Malignant tissues produce divergent antibody glycosylation of relevance

for cancer Gene Therapy Effectiveness

Dominik Brücher^{1,†}, Vojtech Franc^{2,3,†}, Sheena N. Smith^{1,*}, Albert J. R. Heck^{2,3,*}, and Andreas Plückthun^{1,*}

1) Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

2) Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Utrecht, The Netherlands

3) Netherlands Proteomics Center, Utrecht, The Netherlands.

†: These authors contributed equally to this work

Corresponding authors:

Andreas Plückthun (<u>plueckthun@bioc.uzh.ch</u>), Albert J.R. Heck (<u>a.j.r.heck@uu.nl</u>) & Sheena Smith (<u>s.smith@bioc.uzh.ch</u>)

This PDF file includes:

Supplementary Methods

Supplementary Figure S1: Amino acid sequence of the trastuzumab construct

Supplementary Figure S2: Amino acid sequence of the nivolumab construct

Supplementary Figure S3: Homogenous expression of the trastuzumab and nivolumab antibodies

Supplementary Figure S4: Viral particle visualization and antibody expression levels of cancer cells upon viral vector gene expression

Supplementary Figure S5: Comparison of the glycosylation profiles of nivolumab and trastuzumab expressed from either transfected or adenovirus-infected CHO-S cells

Supplementary Figure S6: Glycosylation profiles of nivolumab and trastuzumab expressed in CHO and various human (cancer) cell lines

Supplementary Figure S7: Illustrative example of the quantitative analysis of tryptic glycopeptides of trastuzumab produced in SKBR3 cells

Supplementary Figure S8: Evaluation of the contribution of incompletely glycosylated antibodies produced by the different cell lines using native or peptide-centric MS data

Supplementary Figure S9: Transcript levels of FcRn and different Fc γ -receptors in the used cancer cell lines

Supplementary Figure S10: Transcript levels of ST6Gal1 in the used cancer cell lines

Supplementary Figure S11: Comparison of mannose and galactose profiles of the produced antibodies in the different cell lines.

Other supplementary materials for this manuscript include the following:

Supplementary Table S1: EXCEL file with 132 human glycans (EXCEL file: N-glycan 132 human)

Supplementary Table S2: Relative abundances of all quantified mAb-species as extracted from the native MS data (EXCEL file: Native mAb-species)

Supplementary Table S3: Relative abundances of all quantified peptides species and Skyline data (EXCEL file: Quantified peptides)

Supplementary Data 1:

Manually annotated HCD MS/MS spectra of representative tryptic (glyco)peptides with amino acid sequence TKPREEQYNSTYR (trastuzumab). Sequential fragmentation of the glycan part of the peptide reveals type of glycosylation and glycan composition.

(PDF)

Supplementary Data 2:

Manually annotated HCD MS/MS spectra of representative tryptic (glyco)peptides with amino acid sequence TKPREEQFNSTYR (nivolumab). Sequential fragmentation of the glycan part of the peptide reveals type of glycosylation and glycan composition.

(PDF)

Supplementary Materials and Methods

Chemicals and materials used for MS-based proteomic experiments

Tris-HCl, Tris(2-carboxyethyl)-phosphine (TCEP), sodium deoxycholate (SDC), trifluoroacetic acid (TFA) and trypsin were purchased from Sigma-Aldrich (Steinheim, Germany). LysC was obtained from Wako (Neuss, Germany). Formic acid (FA) was purchased from Merck (Darmstadt, Germany). Acetonitrile (ACN) was acquired from Biosolve (Valkenswaard, The Netherlands). The Oasis PRiME HLB plate was purchased from Waters (Etten-Leur, the Netherlands). Milli-Q water was produced by an in-house system (Millipore, Billerica, MA). For non-MS-based proteomics experiments, common chemicals were purchased from Sigma Aldrich.

