Electronic Supplementary Material (ESI) for RSC Chemical Biology.

Supplementary to "Glycan-Protein Interactions Determine Kinetics of *N*-Glycan Remodeling"

Corina Mathew[1*], R. Gregor Weiß[2*], Christoph Giese[3], Chia-wei Lin[1, 4], Marie-Estelle Losfeld[1], Rudi Glockshuber[3], Sereina Riniker[2] and Markus Aebi[1]

[*] shared first author

Table of Contents

1. Supplementary Methods

Initial Seeding

To generate a diverse set of starting configurations, an initial enhanced sampling protocol was performed with GROMACS 5.1.4 [1] using PLUMED 2.3 [2]. The replica exchange with solute tempering (REST2) [3] scheme was employed for the PDI protein with sites 1 – 4 glycosylated. A cubic box with an edge length of 16 nm with 132'000 solvent molecules including 42 sodium ions was used. To avoid spurious artifacts that have been identified in charged simulation boxes with PME long-range electrostatics [4], the protein and ion electrostatic interactions were scaled together, such that the total charge of the setup was zero in all replicas. The effective replica temperatures ranged from $T_{\min} = 300\ \mathrm{K}$ to $T_{\max} = 400\ \mathrm{K}$ in $n = 30$ steps with intermediate temperatures determined by the replica number i using $T_i = T_{\min}\exp\left(i \cdot \ln\left(T_{\max}/T_{\min}\right)/(n-1)\right)$. First, the separate replicas were equilibrated for 1 ns at the respective replica's temperature. Afterwards, the replica exchange protocol was applied with exchange attempts every 20 ps over a total run time of 100 ns.

Box Size Optimization

After the enhanced conformational sampling with REST2, the optimized box size was determined using the near-densest-lattice-packing (NDLP) algorithm [5]. The algorithm utilizes the REST2 sampled structures to render a triclinic box such that the minimum distance between periodic images (of each input structures) is at least 3.5 nm. The conservative choice of 3.5 nm minimum distance is far beyond twice the cutoff range and presents a compromise between the number of solvent molecules and the hydrodynamic interactions across periodic boundaries. The resulting triclinic box consisted of 63'000 water molecules, 42 ions, and the glycoprotein.

For the second setup with only the a'-domain, the box lengths were manually set to $6.35\ nm \times 6.35\ nm \times 9.5\ nm$. The glycosylation site 5 was aligned along the longer z box edge. In the case of this rectangular, elongated box, the pressure coupling was applied semi-anisotropically, i.e. the x and y directions were coupled to a different barostat than the z direction.

For both setups, the translational and rotational velocity of the center of mass was removed every 100 steps. The subsequent production runs were performed with the GROMACS-2018-rtc [1] branch available on http://github.com/Tsjerk/gromacs/tree/rtc2018.

Branched GM I Reaction Mechanism

To account for the apparent biphasic formation of $Man_5GlcNAc_2$ we extended the consecutive mechanism (figure 4A) by assuming the formation of two different $Man_6GlcNAc_2$ isoforms and therefore adding a branch point after $Man_7GlcNAc_2$. This resulted in two additional apparent rate constants k_3* and k_4* accounting for formation and disappearance of the postulated, second $Man_6GlcNAc_2$ isomer (figure S9A). In comparison to the simpler, unbranched consecutive mechanism from above (figure 4A), the fits obtained using this branched mechanism were clearly improved (in particular with respect to the kinetics of $Man_6GlcNAc_2$/$Man_6GlcNAc_2$* and $Man_5GlcNAc_2$ glycoform formation at site 1, 2, 3 and 5), indicating that processing of $Man_9GlcNAc_2$ by GM I could indeed proceed via a mechanism that involved more than just one $Man_6GlcNAc_2$ isomer (figure S9B). However, our analytical methods do not discriminate between different glycan isomers. Therefore, inclusion of the additional rate constants k_3* and k_4* in the analysis resulted in an underdetermined system, as most obvious in case of k_4* of site 1 and 3 (table S5) and precluded a more thorough analysis. For site 4, $Man_5GlcNAc_2$ formation was too slow to observe two phases. Expectedly, fitting the site 4 data according to the branched mechanism did not show an improvement compared to using the unbranched one.

Freezing of PDI

To determine if freezing, storage and thawing of PDI would have an influence on the site-specific processing of its *N*-glycans, *in vitro* assays with GM I were performed. Therefore, PDI was either used freshly after purification, after being flash-frozen and immediately thawed and after being flash-frozen and stored for ten days. Figure S10 shows that no differences in site-specific processing occurred after freezing, storage and thawing of PDI.

Purification of Glycan Processing Enzymes

Both, the substrate of the *in vitro* assays PDI and the glycan processing enzymes ER Man I, GM I and GnT I were produced in High-Five™ cells using the insect cell/baculovirus system [6]. PDI was retained in the ER and purified like described before [7]. The glycan processing enzymes were secreted and purified from the supernatant *via* their His-tag (figure S11).

2. Supplementary Figures

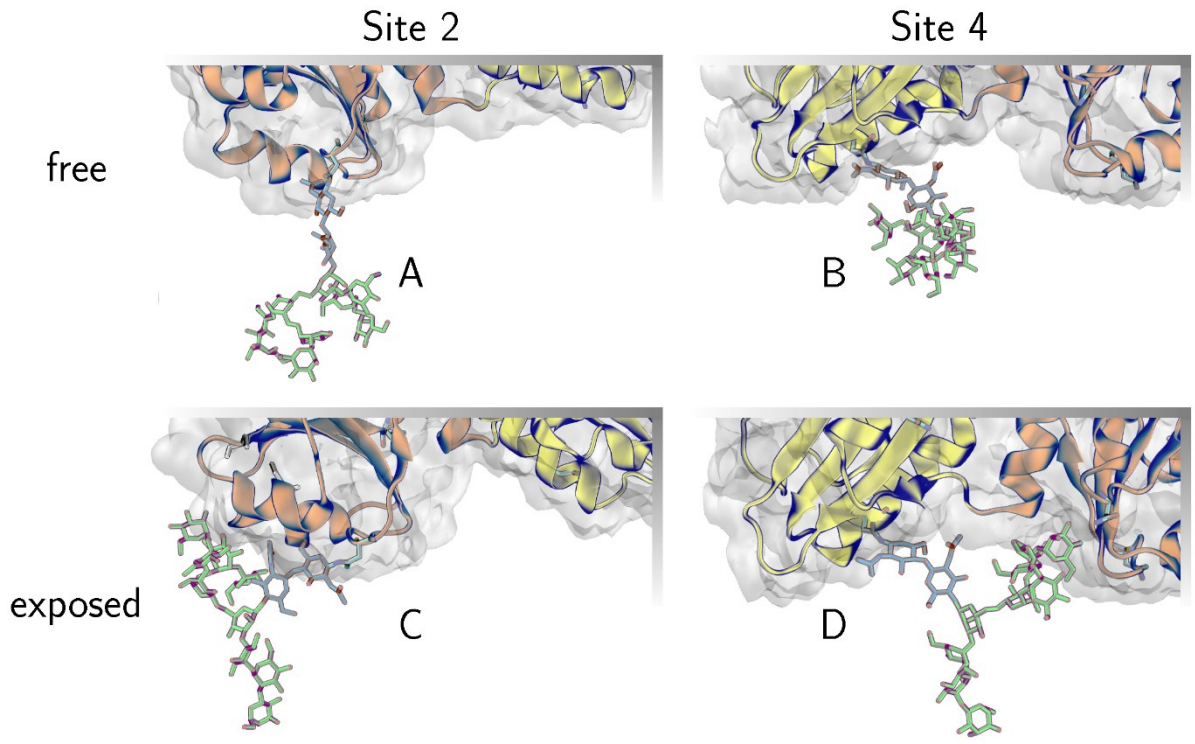


Figure S1: Convex and concave protein surface on sites 2 and 4

The convex and concave protein surface at sites 2 and 4 represent different glycan accessibility.

A: The positive curvature of the protein surface at site 2 allows extended branch conformations if the whole glycan is classified as 'free'.

B: In contrast, the glycan at site 4 is confined by the concave protein surface environment, such that even 'free' conformations are still dominated by collapsed conformations. Hence, at site 4 the steric hindrance of a concave protein surface diminishes the accessible space for 'free & exposed' glycan conformations

C: Extended branch conformations are particularly 'exposed' at site 2 even if the glycan is in 'contact'.

