Supporting Information for the Research Article, "Examination of Glycan Profiles from IgG-Depleted Human Immunoglobulins Facilitated by Microscale Affinity Chromatography"

Martin Svoboda, Benjamin F. Mann, John A. Goetz, Milos V. Novotny*

Department of Chemistry and National Center for Glycomics and Glycoproteomics,

Indiana University, Bloomington, IN 47405

CONTENTS PAGE

Enzymatic Release of N-glycans, Trypsin Digestion of Proteins, and Spin-column Purification of Glycans and Peptides. Dried proteins were first re-suspended, denatured and reduced using 30 µl of 1M guanidine hydrochloride with 5 mM dithiothreitol (DTT) in 10 mM phosphate buffer (pH 7.4) at 60 °C for 1 h. After cooling to room temperature, iodoacetamide (IAA) was added to a final concentration of 25 mM. Alkylation was allowed to proceed for 45 min, after which time excess IAA was quenched by the addition of an aliquot of 1.5 µl of 0.1 M DTT. The N-glycans were enzymatically released from captured immunoglobulins using PNGase F. The reaction volume was adjusted to 60 µl, 1 µl of PNGaseF solution (diluted 1:10) was added, and the sample was allowed to react for 18 h at 37 °C. The approximate concentration was 0.1 mU PNGase F per mg Igs in the Protein G fractions and 1 mU in the Protein L fractions. Subsequently, 1 µl of trypsin solution (1 mg/mL) was added and the digestion reaction was allowed to proceed for an additional 18 h at 37 °C.

The purification of the released glycans from tryptic peptides was accomplished in two steps. In the first step, the peptides were extracted using a C18 microspin column. The column was washed 3 times by 400 µl of an 85%/15%/0.1% acetonitrile/water/TFA solution (solution B), then again 3 times by 400 µl of a 95%/5%/0.1% water/acetonitrile/TFA solution (solution A). The sample was applied onto the column, the vial was rinsed twice by the A solution and both parts were added to the sample solution. The column was centrifuged and the sample solution was twice re-applied onto the column and, after final centrifugation, twice washed by 100 µl of the A solution. The captured peptides were then eluted by applying twice 150 µl of a 50%/50%/0.1% acetonitrile/water/TFA solution. The peptide fraction was then dried.

In the second step, all other impurities were removed by employing a graphitized-carbon microspin column. The column was prepared in the same way as the C18 column. The flow-

through fraction from the C18 step (which contained the released glycans) was applied and then twice reapplied. Before elution, the sample was washed two times by 100 μl of the A solution. Elution was performed using twice 150 μl of a 35%/65%/0.1% acetonitrile/water/TFA solution. The purified glycans were vacuum-centrifuged until dry.

MS Data Processing. The measured MALDI spectra were processed using Data Explorer (ver. 4.9). Baseline correction and a noise filter were applied, and the resultant spectra were exported into ASCII files. Normalization of glycan intensities was performed using an in-house developed software called PeakCalc.

For the processing of acquired data from the LC-MS/MS measurements, we used a few in-house developed programs which are compiled into one software package as ProteinQuant Suite. A detailed description can be found in the relevant publication.[36] In the first step, the raw spectra from Xcalibur data acquisition software were converted into MASCOT generic files (.MGF) with TurboRAW2MGF for submission to the MASCOT database search engine. The converted files were searched against the Uniprot database for *Homo sapiens (human)*; input criteria for searching were one missed cleavage by trypsin, precursor mass tolerance ±0.02 Da and fragment tolerance ±0.8 Da, peptides with charge +2 and +3, fixed modification of carbamidomethylation of cysteines and variable modification of oxidation of methionines. Results of searching were parsed with the ProtParser program according to defined criteria (in this work, we used the same criteria as in ref. 36) and created files with output data were then combined into a master file, which contained a list of all proteins and peptides identified in the LC-MS/MS experiments. This master file was submitted together with mzXML files, created

from RAW data files using the ReAdW utility, to the program ProteinQuant. The final result is a list of quantified proteins and their identified peptides.