Antibodies and cell lines

The commercially available therapeutics Herceptin[®] and Opdivo[®] were obtained from Genentech. Antibodies for western blotting included mouse anti-ST6GAL1 mAb (Invitrogen Cat. No. MA5-11900), mouse anti-GAPDH (Santa Cruz Cat. No. sc-32233), and goat antimouse IgG AP (Sigma Cat. No. A3562). Adherent HEK293 cells (ATCC[®] CRL-1573TM) were used for first-generation virus production. The cell line 116, used for gutless virus production,¹ was kindly provided by Philip Ng (Baylor College of Medicine, Houston). HT1080 cells that overexpress the fibroblast activating protein α (FAP) were kindly provided by Christian Münz (University of Zurich). The cell lines HEK293-F (Thermo Fisher: R79007), CHO-S (Thermo Fisher: R80007), A549 (ATCC[®] CCL-185TM), SKBR3 (ATCC[®] HTB-30TM), BT474 (ATCC[®] HTB-20TM), and Detroit-551 (ATCC[®] CCL-110TM), were obtained from the indicated vendor. All cell lines were cultivated according to the provider's recommendations, and cells were passaged no higher than 15 times from the original stocks.

Antibody cloning:

N-terminal flanking DNA of all antibodies contained a *NheI* restriction site followed by a mammalian Kozak consensus sequence (gccRccAUGG); the C-terminal flanking DNA contained two stop codons followed by a *XhoI* restriction site for subcloning. The plasmids for first-generation adenovirus production were modified from the AdEasy Adenoviral Vector System kit (Agilent Cat. No 240009) as follows: (a) The multiple cloning site (MCS) of the first-generation adenovirus cloning plasmid pShuttle was replaced via Gibson Assembly (New England Biolabs) with a designed synthetic cassette with the following elements into which the antibody constructs were inserted: CMV_promoter-*NheI*_restriction_site-*XhoI*_restriction_site-EMCV_IRES-eGFP-BGH_PolyA. (b) The plasmid containing the adenoviral genome, pAdEasy-1, was modified by the integration of a mutation into the hypervariable loop 7 (HVR7) of the hexon that prevents blood factor X binding and thus reduces liver infection, as previously described.²

Infectious particles determination:

24-well plates were seeded with 50,000 A549 cells per well 24 hours prior to infection in 0.5 mL DMEM medium supplemented with 10 % fetal calf serum (Animed Lot 2-01-F30-1). Purified virus solution was added to the cell medium (3 μ L per well), and cells were incubated at 37°C, 5% CO₂ for 2 hours before harvest. Following incubation, the viruscontaining medium was aspirated, and cells were washed with 1 ml PBS per well. The viral genomes were extracted from cells using a Genekam DNA isolation kit (SB0072) according to the manufacturer's protocol. Adenoviral genome numbers were determined by quantitative PCR with adenovirus serotype-5-specific primers (0.5 μ M each): Ad_fwd: 5'-GTGATAACCGTGTGCTGGAC-3' and Ad_rev: 5'-CAGCTTCATCCCATTCGCAA-3' as well as a sequence-specific Hex labeled probe: 5'-TCCGCGGCGTGCTGGACAGG-3'. The reaction was carried out in 20 μ L using IDT PrimeTime® Gene Expression Master Mix (Cat#: 1055770) and a Stratagene Mx3005P (Agilent) instrument with the following program: 10 min at 95°C, 0.5 min at 95°C, 1 min at 56°C and 1 at 72°C for 50 cycles, each cycle starting at the second step with DNA quantification after the third step. Ct values were determined by the MxPro Mx3005P software using a single threshold mode.

Western blotting:

Cell lines were grown to confluency (80-100%) and lysed in 1X Cell Culture Lysis Reagent (Promega Ref. No. E1531) containing 60 mM DTT and a cocktail of protease inhibitors (4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), 48 μ g/mL; pepstatin-A, 1 μ g/mL; and leupeptin, 5 μ g/mL). Laemmli sample buffer was added to 1X (Bio-Rad Cat. No. 161-0747), and samples were heated to 99°C for 10 minutes. For each cell lysate, 15 μ g total protein was loaded into each well of a Criterion TGXTM Precast Midi Protein Gel (4-20%, Bio-Rad Cat. No. 5678095) in 1x TGX buffer (Bio-Rad Cat. No. 161-0772), then transferred on to an Immobilon-P PVDF transfer membrane (Merck Cat. No. IPVH00010) using a Bio-Rad Trans-Blot Turbo Transfer system (2.5 A, 7 min). The membrane was blocked in 1X Sigma Casein Blocking Solution (Sigma Cat. No. B6429) in TBS buffer (20 mM Tris, 150 mM NaCl, pH 7.5) for 30 min at room temperature with gentle orbital shaking. After blocking, the membrane was cut at approximately 40 kDa, and incubated with either a mouse anti-ST6GAL1 mAb (1:1,000; top piece) or a mouse anti-GAPDH mAb (1:3,000; bottom piece) in antibody dilution buffer (1X Sigma Casein Blocking Solution in TBS-T buffer) for 12-16 hours at 4°C with gentle orbital shaking. Membranes were washed 4X over 20 minutes with TBS-T (1X TBS with 0.05% Tween-20), then incubated with a goat anti-mouse IgG AP-conjugated antibody (1:10,000) in antibody dilution buffer for 1 hour at room temperature. Membranes were washed 4X over 20 minutes with TBS-T, then imaged on a Vilber Fusion Fx Spectrometer after addition of CDP-Star® Chemiluminescent Substrate (Sigma Cat. No. C0712).