D: While at site 4 individually 'exposed' branches are still relatively buried inside the concave protein environment if the glycan is in 'contact' with the protein surface.

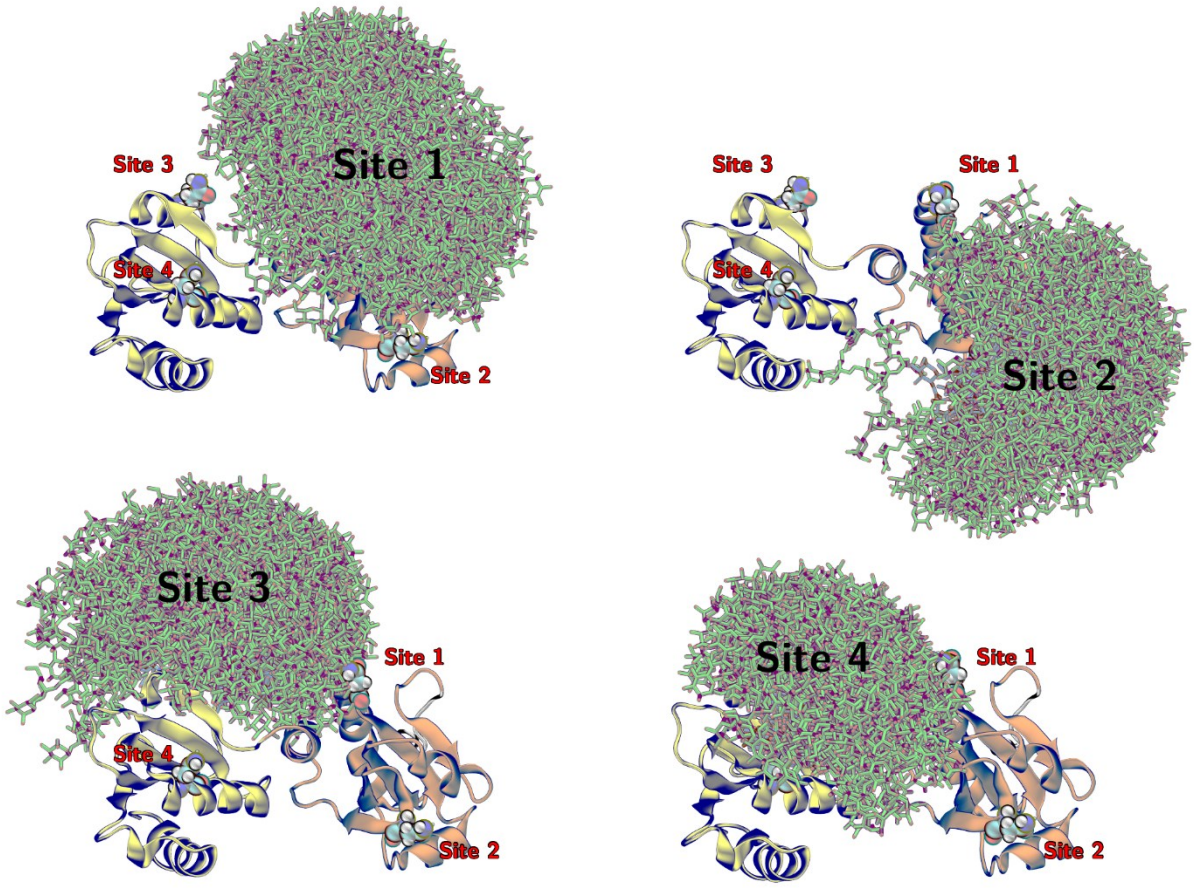


Figure S2: Site-specific *N*-glycan ensemble from simulations

A ribbon representation of the a- and b-domain with the ensembles of glycan conformations for sites 1-4. A random choice of 175 frames are used to illustrate the *N*-glycan ensemble on each site, respectively. The underlying protein domains are taken from one frame. Note that the a- and b-domain are flexibly connected. Hence, the respectively neighbouring domain also dominantly varies in its relative position to a given glycan ensemble.

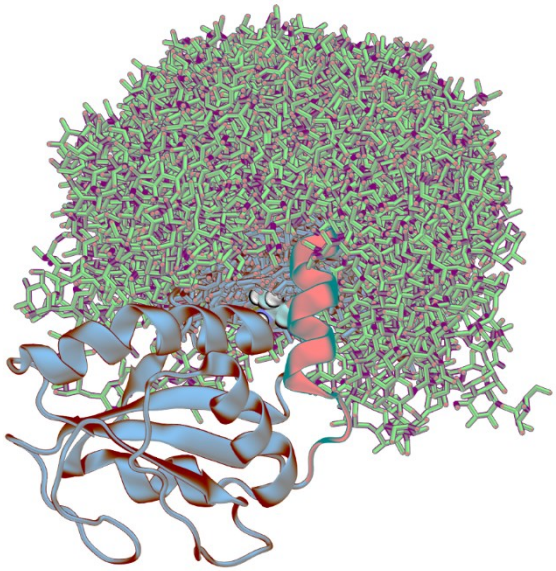


Figure S3: *N*-glycan ensemble on site 5 from simulations

The ensemble of the glycan on site 5 is represented by 175 random frames from simulations. The a'-domain is represented by a single simulation frame. Note, that the C-terminal helix overlaps with the ensemble. This helix is particularly flexible itself and interacts with the glycan in various conformations.

