Alterations in the glycan profile of mouse transferrin: New insights in collageninduced arthritis.

Montserrat Mancera-Arteu¹, Estela Giménez¹, Jaime Sancho², Victoria Sanz-Nebot¹*

¹Department of Chemical Engineering and Analytical Chemistry, Institute for Research on Nutrition and Food Safety (INSA·UB), University of Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain

²Instituto de Parasitología y Biomedicina "López-Neyra" (IPLBN), CSIC, Granada, Spain.

*Corresponding author: <u>vsanz@ub.edu</u> (V.Sanz-Nebot, PhD)

Tel: (+34) 934021283, Fax: (+34) 934021233

Abstract

purification from mice serum samples Transferrin by immunoaffinity chromatography (IAC) was optimized in order to study the possible modifications occurred in its glycans in collagen-induced arthritis (CIA) samples. SDS-PAGE and nanoLC-MS/MS were used to monitor the IAC purification performance. Afterwards, a relative quantification of mouse transferrin (mTf) glycan isomers using $[^{12}C_6]/[^{13}C_6]$ aniline was used to unequivocally detect alterations in the glycan profile of CIA mice. In addition, multivariate data analysis was applied to identify the most meaningful glycan isomers for the discrimination between control and pathological samples. Partial least squares discriminant analysis (PLS-DA) revealed that five out of fifteen mTf glycan isomers could be potential biomarkers of CIA, most of them corresponding to highly sialylated structures (H6N5S3 2, H6N5S3 3 and H5N4S3 2). Moreover, some of these glycan isomers also seemed to be related with the progression of CIA, especially H6N5S2 and H6N5S3 2, as their overexpression increased with the clinical score of the pathology. Hence, the established methodology provides valuable information to find glycan-based biomarkers of CIA, but also leaves the door open to evaluate, in the future, glycosylation changes of many other inflammatory diseases, in which transferrin has been described to be altered.

Keywords: transferrin / collagen-induced arthritis / glycan isomers / immunoaffinity chromatography / multivariate data analysis / sialylation

Introduction

Changes in serum protein glycosylation are early indicators of cellular alterations in many diseases, including the inflammatory arthropathies (IA), providing a good basis for diagnosis and insights into disease progression ^{1,2}. Altered glycomic profile was reported in several proteins isolated from the sera of IA patients, including IgG, transferrin, haptoglobin and α 1-acid glycoprotein ^{1,3}. Transferrin (Tf) is a negative acute phase protein whose main role is the transport of iron through the blood plasma. Changes in sialylation, fucosylation and glycan branching of Tf have been described in many inflammatory diseases³, such as rheumatoid arthritis (RA). RA is a chronic inflammatory autoimmune disease that affects up to 1% of population worldwide, doubling mortality rate of RA patients compared to healthy individuals. However, up to now, there is no single test to confirm the diagnosis of RA and it is based on the symptoms and the measurement of some variables (rheumatoid factor (RF), C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), among others)⁴. Therefore, with the aim of finding a reliable biomarker, several authors addressed the study of Tf glycosylation in RA analyzing the intact glycoforms by isoelectric focusing (IEF)⁵, or more recently, by capillary electrophoresis (CE) ⁶. However, these methods provided ambiguous results: whereas an increase of the highly sialylated glycoforms (S5 and S6) in RA patients was observed by Feelders et al.⁵, Gudowska et al.⁶ reported a decrease of S3 and S5 sialoforms.

CIA is a commonly studied autoimmune mice model of arthritis that closely resembles human RA in terms of disease course, histological findings and also in its response to commonly used anti-arthritic pharmaceuticals ^{7,8}. A chronic form of CIA is induced in C57BL/6 wild-type (WT) mice by immunization with chicken type II collagen in complete Freund's adjuvant (Col-II/CFA) ⁹. Thus, it is considered an adequate model

to study the efficacy of novel drugs and to evaluate glycosylation changes derived from arthritis ¹⁰. Mouse transferrin (mTf) is a serum glycoprotein of about 80 kDa, which presents only one *N*-glycosylation site with complex type *N*-glycans. Due to the multiple similarities between CIA and RA previously mentioned, alterations in the glycosylation pattern of mTf in presence of CIA could be expected. In a previous work of our research group, the analysis of mouse transferrin (mTf) at the glycopeptide level by capillary liquid chromatography-mass spectrometry was carried out. That study demonstrated mTf glycopeptides are modified in presence of CIA, mainly observing an increase of fucosylation and glycan branching ¹¹.