EM measurements:

Prior to imaging, Zeba Spin Desalting Columns (7K MWCO, Thermo Scientific Cat. No. 89877) were used to exchange the adenoviral storage buffer into 20 mM HEPES, 150 mM NaCl, pH 6.8. Viral particles were irradiated under a UV light for 1 min, then samples were negative stained by spotting 10 μ L droplets on copper grids (300 mesh, glow discharged) for 1 min, then overlaid with 10 μ L droplets of 1% uranyl acetate for 1 min. The excess uranyl acetate was removed with filter paper, and the grid was allowed to air dry. Grids were examined with a CM100 transmission electron microscope (Thermo Fisher Scientific, Eindhoven, The Netherlands) at an acceleration voltage of 80 kV using an Orius 1000 digital camera (Gatan, Munich, Germany). Samples were prepared and imaged at the Center for Microscopy and Image Analysis (ZMB) at the University of Zurich.

Enzyme-linked immunosorbent assay (ELISA)

Following antibody production, described in the online methods section Antibody production of the main text, cell supernatants were filtered through a 0.2 μ m filter and stored at -20 °C. Nunc MaxiSorp flat-bottom plates (Thermo Fisher Cat. No. 44-2504-21) were washed in PBS and then coated with 100 μ L per well of 20 nM HER2 extracellular domain (Sino Biological Cat. No. 10004-HOSH) or 100 nM PD1 (Acro Biosystems Cat. No. PD1-H5221) for 12-16 hours at 4°C with gentle rocking. Following coating, plates were blocked in 5% nonfat dry milk (Applichem Cat. No. A0830) made in PBS for 1 hour at 25°C with gentle rocking. Plates were washed five times with PBS-T (1X PBS pH 7.4 + 0.1% Tween) using a Biotek ELx405 plate washer (each cycle: aspirate, shake for 5 s at intensity level 3, soak for 15

s), then standards (Herceptin or Opdivo in PBS), blanks or cell culture supernatants (at 1:5 and/or 1:50 in PBS) were added in triplicate in 100 μ L volumes per well. Samples were incubated at 25°C for 1.5-2 hours with gentle rocking, and then plates were washed five times as before. For detection, goat anti-human Fc-AP (Sigma Cat. No. A9544) was diluted 1:2,000 in 2% nonfat dry milk made in PBS-T and 100 μ L was added to each well. Plates were incubated at 25°C for 1 hour with gentle rocking, then washed 5-times as before. Signal was detected following addition of 100 μ L/well 3 mM para-nitrophenyl-phosphate (4-NPP) substrate in 4-NPP buffer (50 mM NaHCO₃, 50 mM MgCl₂) by reading absorbance at 405 nm using 540 nm as a reference wavelength after approximately 20 minutes of incubation at 25°C. Absorbance was normalized to the reference wavelength and blank wells, and concentrations were determined by interpolation of a corresponding standard curve using a sigmoidal 4-parameter logistic fit [Sigmoidal, 4PL, X is log(concentration)] using Graphpad Prism 8 software for each reading.