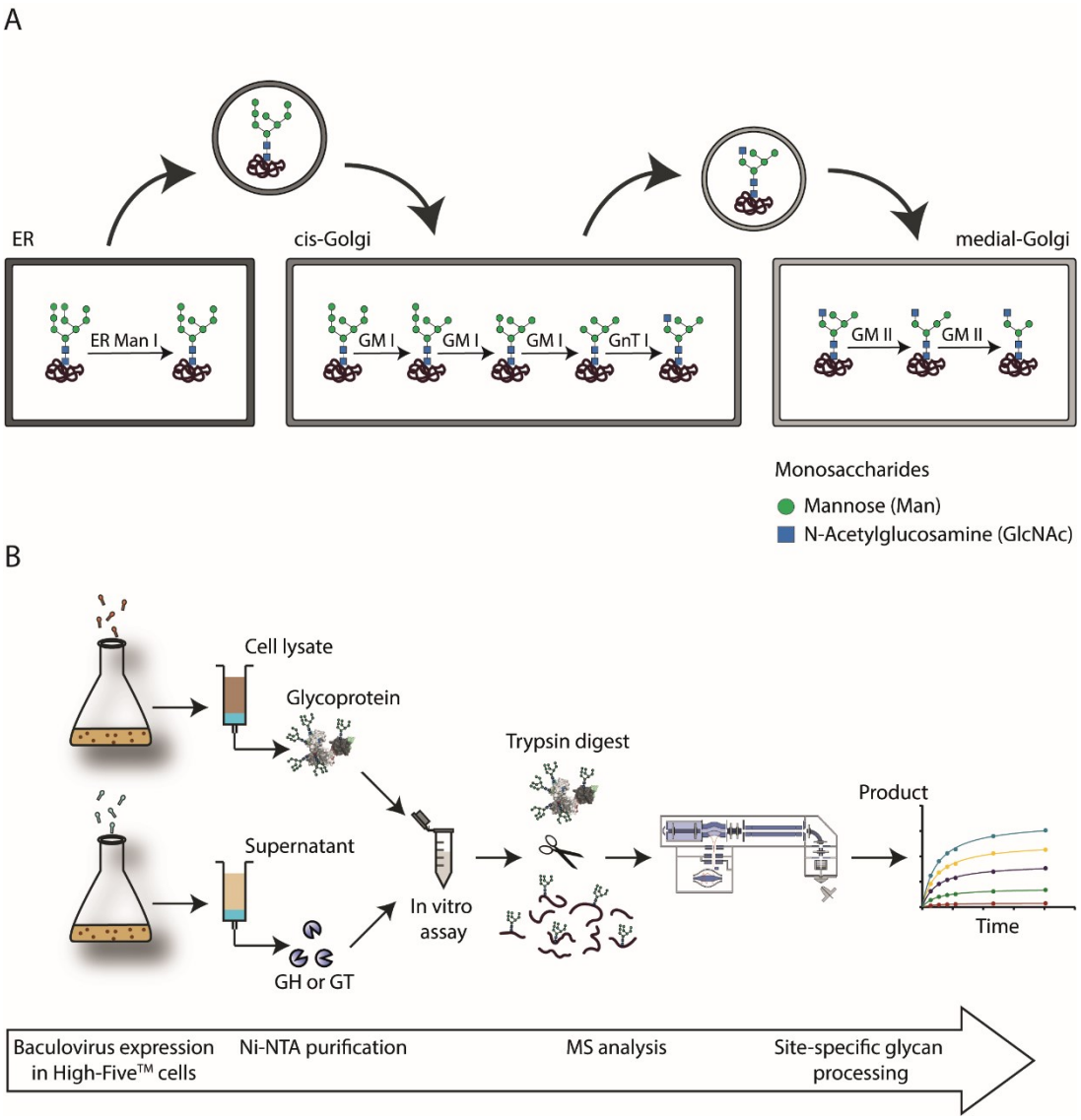


Figure S4: *N*-glycan processing *in vivo* and *vitro*

A: Simplified representation of the N-glycan processing steps in mammalian cells. Shown are only the reactions relevant to this study, catalyzed by ER Man I, GM I, GnT I and GM II. Vesicles transport the processed glycoproteins between the different compartments (ER, cis-Golgi and medial-Golgi). This transport defines the time glycoproteins are exposed to the respective glycan-processing enzymes.

B: Workflow of *in vitro* assays to study site-directed glycan processing. The model protein PDI and different glycan processing enzymes, glycosylhydrolases (GH) and a glycosyltransferase (GT), were produced in High-Five™ cells using the Baculovirus expression system. PDI was retained in the ER of the cells and therefore purified from the cell lysate. Glycan-processing enzymes were secreted and therefore purified from the supernatant with metal chelate (Ni-NTA) affinity chromatography. For the *in vitro* assay PDI and the respective enzymes were incubated, and the reactions were stopped by TCA

precipitation after different incubation times. Tryptic glycopeptides were analyzed by MS and product formation over time was quantified individually for each site.

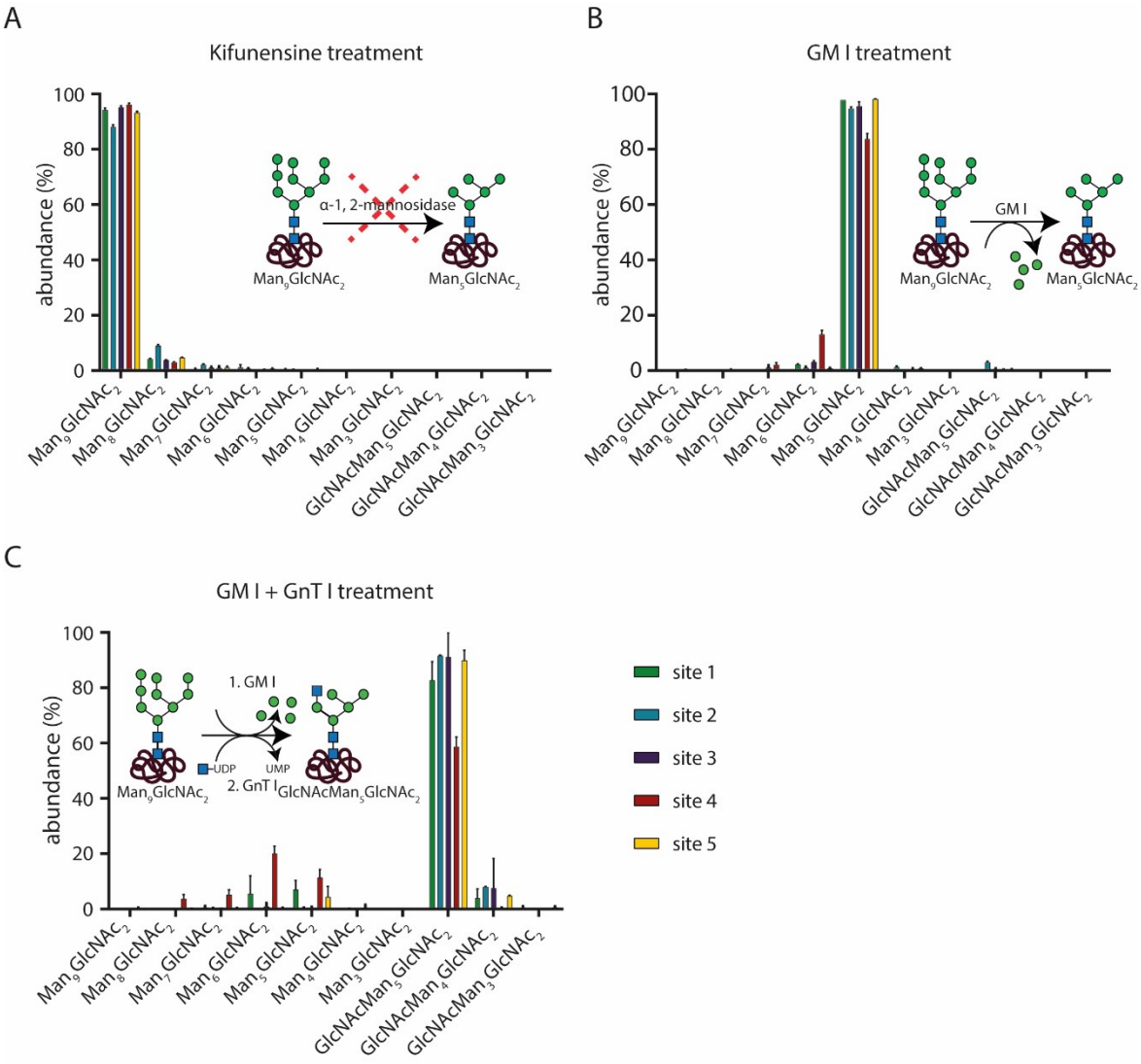


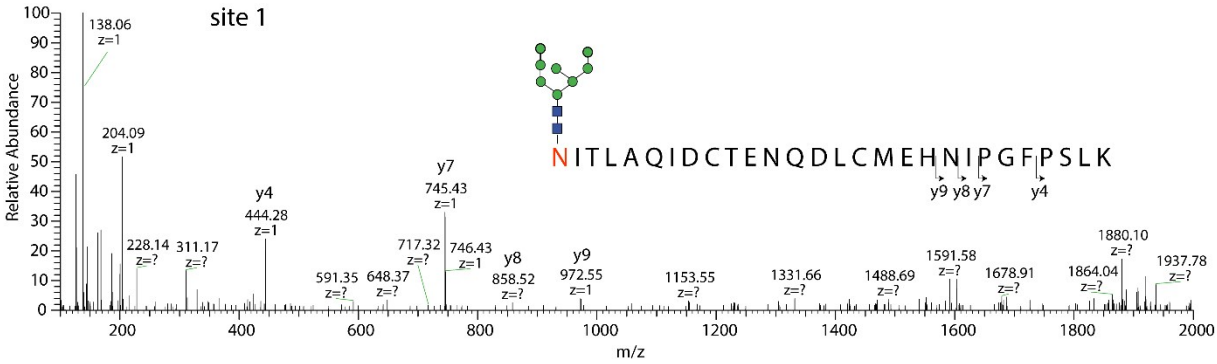
Figure S5: Generation of the PDI glycosubstrate for different glycan processing enzyme *in vitro* assays

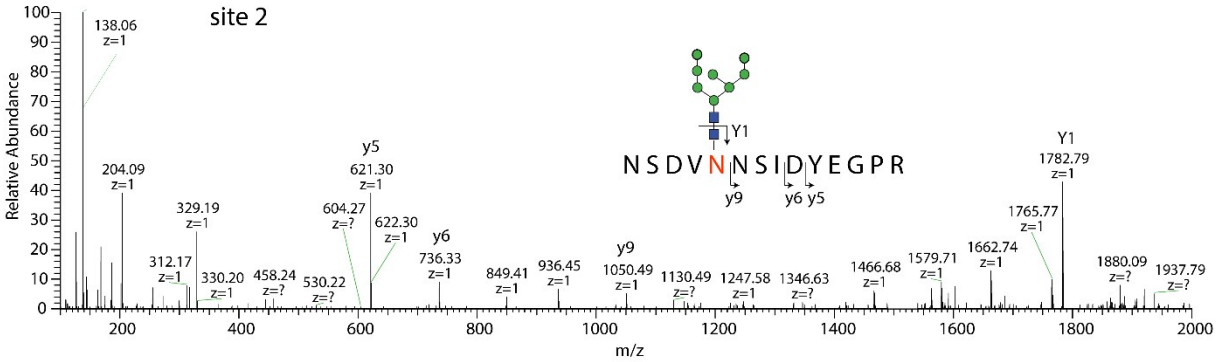
The four different glycan processing enzymes used in the *in vitro* assays require different glyco-substrates. PDI was therefore treated according to these requirements (*n*=3).