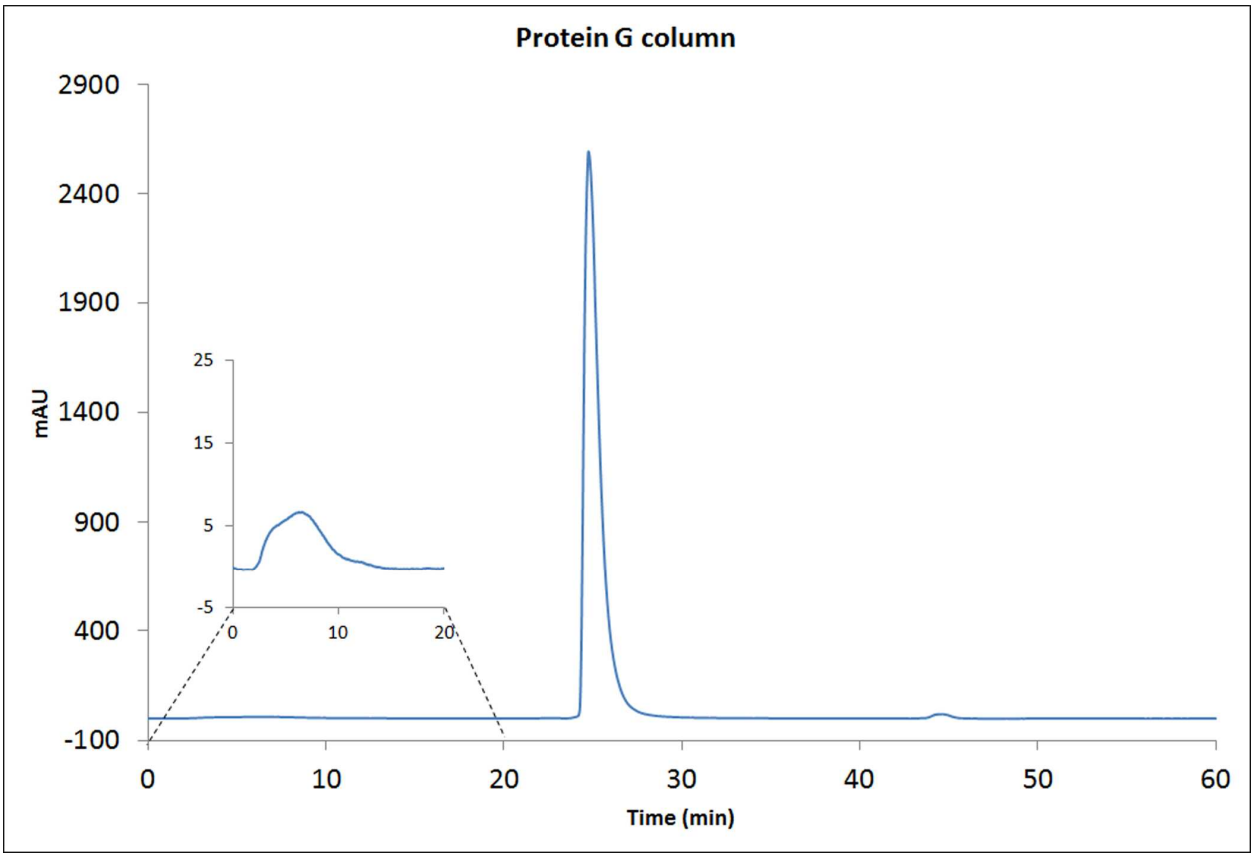

Fig. S-1: Chromatogram of 120 µg of IgG captured with the Protein G microcolumn. At this amount, a flow-through peak appeared for the first time (enlarged area), while it was undetectable at lower IgG amounts.

