collected in the ion trap.

For time-course reactions, a shorter LC gradient was used, and digests of the same reporters were combined prior to analysis for higher throughput. The elution gradient used for PDI, etanercept, EPO, and CD16a proteins was 1% to 80% acetonitrile in 0.1% formic acid over 60 min followed by 5 min of 80% acetonitrile in 0.1% formic acid. The peptides were eluted into the source of an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). The spray voltage was set to 2.25 kV and the temperature of the heated capillary was set to 280  $^{\circ}\text{C}$ . Full MS scans were acquired from  $m/z$  300 to 2000 at 60k resolution, and MS/MS scans following CID at 38% collision energy were collected in the ion trap. The elution gradient used for SARS-CoV-2 spike glycoprotein was 1% to 80% acetonitrile in 0.1% formic acid over 300 min followed by 10 min of 80% acetonitrile in 0.1% formic acid. The peptides were eluted into the source of an Orbitrap Eclipse Tribrid mass spectrometer (Thermo Fisher Scientific). The spray voltage was set to 2.25 kV and the temperature of the heated capillary was set to 275  $^{\circ}\text{C}$ . Full MS scans were acquired from  $m/z$  300 to 1900 at 60k resolution, and MS/MS scans following CID at 38% collision energy were collected in the ion trap.

#### MS data analysis

For the intact glycopeptide analysis, the raw spectra were analyzed using pGlyco3 (58) for database searches with mass tolerance set as 20 ppm for both precursors and fragments.

The database search output was filtered to reach a 1% false discovery rate for glycans and 10% for peptides. The filtered result was further validated by manual examination of the raw spectra. For isobaric glycan compositions, fragments in the MS/MS spectra were evaluated to provide the most probable topologies. Quantitation was performed by calculating spectral counts for each glycan composition at each site. Any N-linked glycan compositions identified by only one spectra were removed from the quantitation. For the deglycosylated glycopeptide analysis, the spectra were analyzed using SEQUEST (Proteome Discoverer 1.4 and 2.5, Thermo Fisher Scientific) with mass tolerance set as 20 ppm for precursors and 0.5 Da for fragments. The search output from Proteome Discoverer 1.4 was filtered using ProteoIQ (v2.7, Premier Biosoft) to reach a 1% false discovery rate at protein level and 10% at peptide level. The search output from Proteome Discoverer 2.5 was filtered within the program to reach a 1% false discovery rate at protein level and 10% at peptide level. Occupancy of each N-linked glycosylation site was calculated using spectral counts assigned to the  $^{18}\text{O}$ -Asp-containing (PNGaseF-cleaved) and/or HexNAc-modified (EndoH-cleaved) peptides and their unmodified counterparts.

For time-course reactions, quantitation was performed through manual inspection of MS1 spectra using Thermo Freestyle 1.7 (Thermo Fischer Scientific). The intensities of monoisotopic peak heights for all observable charge states for reactants and products were determined and then summed and averaged in triplicate to determine percent conversion to product over time. Plots generated using RStudio (1.4.1717).

#### Data availability

All data generated or analyzed during this study are included in this article and supporting information files. The glycopeptide analysis MS data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository with the dataset identifier PXD032149. MS data for time-course reactions available upon request.

---

*Supporting information*—This article contains supporting information (Figs. S1–S43) and (Tables S1–S3).

*Acknowledgments*—We thank Dr Henrik Clausen for providing cell lines for this project.

*Author contributions*—T. M. A., K. W. M., and L. W. conceptualization; T. M. A. and P. Z. methodology; T. M. A. and P. Z. validation; T. M. A. and P. Z. formal analysis; T. M. A., D. C., and P. Z. investigation; D. C. resources; P. Z. data curation; T. M. A. writing—original draft; P. Z., D. C., K. W. M., and L. W. writing—review & editing; T. M. A., D. C., and P. Z. visualization; L. W. supervision; L. W. project administration; K. W. M. and L. W. funding acquisition.

*Funding and additional information*—This work was supported in part by National Institutes of Health Grant 5R01GM130915 from NIGMS (to K. W. M. and L. W.) and the Glycoscience Training Program (5T32GM107004) from NIGMS. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.