However, the glycopeptide approach did not provide information about possible altered glycan isomers. Zwitterionic hydrophilic interaction capillary liquid chromatography-mass spectrometry (CapZIC-HILIC-MS) enables an excellent separation of isomeric glycans, and the use of a glycan reductive isotope labelling (GRIL) strategy with $[^{12}C_6]/[^{13}C_6]$ -aniline allows performing a reliable quantification isomer by isomer, identifying unequivocally major and minor variations in the expression of certain glycan isomers 12,13 . Nevertheless, using this glycomic approach, a specific and selective purification strategy is required in order to selectively capture the target glycoprotein and avoid the contribution of glycans coming from other glycoproteins.

In this work, an immunoaffinity chromatography (IAC) column was developed and optimised to properly isolate mTf from serum samples, testing different conditions to remove the non-specifically retained proteins. SDS-PAGE as well as nanoLC-MS/MS analysis of the eluted IAC fractions were used to monitor the performance of mTf purification. Under the optimized IAC conditions, glycan isomers of mTf purified from commercial control mouse serum were analysed by CapZIC-HILIC-MS to establish a reference glycoprofile of this protein. Afterwards, in order to detect modifications in mTf glycosylation in WT mice suffering from CIA (CIA WT samples), relative quantification of mTf glycans was carried out using the GRIL strategy with $[^{12}C_6]/[^{13}C_6]$ -aniline (CIA or non-immunized WT mice versus control mouse serum samples). To our knowledge, this is the first glycomic study focused on the identification of Tf aberrant glycosylation in such inflammatory diseases. Moreover, given the potential of partial least squares discriminant analysis (PLS-DA) for the analysis of human α 1-acid glycoprotein glycan isomers in pancreatic ductal adenocarcinoma (PDAC) samples ¹³, this chemometric tool was also used in the present study to identify which glycan isomers of mTf enable the differentiation between non-immunized and CIA WT mice. These altered glycans in CIA could be useful, in the future, to find novel glycan biomarkers for the diagnosis of RA in humans.

2. Materials and methods

2.1. Chemicals

All chemicals used in the preparation of buffers and solutions were of analytical reagent grade. Acetic acid (HAc, glacial), formic acid (HFor 98-100%), dimethylsulphoxide (DMSO), glycine ($\geq 99.7\%$), glycerol ($\geq 99.5\%$), sodium dodecyl sulfate (SDS, >99.8%) tris(hydroxymethyl)aminomethane (Tris, >99.9%) and acetone were supplied by Merck (Darmstadt, Germany). Sodium phosphate dodecahydrated (Na₃PO₄·12H₂O), sodium cyanoborohydride (NaBH₃CN), [¹²C₆]-aniline, [¹³C₆]-aniline, sodium chloride (NaCl, 99.5%) and sodium bicarbonate (NaHCO₃) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Hydrogen chloride (HCl, 37%) and 2mercaptoethanol (β-ME) were supplied by Panreac (Barcelona, Spain) and "NP-40 alternative" by Calbiochem (Darmstadt, Germany). Sodium azide (NaN₃, >99.5%) was obtained from Fluka (Madrid, Spain). Bromophenol blue, tetramethylethylenediamine (TEMED), acrylamide/bis solution (30%), ammonium persulfate (APS) and Bio-Safe[™] Coomassie stain were supplied by Bio-Rad (Hercules, USA). BenchMark[™] Protein Ladder was provided by Thermo Fisher Scientific (Waltham, USA). Ammonium acetate (NH₄Ac), acetonitrile (ACN) and water LC-MS quality grade, used for CapZIC-HILIC-MS analysis, were obtained from Merck and Fluka, respectively. Mouse apo-transferrin standard (mTf, ≥98%, commercially purified from control sera) was purchased from Sigma-Aldrich. Peptide N-glycosidase F (PNGase F) was obtained from Roche Diagnostics (Basel, Switzerland). CNBr-activated Sepharose 4B was purchased from GE Healthcare (Waukesha, WI, USA) and goat polyclonal antibody against human Tf (immunogen affinity purified) from Abcam (Cambridge, UK).