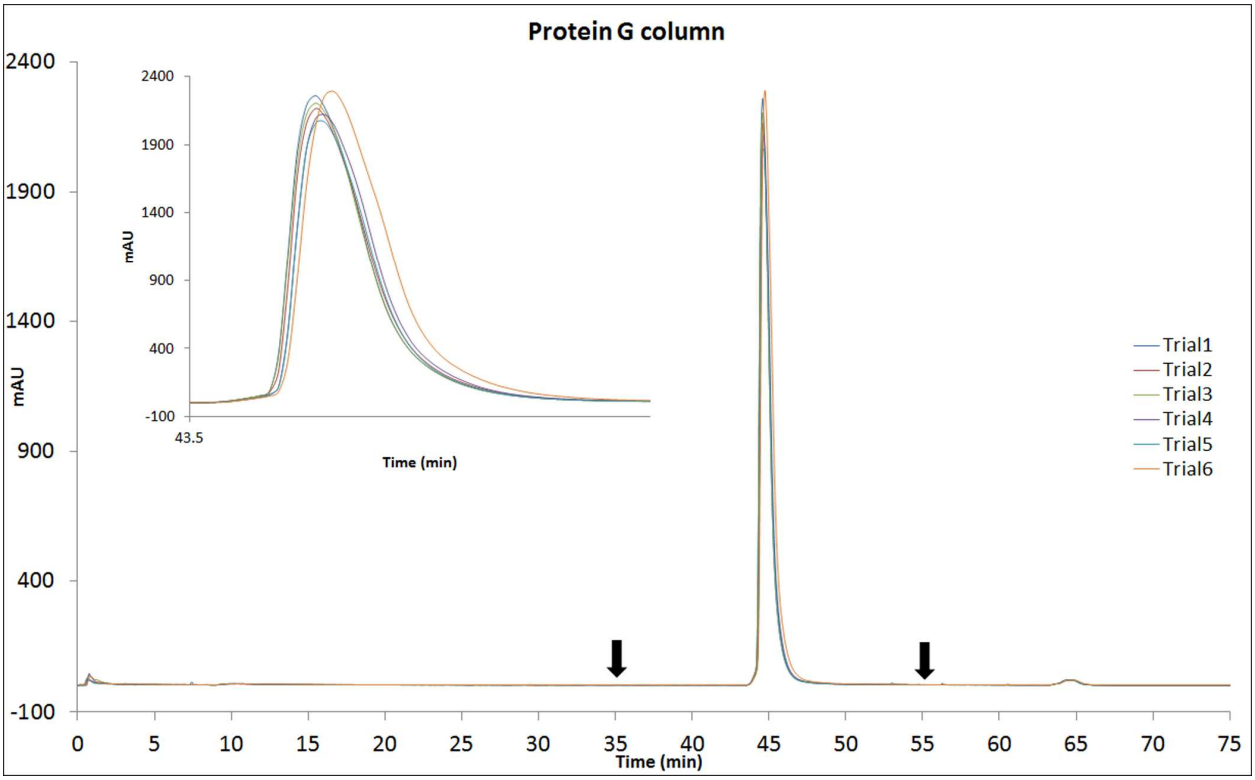

Fig. S-2A: Chromatograms of six trials of pooled female blood sera from the silica-Protein G column showing washing, elution and reequilibration phases. The flow-through peaks for unbound serum proteins have already been eluted from the column during the serial enrichment and are not shown here. Black arrows label where the elution buffer was applied (minute 35) and stopped (minute 55). The inset shows a zoomed view of the time window (43.5 – 48.5 min) surrounding the eluted peaks.