Native MS data analysis

Raw spectra were deconvoluted to zero-charge spectra by Intact Mass software³ (Protein Metrics, CA, USA) using default settings. The mass range and *m/z* range were adjusted to display all glycoproteoform signals. For PTM annotations, average masses were used, including hexose/mannose/galactose (Hex/Man/Gal, 162.1424 Da), N-acetylhexosamine/Nacetylglucosamine (HexNAc/GlcNAc, 203.1950 Da), fucose (Fuc, 146.1430 Da), and Nacetylneuraminic acid (Neu5Ac, 291.2579 Da). All symbols used and nomenclature follows the recommendations of the Consortium for Functional Glycomics.⁴ Based on the known backbone masses of IgG samples (trastuzumab and nivolumab), the biosynthetic pathway of N-glycosylation, and data obtained from glycopeptide analysis, a matrix of possible glycoproteoforms was constructed. We made use of an in-house developed R script for semiautomated peak annotation as described previously.⁵ In short, we first generated a library of proteoforms with specific masses and probability ranks using data (masses and relative intensities) derived from the peptide-centric analysis. All proteoform mass combinations were ranked according to the intensity of the detected PTM peptides from which they were constructed, and the match between their theoretical and experimental masses. The masses of peaks were then matched to the generated library with the highest-ranked proteoforms. The generated lists of proteoforms were manually inspected further and filtered to obtain the most meaningful PTM compositions for each detected proteoform. This allowed us to visualize probable proteoforms in the native mass spectra with their tentative PTM composition. For

estimation of quantities of the major proteoforms, peak intensities from zero-charge deconvoluted spectra were obtained directly from Intact Mass Software.

Flow cytometry analysis:

For flow cytometry analysis 1.0 x 104 HEK293 cells were seeded per well in a 96-well plate one day prior to treatment. Cell infections were performed in triplicates for conditions using different dilutions (0, 1, 5 or 10 infectious particles per cell) of trastuzumab or nivolumab first-generation adenoviral vectors, encoding GFP transcriptionally coupled to antibody expression via an IRES domain. Cells were harvested 48 hours post-transfection by flushing cells of the Petri dish in the culture medium. Subsequently, cells were harvested by centrifugation for 1 min at 800 g at RT. Harvested cells were washed twice with ice-cold 1% PBS/BSA and resuspended in a volume of 120 μ L 1% PBS/BSA for analysis. GFP-reporter expression was monitored on an LSR II Fortessa flow cytometer counting 2,000 - 5,000 single gated cell events per sample. For GFP detection, the Alexa 488 channel was used, and obtained data were analyzed using the software FCS Express 5 Flow and Prism 5 (GraphPad).

LC-MS and MS/MS analysis

All peptides (300 fmol of IgG peptides), measured in triplicates, were separated and analyzed using an Agilent 1290 Infinity HPLC system (Agilent Technologies, Waldbronn, Germany) coupled in-line to an Orbitrap Q exactive HF mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Reversed-phase separation was accomplished using a 100 µm inner diameter, 2 cm trap column (in-house packed with ReproSil-Pur C18-AQ, 3 µm) (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) coupled to a 50 µm inner diameter, 50 cm analytical column (in-house packed with Poroshell 120 EC-C18, 2.7 µm) (Agilent Technologies, Amstelveen, The Netherlands). Mobile-phase solvent A consisted of 0.1% formic acid in water, and mobile-phase solvent B consisted of 0.1% formic acid in acetonitrile. The flow rate was set to 300 nL/min. A 66 min gradient was used as follows: 0-5 min, 100% solvent A; 13-44% solvent B within 35 min; 44-100% solvent B within 3 min; 100% solvent B for 5 min; 100% solvent A for 17 min. For the MS scan, the mass range was set from 375-2,000 m/z at a resolution of 60,000, and the automatic gain control (AGC) target was set to 3×10^{6} . For the MS/MS measurements, higher-energy collision dissociation (HCD) was used and performed with a normalized collision energy of 30%. For the MS/MS scan, the mass range was set from 200-2,000 m/z at a resolution of 30,000; the AGC target was set to 1×10^4 . The precursor isolation width was 1.4 Da, and the maximum injection time was set to 150 ms.