## Monitoring N-glycan microheterogeneity

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: CID, collision-induced dissociation; ER, endoplasmic reticulum; MD, molecular dynamics; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

### References

- Varki, A. (2017) Biological roles of glycans. *Glycobiology* **27**, 3–49
- Moremen, K. W., Tiemeyer, M., and Nairn, A. V. (2012) Vertebrate protein glycosylation: Diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **13**, 448–462
- Stanley, P., Taniguchi, N., and Aebi, M. (2015) N-Glycans. In *Essentials of Glycobiology*, 3rd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY)
- Solá, R. J., and Griebenow, K. (2009) Effects of glycosylation on the stability of protein pharmaceuticals. *J. Pharm. Sci.* **98**, 1223–1245
- Wild, R., Kowal, J., Eyring, J., Ngwa, E. M., Aebi, M., and Locher, K. P. (2018) Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* **359**, 545–550
- Kasturi, L., Eshleman, J. R., Wunner, W. H., and Shakin-Eshleman, S. H. (1995) The hydroxy amino acid in an Asn-X-Ser/Thr sequon can influence N-linked core glycosylation efficiency and the level of expression of a cell surface glycoprotein. *J. Biol. Chem.* **270**, 14756–14761
- Hart, G. W., Brew, K., Grant, G. A., Bradshaw, R. A., and Lennarz, W. J. (1979) Primary structural requirements for the enzymatic formation of the N-glycosidic bond in glycoproteins. Studies with natural and synthetic peptides. *J. Biol. Chem.* **254**, 9747–9753
- Rudd, P. M., and Dwek, R. A. (1997) Rapid, sensitive sequencing of oligosaccharides from glycoproteins. *Curr. Opin. Biotechnol.* **8**, 488–497
- Stanley, P., Moremen, K. W., Lewis, N. E., Taniguchi, N., and Aebi, M. (2022) N-Glycans. In: Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., et al. eds. *Essentials of Glycobiology*, 4th Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY)
- Losfeld, M.-E., Scibona, E., Lin, C.-W., Villiger, T. K., Gauss, R., Morbidelli, M., et al. (2017) Influence of protein/glycan interaction on site-specific glycan heterogeneity. *FASEB J.* **31**, 4623–4635
- Arigoni-Affolter, I., Scibona, E., Lin, C.-W., Brühlmann, D., Souquet, J., Broly, H., et al. (2019) Mechanistic reconstruction of glycoprotein secretion through monitoring of intracellular N-glycan processing. *Sci. Adv.* **5**, 1–12. eaax8930
- Harpaz, N., and Schachter, H. (1980) Control of glycoprotein synthesis. Processing of asparagine-linked oligosaccharides by one or more rat liver Golgi alpha-D-mannosidases dependent on the prior action of UDP-N-acetylglucosamine: Alpha-D-mannoside beta 2-N-acetylglucosaminyltransferase I. *J. Biol. Chem.* **255**, 4894–4902
- Narasimhan, S., Stanley, P., and Schachter, H. (1977) Control of glycoprotein synthesis. Lectin-resistant mutant containing only one of two distinct N-acetylglucosaminyltransferase activities present in wild type Chinese hamster ovary cells. *J. Biol. Chem.* **252**, 3926–3933
- Arima, T., Spiro, M. J., and Spiro, R. G. (1972) Studies on the carbohydrate units of thyroglobulin. *J. Biol. Chem.* **247**, 1825–1835
- Takasaki, S., Mizuochi, T., and Kobata, A. (1982) Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. *Methods Enzymol.* **83**, 263–268
- Harpaz, N., and Schachter, H. (1980) Control of glycoprotein synthesis. Bovine colostrum UDP-N-acetylglucosamine:alpha-D-mannoside beta 2-N-acetylglucosaminyltransferase I. Separation from UDP-N-acetylglucosamine:alpha-D-mannoside beta 2-N-acetylglucosaminyltransferase II, partial purification, and substrate specificity. *J. Biol. Chem.* **255**, 4885–4893
- Gordon, R. D., Sivarajah, P., Satkunarajah, M., Ma, D., Tarling, C. A., Vizitini, D., et al. (2006) X-Ray crystal structures of rabbit N-acetylglucosaminyltransferase I (GnT I) in complex with donor substrate analogues. *J. Mol. Biol.* **360**, 67–79
- Cymer, F., Beck, H., Rohde, A., and Reusch, D. (2018) Therapeutic monoclonal antibody N-glycosylation – structure, function and therapeutic potential. *Biologicals* **52**, 1–11