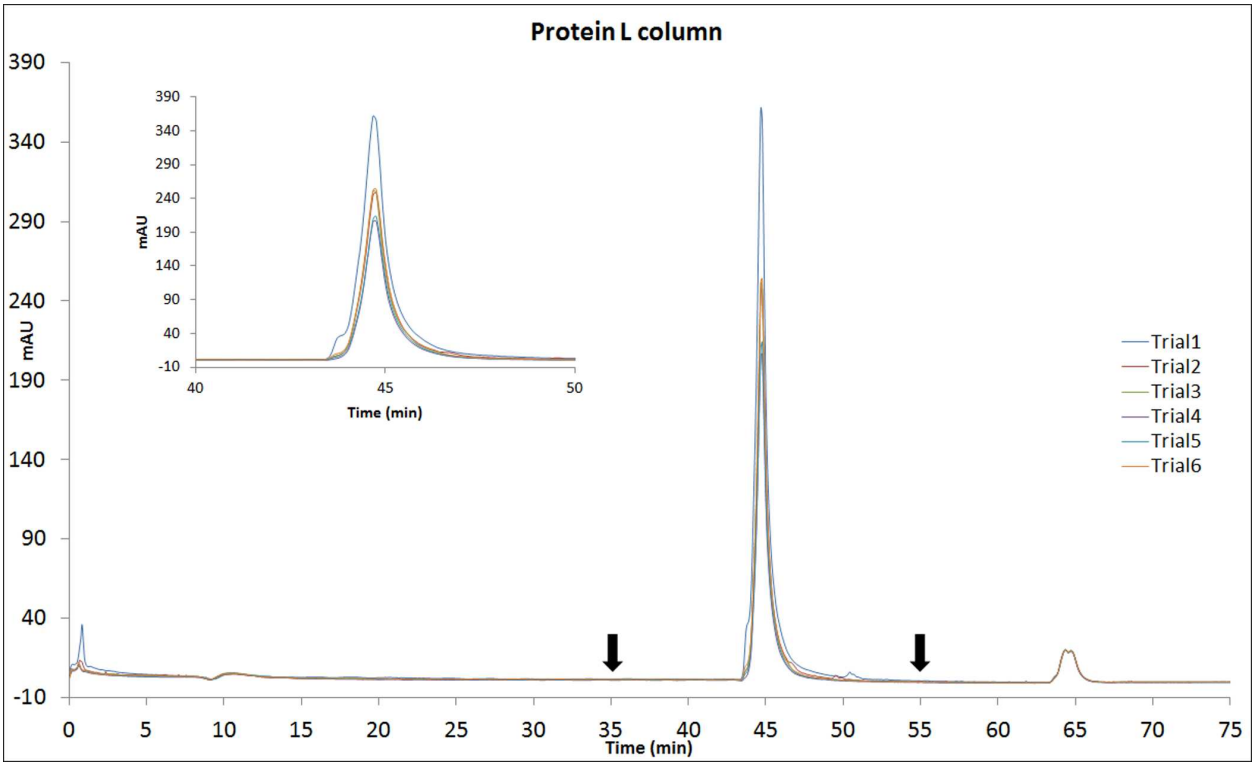

Fig. S-2B: Chromatograms of six trials of pooled female blood sera from the silica-Protein L column showing washing, elution and reequilibration phases. The flow-through peaks for unbound serum proteins have already been eluted from the column during the serial enrichment and are not shown here. Black arrows label where the elution buffer was applied (minute 35) and stopped (minute 55). The inset shows a zoomed view of time window (40 – 50 min) surrounding the eluted peaks. The elution peak area for the six replicates was calculated. After application of Dixion's Q test, we excluded the result for the first trial with 95% confidence. The mean elution peak area for the remaining replicates was 196.8±19.6 mAU*min, and there was no visible trend in increasing peak area with additional experiments, so carryover between experiments was negligible.