2.2. Mice serum samples

Control mouse serum (pooled serum from normal mouse population) was purchased from Sigma-Aldrich. C57BL/6 wild-type (WT) mice were purchased from Harlan Ibérica (Barcelona, Spain). For the induction of CIA, 8-12 weeks-old male mice were immunized with Col-II/CFA 6 as previously described ^{9,14}. Four non-immunized WT mice serum samples: WT1, WT2, WT6, WT7, all of them males of 8-10 weeks old, and five WT mice serum samples with CIA: WT1-CIA (clinical score of 5), WT3-CIA (clinical score of 7), WT4-CIA (clinical score of 6), WT6-CIA (clinical score of 8) and WT7-CIA (clinical score of 8), all of them males of 14-16 weeks old, were analyzed. The clinical score was quantified according to a graded scale of 0–3 as follows: 0= no inflammation (normal joint); 1= detectable local swelling and/or erythema; 2= swelling in >1 joint and pronounced inflammation; 3= swelling of the entire paw and/or ankylosis. Each paw was graded, and the scores were summed (with a maximum possible score 12 per mouse). All studies with live animals were approved by the IPBLN and Universidad de Cantabria Institutional Laboratory Animal Care and Use Committees.

2.3. Immunoaffinity chromatography (IAC)

2.3.1 IAC column preparation

A CNBr-sepharose IAC column was prepared following our previous protocol ¹⁵ with some modifications proposed in the literature ¹⁶. Briefly, 5-column volumes of cold activation buffer (1 mM HCl) were added to 0.15 g of cyanogen bromide-sepharose dry resin and incubated on a nutator mixer for 2 h at 4 °C. After swelling the resin (0.5 mL, approximately), it was centrifuged for 5 min and supernatant was decanted. The commercial buffer containing the anti-Tf antibody was exchanged for the coupling buffer (0.1 M NaHCO₃ and 0.5 M NaCl, pH 8.3) using Microcon YM-10 centrifugal filters as

described in ¹⁷. The resultant antibody solution (650 μ L at a concentration of 1.54 mg·mL⁻¹) was added to the resin and incubated overnight on a nutator mixer at 4 °C. Afterwards, resin was washed several times with coupling buffer, centrifuged and supernatant was removed. In order to block unreacted groups, 5-column volumes of quenching buffer (0.1 M Tris–HCl, pH 8.0) were added and incubated on a nutator mixer for 2 h at room temperature. Then, resin was extensively washed to remove uncoupled antibody repeating this cycle for three times: centrifugation and supernatant removal, addition of 10-column volumes of high pH washing buffer (0.1 M Tris–HCl and 0.5 M NaCl, pH 8.0), centrifugation and supernatant removal and addition of 10-column volumes of low pH washing buffer (0.1 M NaAc and 0.5 M NaCl, pH 4.0). Finally, resin was transferred to a 1.5 mL empty plastic column and stored at 4 °C filled with storage buffer (10 mM Tris–HCl and 0.01% (w/v) NaN₃, pH 7.6–7.7).

2.3.2 Albumin/IgG depletion kit

ProteoExtract[™] Albumin/IgG Removal Kit was purchased from Merck to remove serum albumin and IgG from mice serum samples. Depletion procedure was performed according to manufacturer's instructions ¹⁸. Samples were concentrated by Speed Vac until a proper volume (~600-800 µL) to be loaded later in the IAC column.

2.3.3 mTf purification by IAC

The purification procedure previously described in 15 was modified in order to improve mTf isolation, minimizing non-specific retention of other serum proteins. Firstly, the IAC column was conditioned with 8-column volumes of binding buffer (10 mM Tris– HCl and 0.25 M NaCl, pH 7.6). Serum samples (50-100 µL) were diluted 1/8 in binding buffer and passed ten times through the IAC column (albumin/IgG-depleted samples were loaded directly). After washing with 6-column volumes of washing buffer (10 mM Tris– HCl and 0.5 M NaCl, pH 7.6), retained mTf was eluted with 1-column volumes of elution buffer (100 mM glycine-HCl, pH 2.5). The eluate was collected in a tube containing 0.5 M Tris to immediately neutralize it. Finally, excess of low-molecular mass reagents was removed with Microcon YM-30 centrifugal filters as described in ¹⁷.