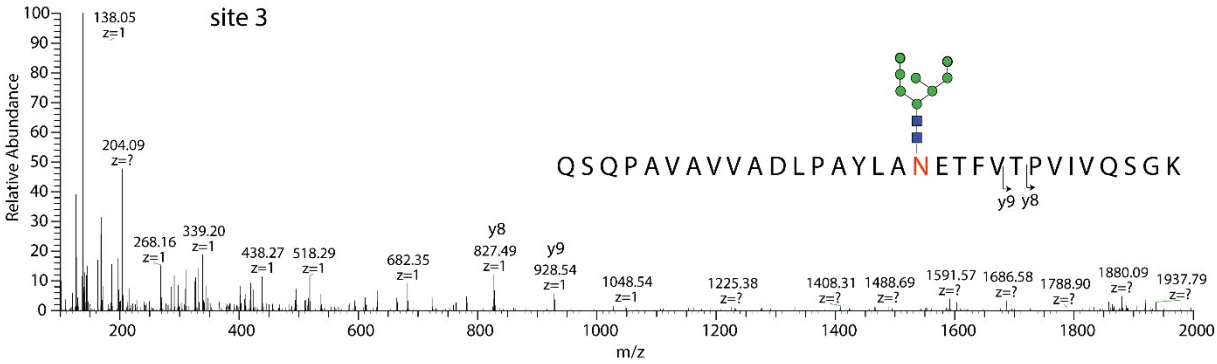
A: Substrate production for ER Man I and GM I *in vitro* assays. During PDI production in High-Five™ cells, the α-1, 2-mannosidase inhibitor kifunensine was added to the culture. It prevented mannose trimming in the ER and cis-Golgi. PDI purified from cells treated with kifunensine therefore showed a homogenous glycoform distribution (mainly $Man_9GlcNAc_2$ on all sites).

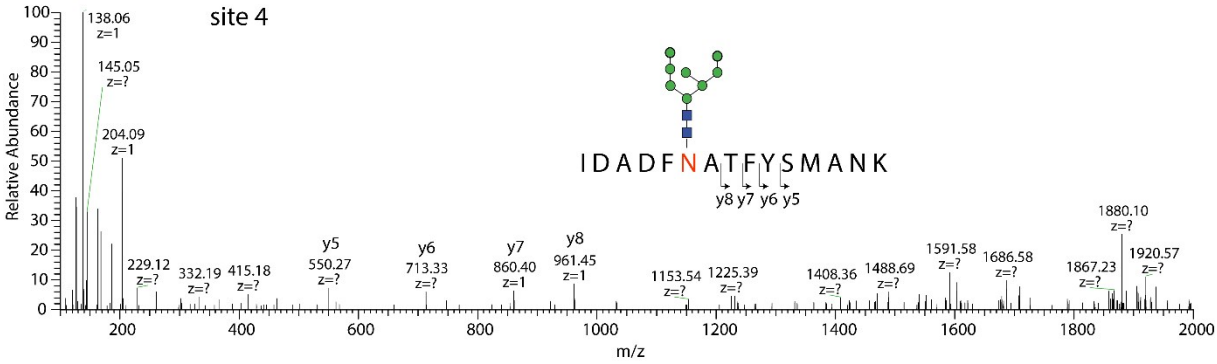
B: Substrate production for GnT I *in vitro* assay. Purified PDI produced by High-Five™ cells not treated with kifunensine was incubated over night with GM I. Naturally occurring site-specific differences in glycoform distribution were adjusted by the mannose trimming activity of GM I, resulting in mainly $Man_5GlcNAc_2$ (the substrate of GnT I) on all sites.

C: Substrate production for GM II *in vitro* assay. Purified PDI produced by High-Five™ cells not treated with kifunensine was incubated with GM I. After buffer exchange to GnT I activity buffer, PDI mainly bearing $Man_5GlcNAc_2$ on all sites was incubated with GnT I to obtain the glyco-substrate of GM II, $GlcNAcMan_5GlcNAc_2$.


site 1
NITLAQIDCTENQDLCMEHNIPGFPSLK
y9 y8 y7 y4
Relative Abundance
m/z


site 2
NSDVNNSIDYEGPR
Y1
y9
y6 y5
Relative Abundance
m/z


site 3
QSQPAVAVVADLPAYLANETFVTPVIVQSGK
y9 y8
Relative Abundance
m/z


site 4
IDADFNATFYSMANK
y8 y7 y6 y5
Relative Abundance
m/z

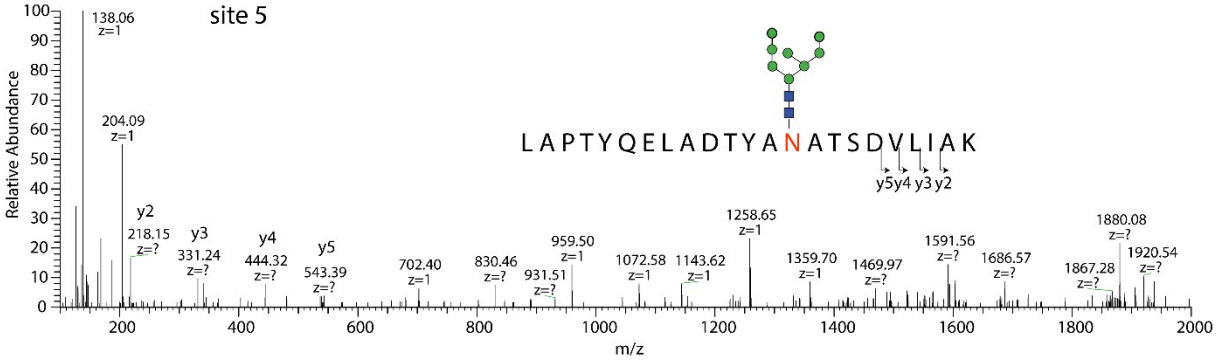

site 5
LAPTYQELADTYANATSDVLIAK
y5y4 y3 y2
Relative Abundance
m/z

Figure S6: MS/MS spectra of the five PDI *N*-glycosylation sites

The five panels show the MS/MS spectra of five different tryptic PDI peptides. Each peptide contains one *N*-glycosylation sites with a $Man_8GlcNAc_2$ glycan attached to the asparagine residue in the Asn-X-Ser/Thr *N*-glycosylation sequon (marked in red). HCD MS/MS spectra were performed with a target value of $5e^5$ by the collision energy setup at a normalized collision energy 40.