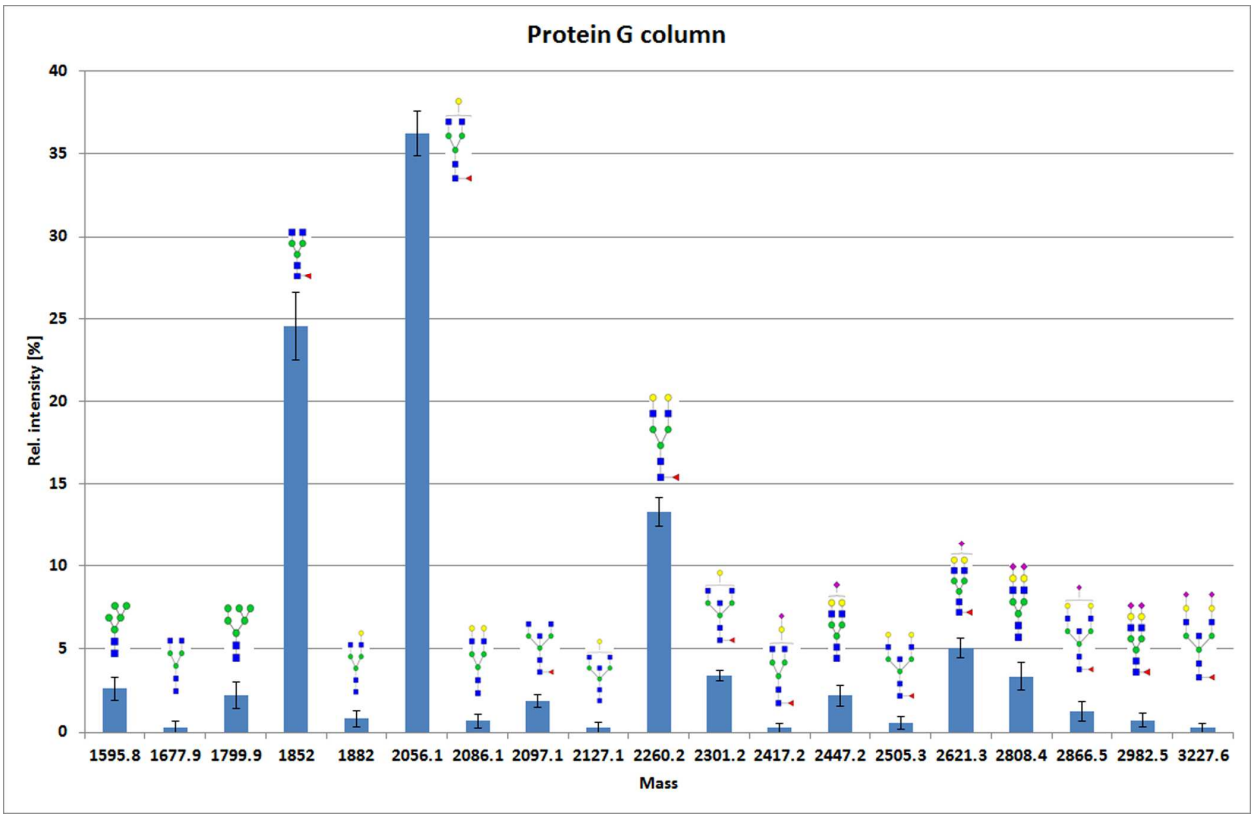

Fig. S-3: Relative abundance of glycan structures expressed on IgG enriched with the Protein G microcolumn using an initial volume of 3 µl of pooled female serum for each analysis. Abundance was evaluated using the in-house developed program PeakCalc. Six technical replicates were performed for the affinity chromatography experiments, while each technical replicate was spotted six times on the MALDI target. The mean relative intensity from the separate enrichment experiments was used to determine the average value and standard deviation for each glycan structure. Monosaccharides are depicted as follows: blue square represents N-acetylglucosamine, green circle is mannose, yellow circle is galactose, red triangle represents fucose, and pink diamond is sialic acid.