LC-MS and MS/MS data analysis

The raw data files containing MS/MS spectra of IgG peptides were processed using Byonic software (version 3.2.0) (Protein Metrics Inc.). The following parameters were used for data searches in Byonic: trypsin digestion with a maximum of 2 missed cleavages, precursor ion mass tolerance, 10 ppm; fragmentation type, HCD; fragment mass tolerance, 20 ppm; variable modifications: methionine oxidation. For glycan analysis, we used a Byonic database of 132 human glycans (Sup. Table 1). The maximum number of precursors per scan was set to one and the FDR as 1%. Only PSMs matched to HCD spectra with non-negligible error probabilities |log Prob|>4.0 and Byonic score >=200 were accepted for further manual data filtering. Inspection of the MS/MS data did not reveal evidence of N-glycolylneuraminic acid $(m/z \ 209.087 \text{ or } 308.098)$; therefore, sialylation was presumed to only occur in the Nacetylneuraminic acid variant. The FASTA file used for the peptide searches contained trastuzumab and nivolumab amino acid sequences (Sup. Fig. 1-2). The software Skyline (version 4.2.0.18305) was used for site-specific relative quantification of the N-peptide glycoforms in the Fc region. Two major and one less frequent Fc region peptides, EEQYNSTYR, TKPREEQYNSTYR, and PREEQYNSTYR (trastuzumab), and EEQFNSTYR, TKPREEQFNSTYR, and PREEQFNSTYR (nivolumab), were used for relative quantification of IgG glycoforms. All glycoforms of these peptides annotated by Byonic were loaded into the Skyline software interface and their retention time, ppm error to the theoretical mass and isotopic pattern were checked. Curation criteria were as followed: (1) \leq 5 ppm error to the theoretical mass, (2) eluting within \pm 1 min of the mean retention time for that peptide, (3) having no apparent overlapping isotopic patterns. The resulting list of glycopeptides with all their detected charge states was further used for relative quantification. The first three isotopes were taken from each manually validated peptide glycoform for the calculation of the peak areas. For each sample, peak areas of all glycoforms of all abovedescribed peptides with all detected charge states were summed up and set to 100%. The peak area calculated for each curated peptide glycoform was normalized to the total area of all glycoforms including naked peptide and thus, glycoform-relative abundances were obtained. The average peptide ratios from all measurements were taken as a final estimation of the abundance, and only glycoforms with relative abundance >0.5% were further considered. Additionally, we performed manual inspection of the resulting list of glycoforms by checking representative HCD MS/MS spectra of the glycoforms with amino acid sequence TKPREEQ(Y/F)STYR. Selected MS/MS spectra were checked for the presence of the group

of fragments with a mass difference pattern of 17-83-120, which is typical for the fragmentation of N-glycopeptides and serves for unambiguous confirmation of glycopeptide identity. Further manual sequencing of the glycan moiety provided more accurate glycan compositions of the final glycoform list. Here, reported glycan structures are depicted without the linkage type of the glycan units since our acquired MS/MS data do not directly provide such information. The resulting glycopeptide list with relative abundance was also later used in the processing of native spectra annotations and visualization of N-glycosylation patterns.

Supplementary Figures

HC signal sequence
MDWTWRFLFVVAAATGVQSEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLE VH
WVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWG VH CH1
QGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAV CH1 HINGE CH2
LQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKK <mark>v</mark> epkscdkthtcppcpapellggp CH2
SVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVV CH2
SVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCL CH3
VKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALH Furin site F2A site LC signal sequence
NHYTQKSLSLSPGKRKRRAPVKQTLNFDLLKLAGDVESNPGPMDMRVPAQLLGLLLWLSGARCD VL
IQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSR VL
SGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTAS ^{CL}
VVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEV CL

THQGLSSPVTKSFNRGEC

Figure S1. Amino acid sequence of the trastuzumab construct. The heavy chain (VH, CH1, CH2, CH3) is presented in dark red while the light chain (VL, CL) is presented in light red. The leader sequences of the heavy and light chains (green), the furin site (orange), and the ribosome skipping site F2A (black) are processed during antibody maturation.