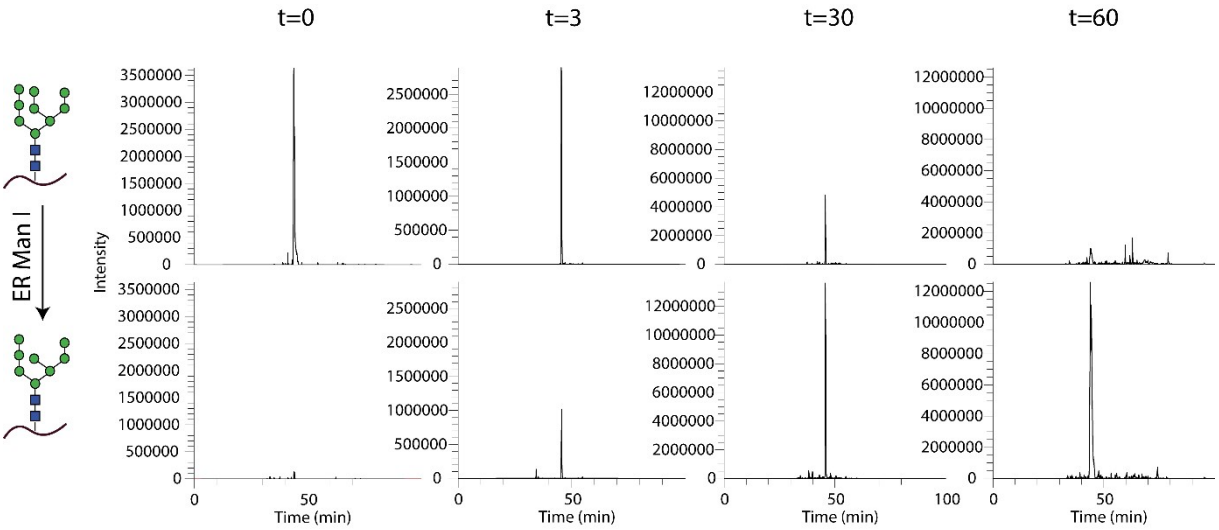


Figure S7: XIC of site 4 *in vitro* processing by ER Man I

PDI purified from kifunensine treated cells was incubated with ER Man I. Shown are the extracted ion chromatograms (XIC) of the site 4 PDI glycopeptide at different time points. The top panel shows the decrease of $Man_9GlcNAc_2$ while the lower panel shows the respective increase in $Man_8GlcNAc_2$.

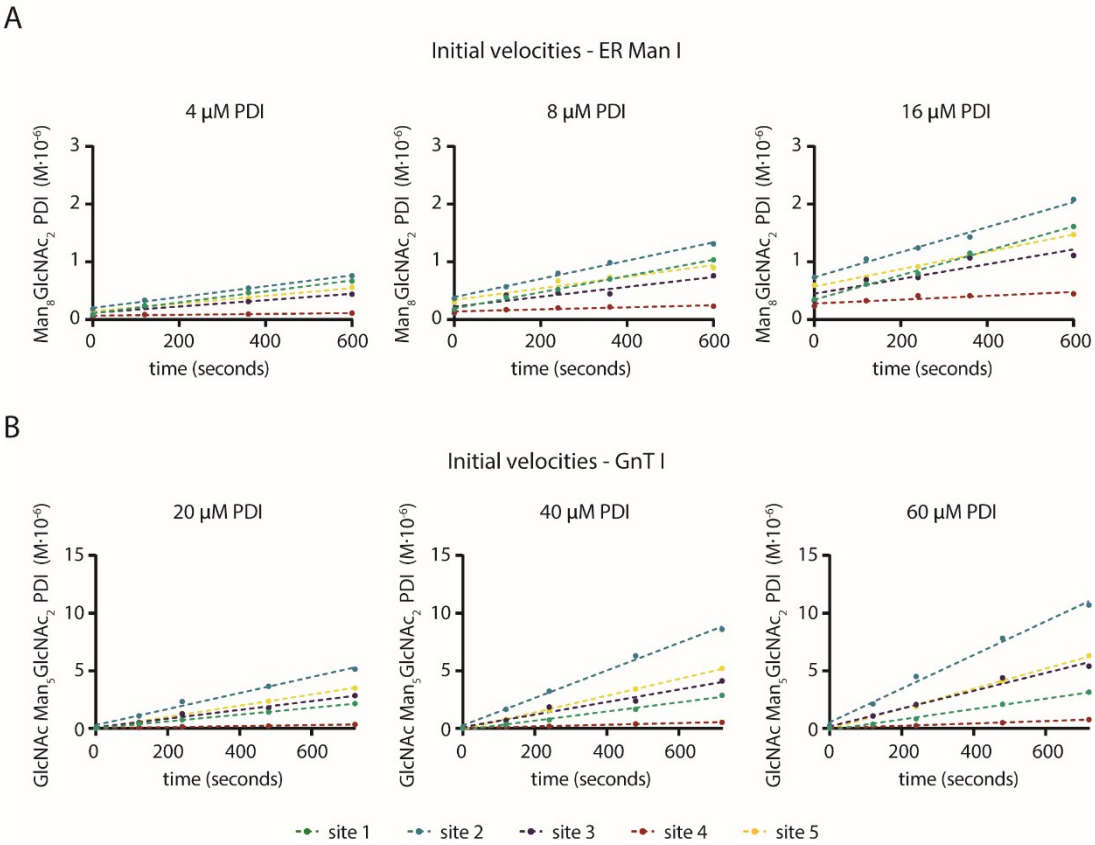


Figure S8: Linear phase of reaction for ER Man I and GnT I at different initial concentrations of PDI

A: ER Man I: The substrate $Man_9GlcNAc_2$-PDI (4, 8 and 16 µM) was incubated in presence of 0.6 nM ER Man I. The generation of the product $Man_8GlcNAc_2$ (in 10^{-6} M) on the five different sites is plotted against the incubation time (*n*=1). Linear regression lines were fitted using GraphPad Prism software.

B: GnT I: The substrate $Man_5GlcNAc_2$-PDI (20, 40 and 60 µM) was incubated in presence of 67 nM GnT I and 5 mM UDP-GlcNAc. The generation of the product $GlcNAcMan_5GlcNAc_2$ (in 10^{-6} M) on the five different sites is plotted against the incubation time (*n*=1).