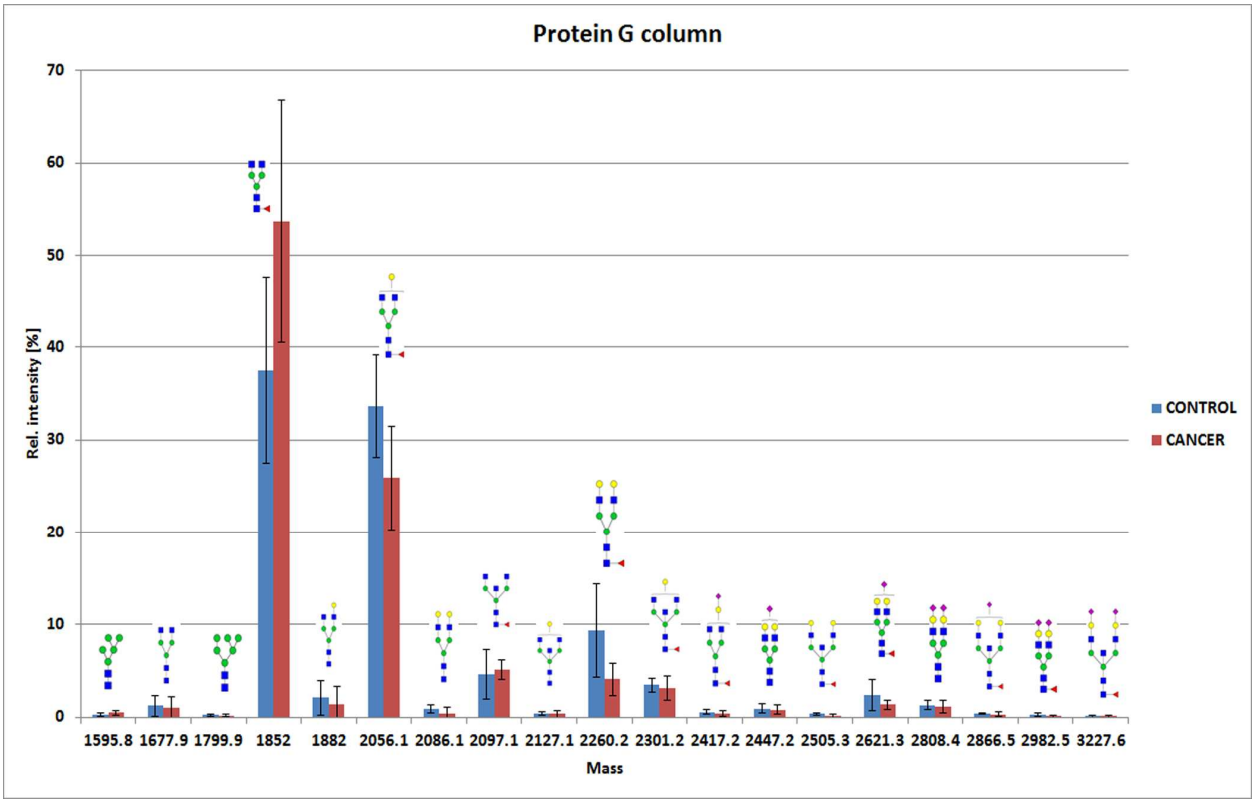

Fig. S-4: A comparison of the relative intensities of the glycan structures from IgG. Abundance was evaluated using the in-house developed program PeakCalc. Six technical replicates were performed for the affinity chromatography experiments, while each technical replicate was spotted six times on the MALDI target. The mean relative intensity from the separate enrichment experiments was used to determine the average value and standard deviation for each glycan structure. Monosaccharides are depicted the same as in Fig. S-3.

Table S-1: Patient Information of Individuals from the Lung Cancer Experiments

Patient ID	**Age**	**Years Smoked**	**Packs per Day**
1 Cancer	65	45	LESS1[a]
2 Cancer	61	45	LESS1[a]
3 Cancer	60	28	ONE_TWO[b]
4 Cancer	61	20	ONE_TWO[b]
5 Cancer	66	35	LESS1[a]
6 Cancer	66	50	OVER_TWO[c]
	63.2		
1 Control	58	10	LESS1[a]
2 Control	59	2	LESS1[a]
3 Control	61	25	LESS1[a]
4 Control	57	4	NAN[d]
5 Control	64	8	NAN[d]
6 Control	62	10	OVER_TWO[c]
	60.2		

[a] *LESS1-Less than one pack per day*

[b] *ONE_TWO: One to two packs per day*

[c] *OVER TWO: More than two packs per day*

[d] *NAN: Not answered*

Table S-2: Proteomic Profile of the Protein G-enriched Fraction of Blood Serum

UniProt accession	Proteins from Protein G column	Protein area (%)	SD	CV (%)
	Immunoglobulins			
P0CG04	Ig lambda chain C region	21.00	9.91	47.2
P01834	Ig kappa chain C region	18.69	6.67	35.7
P01857	Ig gamma-1 chain C region	17.89	7.74	43.3
P01860	Ig gamma-3 chain C region	9.96	1.52	15.2
P01861	Ig gamma-4 chain C region	5.95	1.86	31.3
P01859	Ig gamma-2 chain C region	4.36	1.76	40.5
P01876	Ig alpha-1 chain C region	0.53	0.37	70.2
P01871	Ig mu chain C region	0.22	0.25	112.0
	Ig light chains V region	14.82		
	Ig heavy chains V region	1.36		
	Total	94.78		
	Non-glycoproteins			
P02768	Serum albumin	0.14	0.04	26.6
	Other proteins	2.21		
	Total	2.35		
	N-linked Glycoproteins			
P04004	Vitronectin	2.67	0.31	11.7
	Other proteins	0.21		
	Total	2.87		