MEFGLSWVFLVALFRGVQCQV	QLVESGGGVVQPGRSLRLDCKASGI	TFSNSGMHWVRQAPGKGL
∣VH		
EWVAVIWYDGSKRYYADSVKG	RETISRDNSKNTLELQMNSLRAEDT	AVYYCATNDDYWGQGTLV
.VH .CH1		
TVSSASTKGPSVFPLAPCSRS	TSESTAALGCLVKDYFPEPVTVSWN	SGALTSGVHTFPAVLQSS
СН1	Hinge	CH2
GLYSLSSVVTVPSSSLGTKTY	TCNVDHKPSNTKVDKRVESKYGPPC	PPCPAPEFLGGPSVFLFP
ICH2		
	VSOEDBEVOENWYVDGVEVHNAKTK	PREFORNSTYRVVSVLTV
CH2	СПЗ	
LHQDWLNGKEYKCKVSNKGLP	SSIEKTISKAKGQPREPQVYTLPPS	QEEMTKNQVSLTCLVKGF
ICH3		
YPSDIAVEWESNGQPENNYKT	TPPVLDSDGSFFLYSRLTVDKSRWQ	EGNVFSCSVMHEALHNHY
YPSDIAVEWESNGQPENNYKT	TPPVLDSDGSFFLYSRLTVDKSRWQ	EGNVFSCSVMHEALHNHY
YPSDIAVEWESNGQPENNYKT Furin site CH3 P2A-site	TPPVLDSDGSFFLYSRLTVDKSRWQ _I LC signal sequen	EGNVFSCSVMHEALHNHY ce _I VL
YPSDIAVEWESNGQPENNYKT CH3	TPPVLDSDGSFFLYSRLTVDKSRWQ	EGNVFSCSVMHEALHNHY ce VL
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen	EGNVFSCSVMHEALHNHY ce VL
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNF	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA	EGNVFSCSVMHEALHNHY ce VL Gllllaaqpamaeivltq
YPSDIAVEWESNGQPENNYKT CH3 TQKSLSLSLGRRKRGSGATNF3	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNFS VL	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA	EGNVFSCSVMHEALHNHY ce vl gllllaaqpamaeivltq
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNFS VL SPATLSLSPGERATLSCRASQ	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ NRATGIPARFSGSGSGTD
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNF VL SPATLSLSPGERATLSCRASQ VL	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ NRATGIPARFSGSGSGTD
YPSDIAVEWESNGQPENNYKT CH3 TQKSLSLSLGRRKRGSGATNF VL SPATLSLSPGERATLSCRASQ VL	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS CL	EGNVFSCSVMHEALHNHY ce Gllllaaqpamaeivltq NRATGIPARFSGSGSGTD
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNF3 VL SPATLSLSPGERATLSCRASQ3 VL FTLTISSLEPEDFAVYYCQQS	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS CL SNWPRTFGQGTKVEIKRTVAAPSVF	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ NRATGIPARFSGSGSGTD IFPPSDEQLKSGTASVVC
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNFS VL SPATLSLSPGERATLSCRASQS VL FTLTISSLEPEDFAVYYCQQSS	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS CL SNWPRTFGQGTKVEIKRTVAAPSVF	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ NRATGIPARFSGSGSGTD IFPPSDEQLKSGTASVVC
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNFS VL SPATLSLSPGERATLSCRASQS VL FTLTISSLEPEDFAVYYCQQSS CL	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS CL SNWPRTFGQGTKVEIKRTVAAPSVF	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ NRATGIPARFSGSGSGTD IFPPSDEQLKSGTASVVC
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLGRRKRGSGATNF VL SPATLSLSPGERATLSCRASQ VL FTLTISSLEPEDFAVYYCQQS CL	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS CL SNWPRTFGQGTKVEIKRTVAAPSVF	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ NRATGIPARFSGSGSGTD IFPPSDEQLKSGTASVVC
YPSDIAVEWESNGQPENNYKT CH3 Furin site P2A-site TQKSLSLSLSLGRRKRGSGATNF3 VL SPATLSLSPGERATLSCRASQ3 VL FTLTISSLEPEDFAVYYCQQS3 CL LLNNFYPREAKVQWKVDNALQ3	TPPVLDSDGSFFLYSRLTVDKSRWQ LC signal sequen SLLKQAGDVEENPGPMKYLLPTAAA SVSSYLAWYQQKPGQAPRLLIYDAS CL SNWPRTFGQGTKVEIKRTVAAPSVF	EGNVFSCSVMHEALHNHY ce VL GLLLLAAQPAMAEIVLTQ NRATGIPARFSGSGSGTD IFPPSDEQLKSGTASVVC

QGLSSPVTKSFNRGEC

Figure S2. Amino acid sequence of the nivolumab construct. The coloring of the nivolumab amino acid sequence is analogous to the one of trastuzumab in Supplementary Figure 1. Here, however, a P2A was used instead of an F2A site.