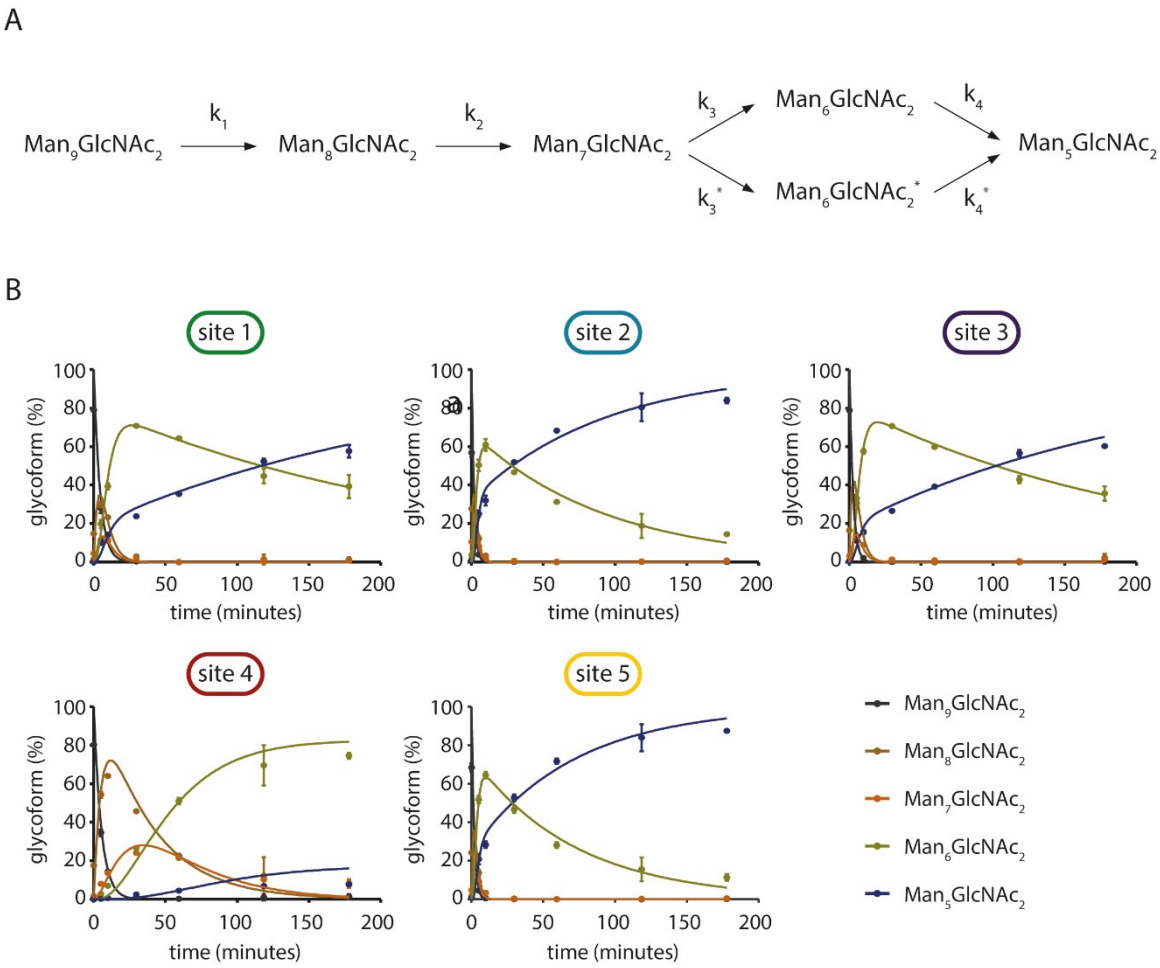


Figure S9: Site-specific processing of PDI-$Man_9GlcNAc_2$ by GM I

A: Consecutive mechanism for the stepwise removal of terminal mannoses from the $Man_9GlcNAc_2$ glycans of PDI, assuming a branch point after $Man_7GlcNAc_2$ and formation of two different $Man_6GlcNAc_2$ isomers. Conversion of the $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$, $Man_8GlcNAc_2$ to $Man_7GlcNAc_2$, $Man_7GlcNAc_2$ to Man_6 /$Man_{6*}GlcNAc_2$, and Man_6 /$Man_{6*}GlcNAc_2$ to the $Man_5GlcNAc_2$ glycoform is described by apparent rate constants k_1, k_2, k_3, k_3*, k_4 and k_4*.

B: Kinetics of glycan processing for each of the five glycosylation sites of PDI measured at 37 °C and pH 7. Initial concentrations were 20 µM PDI-$Man_9GlcNAc_2$ and 0.1 µM GM I (*n*=3). Solid lines represent a global fit of the data from each glycosylation site according to the mechanism in (**A**).

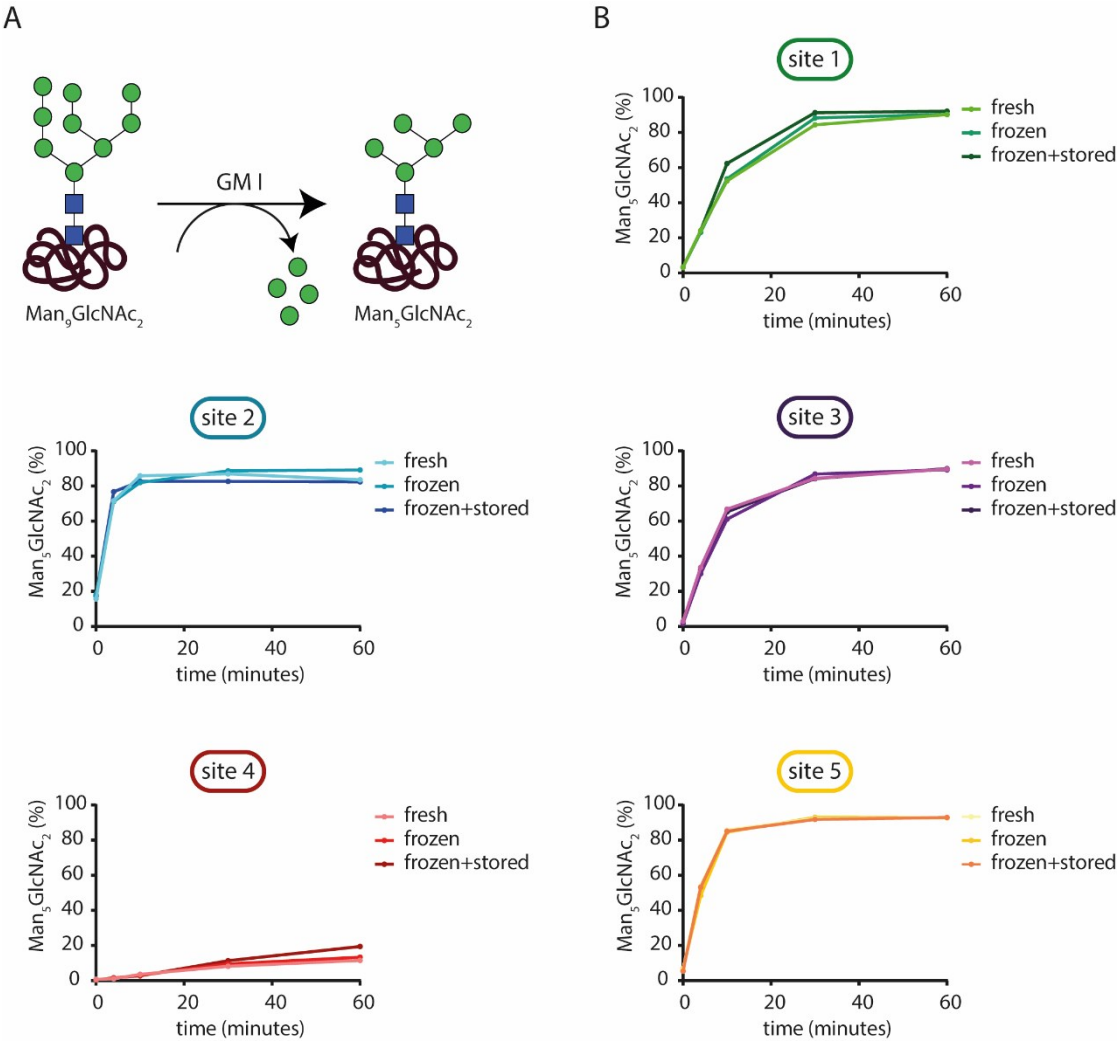


Figure S10: Site-specific processing of frozen PDI

A: Reaction catalyzed by GM I: Removal of four α-1, 2-linked mannoses from $Man_9GlcNAc_2$ glycan attached to PDI resulting in $Man_5GlcNAc_2$.

B: After Ni-NTA purification, PDI was buffer exchanged to GM I activity buffer and either used immediately for a GM I *in vitro* assay, flash frozen in liquid nitrogen, thawed and used for the assay or flash frozen and stored for ten days and then used for the assay. Shown is the conversion of $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$ on each site over 60 minutes. Samples were taken at indicated time points (*n*=1).

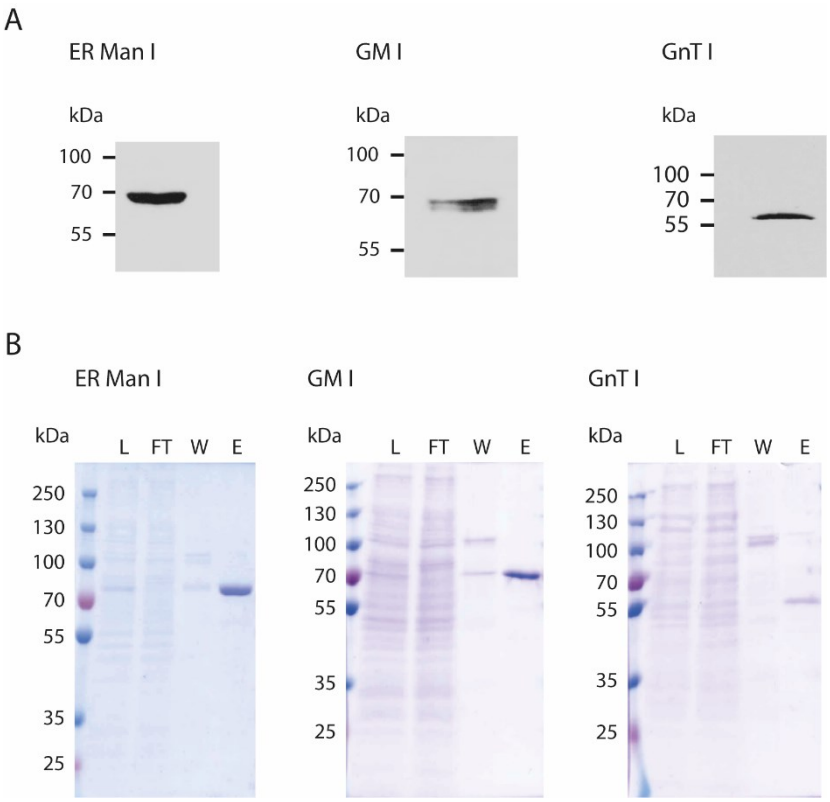


Figure S11: Production and purification of glycan processing enzymes

A: Supernatant of High-Five™ cell culture infected with baculovirus encoding for ER Man I, GM I or GnT I respectively was probed with an anti-His antibody and subsequently with an anti-mouse HRP-conjugated antibody. The immunoblot shows expression and secretion of the glycan processing enzymes encoded by the three different baculovirus stocks.