- Subedi, G. P., and Barb, A. W. (2015) The structural role of antibody N-glycosylation in receptor interactions. *Structure* **23**, 1573–1583
- Li, T., DiLillo, D. J., Bournazos, S., Giddens, J. P., Ravetch, J. V., and Wang, L.-X. (2017) Modulating IgG effector function by Fc glycan engineering. *Proc. Natl. Acad. Sci.* **114**, 3485–3490
- Hayes, J. M., Frostell, A., Karlsson, R., Müller, S., Martín, S. M., Pauers, M., et al. (2017) Identification of Fc gamma receptor glycoforms that produce differential binding kinetics for rituximab. *Mol. Cell. Proteomics* **16**, 1770–1788
- Patel, K. R., Roberts, J. T., Subedi, G. P., and Barb, A. W. (2018) Restricted processing of CD16a/Fc  $\gamma$  receptor IIIa N-glycans from primary human NK cells impacts structure and function. *J. Biol. Chem.* **293**, 3477–3489
- Higel, F., Seidl, A., Sörgel, F., and Friess, W. (2016) N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. *Eur. J. Pharm. Biopharm.* **100**, 94–100
- Azevedo, V., Hassett, B., Fonseca, J. E., Atsumi, T., Coindreau, J., Jacobs, I., et al. (2016) Differentiating biosimilarity and comparability in biotherapeutics. *Clin. Rheumatol.* **35**, 2877–2886
- Zhou, T., Doria-Rose, N. A., Cheng, C., Stewart-Jones, G. B. E., Chuang, G.-Y., Chambers, M., et al. (2017) Quantification of the impact of the HIV-1-Glycan shield on antibody elicitation. *Cell Rep* **19**, 719–732
- Broszeit, F., van Beek, R. J., Unione, L., Bestebroer, T. M., Chapla, D., Yang, J.-Y., et al. (2021) Glycan remodeled erythrocytes facilitate antigenic characterization of recent A/H3N2 influenza viruses. *Nat. Commun.* **12**, 1–12, 5449
- Zhao, P., Praissman, J. L., Grant, O. C., Cai, Y., Xiao, T., Rosenbalm, K. E., et al. (2020) Virus-receptor interactions of glycosylated SARS-CoV-2 spike and human ACE2 receptor. *Cell Host Microbe* **28**, 586–601.e6
- Swiedler, S. J., Freed, J. H., Tarentino, A. L., Plummer, T. H., and Hart, G. W. (1985) Oligosaccharide microheterogeneity of the murine major histocompatibility antigens. Reproducible site-specific patterns of sialylation and branching in asparagine-linked oligosaccharides. *J. Biol. Chem.* **260**, 4046–4054
- Faye, L., Sturm, A., Bollini, R., Vitale, A., and Chrispeels, M. J. (1986) The position of the oligosaccharide side-chains of phytohemagglutinin and their accessibility to glycosidases determines their subsequent processing in the Golgi. *Eur. J. Biochem.* **158**, 655–661
- Brisson, J. R., and Carver, J. P. (1983) The relation of three-dimensional structure to biosynthesis in the N-linked oligosaccharides. *Can. J. Biochem. Cell Biol.* **61**, 1067–1078
- Savidou, G., Klein, M., Grey, A. A., Dorrington, K. J., and Carver, J. P. (1984) Possible role for peptide-oligosaccharide interactions in differential oligosaccharide processing at asparagine-107 of the light chain and asparagine-297 of the heavy chain in a monoclonal IgG1 kappa. *Biochemistry* **23**, 3736–3740
- Watanabe, Y., Bowden, T. A., Wilson, I. A., and Crispin, M. (2019) Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta Gen. Subj.* **1863**, 1480–1497
- Patel, K. R., Nott, J. D., and Barb, A. W. (2019) Primary human natural killer cells retain proinflammatory IgG1 at the cell surface and express CD16a glycoforms with donor-dependent variability. *Mol. Cell. Proteomics* **18**, 2178–2190
- Hang, L., Lin, C., Grant, O. C., Fleurkens, S., Villiger, T. K., Soos, M., et al. (2015) Analysis of site-specific N-glycan remodeling in the endoplasmic reticulum and the Golgi. *Glycobiology* **25**, 1335–1349
- Weiß, R. G., Losfeld, M.-E., Aebi, M., and Riniker, S. (2021) N-glycosylation enhances conformational flexibility of protein disulfide isomerase revealed by microsecond molecular dynamics and Markov state modeling. *J. Phys. Chem. B.* **125**, 9467–9479
- Thaysen-Andersen, M., and Packer, N. H. (2012) Site-specific glycoproteomics confirms that protein structure dictates formation of N-glycan type, core fucosylation and branching. *Glycobiology* **22**, 1440–1452