Figure S3. Homogenous expression of the trastuzumab and nivolumab antibodies. a,b) Size-exclusion chromatography / multi-angle light scattering (SEC-MALS) analysis of nivolumab (a) and trastuzumab (b) produced by the furin/2A system (bottom panels) in CHO-S cells, following transient transfection, as compared to the corresponding commercial products (top panels). The design of the constructs is schematically shown in (c) and (d). Both the nivolumab P2A construct and the trastuzumab F2A construct led to a monodisperse protein peak comparable to the correct processing of the signal sequences (shades of green), furin (orange), and 2A sites (black), which are post-translationally cleaved off.



Figure S4. Viral particle visualization and antibody expression levels of cancer cells upon viral vector gene expression. a) Purified first-generation adenoviral virions encoding trastuzumab or nivolumab were analyzed by TEM following negative staining, revealing high purity. b) Signal

from a GFP-reporter, which is transcriptionally coupled to antibody expression via an IRES domain, was used as a flow cytometry read-out for infection of HEK293 cells with different infectious viral particles (ivp)/ cell. Quantification was done in technical replicates (n=3) for each condition, and error bars represent SEM. c) The antibody concentration in the cell supernatant of all infected cell lines was determined by an antigen-capture ELISA, using either HER2 for trastuzumab or PD-1 for nivolumab as immobilized capture antigens. Error bars represent SEM for technical triplicates (n=3). Cell numbers, viral particle numbers, and retargeting adaptor concentration (used for HER2 positive MCF7, SKBR3 and BT474 cells) were kept consistent for all cell lines. While small relative differences in the expression levels of antibodies produced from different cell lines were observed, *i.e.*, nivolumab produced from HEK compared to nivolumab produced from SKBR3 cells, no distinguishable differences were detected in the amount of either trastuzumab or nivolumab produced from the same cell line (except for MCF7).



Figure S5. Comparison of the glycosylation profiles of nivolumab and trastuzumab expressed from either transfected or adenovirus-infected CHO-S cells. Native mass analysis of the CHO-S cell-expressed nivolumab (a) or trastuzumab (b) revealed only minor overall differences between antibodies produced from transfected or infected cells using either naked DNA (transfection) or a first-generation adenoviral vector (infection) for gene delivery. The most apparent difference was visible for trastuzumab expression, where antibody variants with a residual arginine (indicating incomplete carboxypeptidase digestion) could be detected in low amounts from CHO-S cells transfected with the trastuzumab expression plasmid, while none were present from virus-infected CHO-S cells. Conversely, low amounts of half-glycosylated, high-mannose forms of trastuzumab were detected in the virus-infected CHO-S samples, but not in the plasmid-transfected ones.



Figure S6. Glycosylation profiles of nivolumab and trastuzumab expressed in CHO and various human (cancer) cell lines. Trastuzumab (left panel) and nivolumab (right panel)

expressed from different cell lines (indicated in the upper left corner of each row) were analyzed by native mass spectrometry. In the top row for reference, the native MS spectra are shown of the two clinically used commercial products (Herceptin and Opdivo, respectively). The antibodies produced from cancer cells showed, in general, more complex glycosylation patterns than those produced in producer cell lines. Furthermore, SKBR3 and BT474 showed a higher abundance of partially and fully non-glycosylated antibody species. The relative abundances of quantified native mAb-species, as extracted from this data, can be found in Supplementary Table 2.



Figure S7. Illustrative example of the quantitative analysis of tryptic glycopeptides of trastuzumab produced in SKBR3 cells. (a) Averaged MS1 spectrum across the retention time window 15.50-17 min (m/z region of 900-1300) in which the N-glycopeptides with the amino acid sequence of TKPREEQYNSTYR eluted. (b) Chromatographic separation of two isobaric forms of the peptide glycoforms with two fucoses or one sialic acid, respectively. These extracted ion chromatograms of each peptide glycoforms were used for the relative quantification. (c) Typical fragmentation pattern of N-glycopeptides occurring within the first HexNAC of the glycan to the corresponding Asn residue. HCD type of fragmentation gives rise to a group of ions with a mass difference of 17-83-120, which were used for validation of glycopeptide identity and used for their unambiguous annotation. (d) HCD MS/MS spectra of the two glycopeptides from the isobaric precursors m/z 1074.7947 and 1074.4556. Characteristic oxonium ions originating from the fragmentation of the glycan of glycopeptides are detected in the lower m/z region. The middle part of the spectra shows the mass difference

patterns. The composition of the glycan moiety attached to a peptide backbone can be deduced from its sequential fragmentation.