B: The different steps of Ni-NTA purification of three different glycan processing enzymes (ER Man I, GM I and GnT I) were analyzed by SDS-PAGE. The gels show Load (L) which is the supernatant of a baculovirus infected High-Five™ cell culture, Flow Through (FT) which is material not binding to the Ni-NTA beads, Wash (W) which is material unspecifically bound to Ni-NTA beads and washed off with washing buffer and Elution (E) which is the glycan processing enzyme of interest which specifically binds *via* its His-tag to the Ni-NTA beads and can be eluted by the addition of the elution buffer.

3. Supplementary Tables

Table S1. Glycosite occupancy of PDI produced in High-Five™ cells using the insect cell/baculovirus system.

Site	Sequence	m/z	Peak area	Site occupancy
1	NITLAQIDCTENQDLCMEHNIPGFPSLK	1086.8429^{3+}	15518895646	66.2%
	NITLAQIDCTENQDLCMEHNIPGFPSLK + HexNAc	1154.5364^{3+}	30400801778	
2	NSDVNNSIDYEGPR	790.3498^{2+}	61943778262	60.7%
	NSDVNNSIDYEGPR + HexNAc	891.8904^{2+}	95964490971	
3	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	1071.5837^{3+}	17604299292	71.2%
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK + HexNAc	1139.2675^{3+}	43465051276	
4	IDADFNATFYSMANK	854.3875^{2+}	2.07751E+11	78.2%
	IDADFNATFYSMANK + HexNAc	955.9257^{2+}	51943606640	
5	LAPTYQELADTYANATSDVLIAK	1234.635^{2+}	51943606640	65.8%
	LAPTYQELADTYANATSDVLIAK + HexNAc	1336.1693^{2+}	1.00148E+11	

Table S2. Site-specific initial ER Man I velocities (in 10^{-9} M/s ± SEM) for different substrate (PDI) concentrations.

	site 1	site 2	site 3	site 4	site 5
4 µM PDI	0.95 ± 0.05	0.94 ± 0.04	0.55 ± 0.08	0.08 ± 0.02	0.67 ± 0.04
8 µM PDI	1.41 ± 0.04	1.58 ± 0.07	0.85 ± 0.18	0.18 ± 0.05	1.01 ± 0.14
16 µM PDI	2.12 ± 0.09	2.17 ± 0.14	1.28 ± 0.31	0.33 ± 0.1	1.5 ± 0.1

Table S3. Site-specific initial GnT I velocities (in 10^{-9} M/s ± SEM) for different substrate (PDI) concentrations.

	site 1	site 2	site 3	site 4	site 5
20 µM PDI	2.95 ± 0.04	6.94 ± 0.46	3.80 ± 0.32	0.45 ± 0.05	4.9 ± 0.1
40 µM PDI	3.98 ± 0.25	11.9 ± 0.44	5.43 ± 0.57	0.72 ± 0.06	7.26 ± 0.16
60 µM PDI	4.5 ± 0.2	14.57 ± 0.72	7.72 ± 0.64	1.03 ± 0.07	8.69 ± 0.21

Table S4. Apparent rate constants k_1, k_2, k_3 and k_4 for formation of the $Man_8GlcNAc_2$, $Man_7GlcNAc_2$, $Man_6GlcNAc_2$ and $Man_5GlcNAc_2$ glycoforms by GM I on each glycosylation site, respectively.

	site 1	**site 2**	**site 3**	**site 4**	**site 5**
k_1 (min^{-1})	0.21 ± 0.03	0.43 ± 0.18	0.37 ± 0.08	0.19 ± 0.02	0.45 ± 0.16
k_2 (min^{-1})	0.22 ± 0.04	0.57 ± 0.28	0.26 ± 0.05	$(2.7 \pm 0.3)\cdot10^{-2}$	0.55 ± 0.21
k_3 (min^{-1})	0.45 ± 0.16	1.10 ± 1.0	0.59 ± 0.20	$(4.1 \pm 0.7)\cdot10^{-2}$	1.10 ± 0.7
k_4 (min^{-1})	$(7.0 \pm 0.7)\cdot10^{-3}$	$(2.4 \pm 0.4)\cdot10^{-2}$	$(7.6 \pm 0.7)\cdot10^{-3}$	$(1.7 \pm 0.4)\cdot10^{-3}$	$(2.7 \pm 0.4)\cdot10^{-2}$

Table S5. *Apparent rate constants k_1, k_2, k_3, k_3*, k_4 and k_4* for formation of the $Man_8GlcNAc_2$, $Man_7GlcNAc_2$, Man_6 /Man_6*$GlcNAc_2$ and $Man_5GlcNAc_2$ glycoforms on each glycosylation site by GM I.*

Apparent rate constant	**site 1**	**site 2**	**site 3**	**site 4**	**site 5**
k_1 (min^{-1})	0.23 ± 0.03	0.63 ± 0.41	0.41 ± 0.07	0.19 ± 0.02	0.60 ± 0.28
k_2 (min^{-1})	0.25 ± 0.04	0.65 ± 0.36	0.29 ± 0.04	$(2.7 \pm 0.3)\cdot10^{-2}$	0.62 ± 0.24
k_3 (min^{-1})	0.46 ± 0.11	0.88 ± 0.80	0.60 ± 0.18	$(0.8 \pm 4.1)\cdot10^{-2}$	0.90 ± 0.58
k_{3*} (min^{-1})	0.13 ± 0.08	0.48 ± 0.44	0.17 ± 0.06	$(3.6 \pm 4.4)\cdot10^{-2}$	0.38 ± 0.25
k_4 (min^{-1})	$(4.1 \pm 0.7)\cdot10^{-3}$	$(1.1 \pm 0.3)\cdot10^{-2}$	$(4.7 \pm 0.7)\cdot10^{-3}$	$(2.9 \pm 21)\cdot10^{-2}$	$(1.4 \pm 0.3)\cdot10^{-2}$
k_{4*} (min^{-1})	$(0.065 \pm 87)\cdot10^{4}$	1.6 ± 3.8	$(0.61 \pm 68)\cdot10^{2}$	$1\cdot10^{-8} \pm 8\cdot10^{-3}$	1.5 ± 3.3

Table S6. *Apparent rate constants k_1 and k_2 for formation of the $GlcNAcMan_4GlcNAc_2$ and $GlcNAcMan_3GlcNAc_2$ glycoforms by GM II on each glycosylation site, respectively.*

	site 1	**site 2**	**site 3**	**site 4**	**site 5**
k_1 (min^{-1})	0.19 ± 0.01	0.18 ± 0.01	0.18 ± 0.02	$(7.0 \pm 0.1)\cdot10^{-3}$	0.21 ± 0.01
k_2 (min^{-1})	0.27 ± 0.03	0.24 ± 0.03	0.15 ± 0.02	$(2.7 \pm 0.4)\cdot10^{-3}$	0.31 ± 0.03

Table S7. *Site specific choice of neighboring glycan and protein residues for the featurization step. Column one labels the considered site/glycan. The second column lists the neighboring glycans and the third column lists the residue indices of amino acids that are considered for the inverse distance featurization.*

Site	Glycan – Glycan	Glycan – Amino Acid
1	2, 3, 4	32 – 52, 87 – 107, 112 – 162
2	1, 4	32 – 52, 87 – 107, 112 – 126, 167 – 186, 218 – 232, 245 – 255, 317 – 327
3	1, 4	32 – 52, 87 – 107, 127 – 162, 187 – 192, 268 – 285, 345 – 355
4	1, 2, 3	32 – 52, 87 – 107, 112 – 192, 218 – 232
5		405 – 429, 483 – 504

Table S8. *Amino-acid specific side-chain atoms that were used in hexose-amino acid distance calculations. The first column lists the amino acids and the second column lists the respective side-chain atoms, which were taken for the inverse distance calculation to respective O_5 hexose atoms.*