37. Mathew, C., Weiß, R. G., Giese, C., Lin, C., Losfeld, M.-E., Glockshuber, R., *et al.* (2021) Glycan–protein interactions determine kinetics of N-glycan remodeling. *RSC Chem. Biol.* **2**, 917–931
38. Patel, K. R., Rodriguez Benavente, M. C., Lorenz, W. W., Mace, E. M., and Barb, A. W. (2020) Fc  $\gamma$  receptor IIIa/CD16a processing correlates with the expression of glycan-related genes in human natural killer cells. *J. Biol. Chem.* **296**, 1–11
39. Liu, L., Gomathinayagam, S., Hamuro, L., Prueksaritanont, T., Wang, W., Stadheim, T. A., *et al.* (2013) The impact of glycosylation on the pharmacokinetics of a TNFR2:fc fusion protein expressed in glycoengineered *pichia pastoris*. *Pharm. Res.* **30**, 803–812
40. Čaval, T., Tian, W., Yang, Z., Clausen, H., and Heck, A. J. R. (2018) Direct quality control of glycoengineered erythropoietin variants. *Nat. Commun.* **9**, 1–8, 3342
41. Sinclair, A. M., and Elliott, S. (2005) Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J. Pharm. Sci.* **94**, 1626–1635
42. Banks, D. D. (2011) The effect of glycosylation on the folding kinetics of erythropoietin. *J. Mol. Biol.* **412**, 536–550
43. Sztain, T., Ahn, S.-H., Bogetti, A. T., Casalino, L., Goldsmith, J. A., Seitz, E., *et al.* (2021) A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat. Chem.* **13**, 963–968
44. García-García, A., Serna, S., Yang, Z., Delso, I., Taleb, V., Hicks, T., *et al.* (2021) FUT8-Directed core fucosylation of N-glycans is regulated by the glycan structure and protein environment. *ACS Catal.* **11**, 9052–9065
45. Narimatsu, Y., Büll, C., Chen, Y.-H., Wandall, H. H., Yang, Z., and Clausen, H. (2021) Genetic glycoengineering in mammalian cells. *J. Biol. Chem.* **296**, 1–25, 100448
46. Hirschberg, K., and Lippincott-Schwartz, J. (1999) Secretory pathway kinetics and *in vivo* analysis of protein traffic from the Golgi complex to the cell surface. *FASEB J.* **13**(Suppl 2), S251–S256
47. Burleigh, S. C., van de Laar, T., Stroop, C. J. M., van Grunsven, W. M. J., O'Donoghue, N., Rudd, P. M., *et al.* (2011) Synergizing metabolic flux analysis and nucleotide sugar metabolism to understand the control of glycosylation of recombinant protein in CHO cells. *BMC Biotechnol.* **11**, 1–17, 95
48. Yu, X., Baruah, K., Harvey, D. J., Vasiljevic, S., Alonzi, D. S., Song, B.-D., *et al.* (2013) Engineering hydrophobic protein-carbohydrate interactions to fine-tune monoclonal antibodies. *J. Am. Chem. Soc.* **135**, 9723–9732
49. Cobb, B. A. (2020) The history of IgG glycosylation and where we are now. *Glycobiology* **30**, 202–213
50. Rose, D. R. (2012) Structure, mechanism and inhibition of Golgi  $\alpha$ -mannosidase II. *Curr. Opin. Struct. Biol.* **22**, 558–562
51. Mukai, Y., Nakamura, T., Yoshikawa, M., Yoshioka, Y., Tsunoda, S., Nakagawa, S., *et al.* (2010) Solution of the structure of the TNF-TNFR2 complex. *Sci. Signal.* **3**, 1–10
52. Syed, R. S., Reid, S. W., Li, C., Cheetham, J. C., Aoki, K. H., Liu, B., *et al.* (1998) Efficiency of signalling through cytokine receptors depends critically on receptor orientation. *Nature* **395**, 511–516
53. Isoda, Y., Yagi, H., Satoh, T., Shibata-Koyama, M., Masuda, K., Satoh, M., *et al.* (2015) Importance of the side chain at position 296 of antibody Fc in interactions with Fc $\gamma$ RIIIa and other fc $\gamma$  receptors. *PLoS One* **10**, 1–17, e0140120
54. Tian, G., Xiang, S., Noiva, R., Lennarz, W. J., and Schindelin, H. (2006) The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell* **124**, 61–73
55. Moremen, K. W., Ramiah, A., Stuart, M., Steel, J., Meng, L., Forouhar, F., *et al.* (2018) Expression system for structural and functional studies of human glycosylation enzymes. *Nat. Chem. Biol.* **14**, 156–162
56. Xiao, T., Lu, J., Zhang, J., Johnson, R. I., McKay, L. G. A., Storm, N., *et al.* (2021) A trimeric human angiotensin-converting enzyme 2 as an anti-SARS-CoV-2 agent. *Nat. Struct. Mol. Biol.* **28**, 202–209
57. Meng, L., Forouhar, F., Thieker, D., Gao, Z., Ramiah, A., Moniz, H., *et al.* (2013) Enzymatic basis for N-glycan sialylation: Structure of rat  $\alpha$ 2,6-sialyltransferase (ST6GAL1) reveals conserved and unique features for glycan sialylation. *J. Biol. Chem.* **288**, 34680–34698
58. Zeng, W.-F., Cao, W.-Q., Liu, M.-Q., He, S.-M., and Yang, P.-Y. (2021) Precise, fast and comprehensive analysis of intact glycopeptides and modified glycans with pGlyco3. *Nat. Methods* **18**, 1515–1523