Figure S8. Evaluation of the contribution of incompletely glycosylated antibodies produced by the different cell lines using native or peptide-centric MS data. Quantification of modified and non-modified glycopeptides in the peptide-centric analysis (black bars) validates the observations made by native MS (yellow bars).



Figure S9. Transcript levels of neonatal Fc receptor (FcRn), different Fc γ -receptors (Fc γ R-1A, Fc γ R-1B, Fc γ R-2A, Fc γ R-2B) and HER2 of cancer cell lines used in this study (HT1080, A549, MCF7, SKBR3, BT474). FcRn and Fc γ R transcripts are low compared to HER2 receptor expression in HER2 positive cell lines (MCF7, SKBR3 and BT474 cells) and did not correlate with non-glycosylated trastuzumab expressed from SKBR3 and BT474 cells. RNA transcript numbers were taken from the cancer cell line encyclopedia^{6,7} and are depicted as Transcripts Per Million (TPM).



Figure S10. Transcript levels of ST6Gal1 in the used cancer cell lines. Cells lines showing elevated levels of antibody sialylation in this study also previously showed increased transcript levels of α 2-6 sialic acid transferase ST6Gal1 in other studies. RNA numbers were taken from the cancer cell line encyclopedia^{6,7} and are depicted as Transcripts Per Million (TPM).



Figure S11. Comparison of mannose and galactose profiles of the produced antibodies in the different cell lines. The amount of high mannose, G0, G1, G2, and terminal galactose (TG) was determined for nivolumab (left heatmap) and trastuzumab (right heatmap). No clear differences between the non-cancerous cell lines (lower panel) and the cancerous cell lines (upper panel) were observed.

References

- Palmer D, Ng P. Improved system for helper-dependent adenoviral vector production. Mol Ther. 2003; 8: 846–852. doi: 10.1016/j.ymthe.2003.08.014.
- Schmid M, Ernst P, Honegger A, Suomalainen M, Zimmermann M, Braun L, Stauffer S, Thom C, Dreier B, Eibauer M et al. Adenoviral vector with shield and adapter increases tumor specificity and escapes liver and immune control. Nat Commun. 2018; 9: 450. doi: 10.1038/s41467-017-02707-6
- Bern M, Caval T, Kil Y, Tang W, Becker C, Carlson E, Kletter D, Sen K, Galy N, Hagemans D et al. Parsimonious charge deconvolution for native mass spectrometry. J Proteome Res. 2018; 17: 1216–1226. doi: 10.1021/acs.jproteome.7b00839
- Varki A, Cummings R, Aebi M, Packer N, Seeberger P, Esko J, Stanley P, Hart G, Darvill A, Kinoshita T et al. Symbol nomenclature for graphical representations of glycans. Glycobiology. 2015; 25: 1323–1324. doi: 10.1093/glycob/cwv091.
- Lin Y, Zhu J, Meijer S, Franc V, Heck A. Glycoproteogenomics: A frequent gene polymorphism affects the glycosylation pattern of the human serum fetuin/α-2-HSglycoprotein. Mol Cell Proteomics. 2019; 18: 1479–1490. doi 10.1074/mcp.ra119.001411.
- Petryszak R, Keays M, Tang Y, Fonseca N, Barrera E, Burdett T, Füllgrabe A, Fuentes A, Jupp S, Koskinen S et al. Expression Atlas update - An integrated database of gene and protein expression in humans, animals and plants. Nucleic Acids Res. 2016; 44: D746–D752. doi: 10.1093/nar/gkv1045
- Barretina J, Caponigro G, Venkatesan K, Margolin A, Kim S, Wilson C, Lehár J, Kryukov G, Sonkin D et al., The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483: 603–607. doi: 10.1038/nature11003