Amino Acid	Side-Chain Atom
Arginine (ARG)	C_ζ
Asparagine (ASN)	C_γ
Aspartic Acid (ASP)	C_γ
Glutamine (GLN)	C_δ
Glutamic Acid (GLU)	C_δ
Histidine (HID)	$C_{\epsilon 1}$
Lysine (LYS)	N_ζ
Serine (SER)	O_γ
Threonine (THR)	$O_{\gamma 1}$
Tryptophan (TRP)	$N_{\epsilon 1}$
Tyrosine (TYR)	O_η

Table S9. *Related to material and methods. Primers used to generate PDI constructs.*

Purpose	**Primer name**	**Sequence 5` to 3`**
Removal of "LELQLEHDEL" sequence	prHI_14_fw	TAA AAG CTT GTC GAG AAG TAC TAG AGG ATC ATA ATC AGC CAT ACC
	prHI_15_rev	CAA TTC ATC GTG AAT GGC ATC TTC TTC GTC AGC CAA TTC AG

Table S10. *Related to material and methods. Composition of enzyme activity buffers used for in vitro assays of ER Man I, GM I, GnT I and GM II.*

Name	**Composition**	**pH**	**Reference**
ER Man I Activity buffer	20 mM MES; 150 mM NaCl; 5 mM $CaCl_2$	6.7	Aikawa *et al.* [8]
GM I Activity buffer	PBS, 0.9 mM $CaCl_2$, 0.5 mM $MgCl_2$	7	This study
GnT I Activity buffer	50 mM MES, 150 mM NaCl, 10 mM $MnCl_2$	6.7	Geisler *et al.* [9]
GM II Activity buffer	50 mM MES, 150 mM NaCl, 10 mM $MnCl_2$	6.7	This study

Table S11. *Related to material and methods. Amino acid sequence, expected glycan structures, charge state and mass over charge ratios (m/z) of the five tryptic peptides of PDI containing the N-glycosylation sites.*

	Peptide sequence	**Glycan structure**	**Charge**	**m/z**
site 1				
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$Man_9GlcNAc_2$	4	1295.7967
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$Man_8GlcNAc_2$	4	1255.2814
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$Man_7GlcNAc_2$	4	1214.7679
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$Man_6GlcNAc_2$	4	1174.2572
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$Man_5GlcNAc_2$	4	1133.7435
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$Man_4GlcNAc_2$	4	1093.23
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$Man_3GlcNAc_2$	4	1052.7254
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$GlcNAcMan_5GlcNAc_2$	4	1184.5135
	NITLAQIDCTENQDLCMEHNIPGFPSLK	$GlcNAcMan_4GlcNAc_2$	4	1143.9945

	NITLAQIDCTENQDLCMEHNIPGFPSLK	$GlcNAcMan_3GlcNAc_2$	4	1103.4802
site 2				
	NSDVNNSIDYEGPR	$Man_9GlcNAc_2$	3	1148.7865
	NSDVNNSIDYEGPR	$Man_8GlcNAc_2$	3	1094.7689
	NSDVNNSIDYEGPR	$Man_7GlcNAc_2$	3	1040.7494
	NSDVNNSIDYEGPR	$Man_6GlcNAc_2$	3	986.7343
	NSDVNNSIDYEGPR	$Man_5GlcNAc_2$	3	932.7157
	NSDVNNSIDYEGPR	$Man_4GlcNAc_2$	3	878.6966
	NSDVNNSIDYEGPR	$Man_3GlcNAc_2$	3	824.6952
	NSDVNNSIDYEGPR	$GlcNAcMan_5GlcNAc_2$	3	1000.4126
	NSDVNNSIDYEGPR	$GlcNAcMan_4GlcNAc_2$	3	946.3949
	NSDVNNSIDYEGPR	$GlcNAcMan_3GlcNAc_2$	3	892.3996
site 3				
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$Man_9GlcNAc_2$	4	1270.0997
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$Man_8GlcNAc_2$	4	1229.5882
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$Man_7GlcNAc_2$	4	1189.0757
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$Man_6GlcNAc_2$	4	1148.561
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$Man_5GlcNAc_2$	4	1108.049
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$Man_4GlcNAc_2$	4	1067.5275
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$Man_3GlcNAc_2$	4	1027.0229
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$GlcNAcMan_5GlcNAc_2$	4	1158.8162
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$GlcNAcMan_4GlcNAc_2$	4	1118.542
	QSQPAVAVVADLPAYLANETFVTPVIVQSGK	$GlcNAcMan_3GlcNAc_2$	4	1077.8029
site 4				
	IDADFNATFYSMANK	$Man_9GlcNAc_2$	3	1191.478
	IDADFNATFYSMANK	$Man_8GlcNAc_2$	3	1137.4576
	IDADFNATFYSMANK	$Man_7GlcNAc_2$	3	1083.4396
	IDADFNATFYSMANK	$Man_6GlcNAc_2$	3	1029.4253
	IDADFNATFYSMANK	$Man_5GlcNAc_2$	3	975.407
	IDADFNATFYSMANK	$Man_4GlcNAc_2$	3	921.389
	IDADFNATFYSMANK	$Man_3GlcNAc_2$	3	867.3855
	IDADFNATFYSMANK	$GlcNAcMan_5GlcNAc_2$	3	1043.1003

	IDADFNATFYSMANK	$GlcNAcMan_4GlcNAc_2$	3	989.0826
	IDADFNATFYSMANK	$GlcNAcMan_3GlcNAc_2$	3	935.0586
site 5				
	LAPTYQELADTYANATSDVLIAK	$Man_9GlcNAc_2$	3	1444.9727
	LAPTYQELADTYANATSDVLIAK	$Man_8GlcNAc_2$	3	1390.9599
	LAPTYQELADTYANATSDVLIAK	$Man_7GlcNAc_2$	3	1336.9418
	LAPTYQELADTYANATSDVLIAK	$Man_6GlcNAc_2$	3	1282.9224
	LAPTYQELADTYANATSDVLIAK	$Man_5GlcNAc_2$	3	1228.9084
	LAPTYQELADTYANATSDVLIAK	$Man_4GlcNAc_2$	3	1174.8845
	LAPTYQELADTYANATSDVLIAK	$Man_3GlcNAc_2$	3	1120.8772
	LAPTYQELADTYANATSDVLIAK	$GlcNAcMan_5GlcNAc_2$	3	1296.5946
	LAPTYQELADTYANATSDVLIAK	$GlcNAcMan_4GlcNAc_2$	3	1242.5693
	LAPTYQELADTYANATSDVLIAK	$GlcNAcMan_3GlcNAc_2$	3	1188.5502

4. References

1. Berendsen, H.J.C., D. van der Spoel, and R. Vandrunen, *Gromacs - A Message-Passing Parallel Molecular-Dynamics Implementation.* Computer Physics Communications, 1995. **91**(1-3): p. 43-56.

2. Tribello, G.A., et al., *PLUMED 2: New Feathers for an Old Bird.* Computer Physics Communications, 2014. **185**(2): p. 604-613.

3. Wang, L.L., R.A. Friesner, and B.J. Berne, *Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2).* Journal of Physical Chemistry B, 2011. **115**(30): p. 9431-9438.

4. Hub, J.S., et al., *Quantifying Artifacts in Ewald Simulations of Inhomogeneous Systems with a Net Charge.* Journal of Chemical Theory and Computation, 2014. **10**(1): p. 381-390.

5. Wassenaar, T.A., et al., *SQUEEZE-E: The Optimal Solution for Molecular Simulations with Periodic Boundary Conditions.* Journal of Chemical Theory and Computation, 2012. **8**(10): p. 3618-3627.

6. Jarvis, D.L., *Chapter 14 Baculovirus–Insect Cell Expression Systems*, in *Methods in Enzymology*, R.R. Burgess and M.P. Deutscher, Editors. 2009, Academic Press. p. 191-222.

7. Hang, I., et al., *Analysis of site-specific N-glycan remodeling in the endoplasmic reticulum and the Golgi.* Glycobiology, 2015. **25**(12): p. 1335-1349.

8. Aikawa, J.-i., I. Matsuo, and Y. Ito, *In vitro mannose trimming property of human ER α-1,2 mannosidase I.* Glycoconjugate Journal, 2012. **29**(1): p. 35-45.

9. Geisler, C. and D.L. Jarvis, *Substrate Specificities and Intracellular Distributions of Three N-glycan Processing Enzymes Functioning at a Key Branch Point in the Insect N-Glycosylation Pathway.* Journal of Biological Chemistry, 2012. **287**(10): p. 7084-7097.