After IAC, purified mTf samples were quantified by capillary electrophoresis with ultraviolet detection (CE-UV, Agilent Technologies, Waldbronn, Germany) using a bare fused-silica capillary of 60 cm total length (LT) x 75 μ m internal diameter (I.D.) x 360 μ m outer diameter (O.D.) (Polymicro Technologies, Phoenix, AZ, USA). This protein quantification approach was used due to its simplicity, low sample consumption, quickness and possibility of automation for consecutive quantification of many serum samples. New capillaries were activated off-line by flushing (930 mbar) sequentially for 15 min each with 1M NaOH and water. Samples were injected for 10 s at 50 mbar and CE was only used to introduce the protein sample throughout the capillary. Experiments were performed without supplying voltage, only applying 50 mbar of pressure, and measuring absorbance at 214 nm. Calibration curve was performed with mouse apotransferrin standard at concentrations between 50 and 1000 μ g·mL⁻¹. Afterwards, purified mTf samples were evaporated to dryness by Speed Vac and stored at -20 °C until its use.

2.4. SDS-PAGE

SDS-PAGE was performed on a vertical system Mini-PROTEAN® Tetra Cell with a PowerPacTM HC Power Supply (Bio-Rad, Hercules, USA) using in-house 10% SDSpolyacrylamide gels (30% (v/v) Acrylamide/bis solution, 0.375 M Tris-HCl pH 8.8, 10% (v/v) SDS, 10% (v/v) APS and 0.004 % (v/v) TEMED). Fifteen μ L of each IAC purified serum sample or mouse apo-transferrin standard (~10 μ g of protein) were reduced and denatured with 5 μ L of reducing-Laemmli sample buffer (0.25 M Tris-HCl pH 6.8, 4% (m/v) SDS, 20% (v/v) glycerol, 10% (v/v) β-ME and 1% (v/v) bromophenol blue), incubating the mixture in a thermoshaker at 100°C for 5 min. Then, samples were loaded into the gel to perform the protein separation. Ten μ L of protein ladder (BenchMarkTM Protein Ladder) were also loaded in one lane in order to assign the molecular weight to the bands. Gel electrophoresis was performed at 120 V for 2 h at room temperature using a running buffer consisted of 25 mM Tris-base, 250 mM glycine and 0.1% SDS. After SDS-PAGE, gel was fixed in 40% (v/v) ethanol and 10% (v/v) HAc for 30 min and then rinsed in Milli-Q water (3 x 5 min). Gel was incubated with Coomassie blue staining solution at room temperature for 1 h with agitation, and then rinsed in Milli-Q water until a proper degree of staining was achieved.

2.5. nanoLC-MS/MS

IAC purified samples (~5 μ g of mTf) as well as mouse apo-transferrin standard (10 μ g of mTf) were reduced, alkylated and subjected to trypsin digestion in the presence of Rapigest® as described in ¹⁵.

Prepared tryptic digests were diluted in 3% ACN, 1% HFor at a final concentration of 0.35 pmol/ μ L. One μ L of each sample was loaded into a 300 μ m × 5 mm PepMap100, 5 μ m, 100 Å, C18 μ -precolumn (Thermo Scientific) at a flow rate of 15 μ L/min using a Thermo Scientific Dionex Ultimate 3000 chromatographic system (Thermo Scientific). Peptides were separated using a C18 analytical column (Acclaim PepMap® RSLC 75 μ m × 50 cm, nanoViper, C18, 2 μ m, 100Å, Thermo Scientific) with gradient elution at a flow rate of 250 nL·min⁻¹. Eluting solvents were A: 0.1% HFor in water and B: 0.1% HFor in ACN. The following gradient conditions were used: from 3 to 35% B in 60 min, from 35 to 50% B in 5 min, and from 50% to 85% B in 2 min, followed by isocratic elution at 85% B in 5 min and stabilization to initial conditions. LC-MS coupling was performed with an Advion TriVersa NanoMate (Advion) fitted on an Orbitrap Fusion Lumos[™] Tribrid (Thermo Scientific). The mass spectrometer operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the orbitrap with the resolution (defined at 200 m/z) set to 120,000. The lock mass was user-defined at 445.12 m/z in each Orbitrap scan. The top speed (most intense) ions per scan were fragmented by HCD and detected in the Orbitrap. The ion count target value was 400,000 and 50,000 for the survey scan and for the MS/MS scan, respectively. Target ions already selected for MS/MS were dynamically excluded for 30s. Spray voltage in the NanoMate source was set to 1.60 kV. RF Lens were tuned to 30%. Minimal signal required to trigger MS to MS/MS switch was set to 20,000. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for fragmentation.

A database search was performed with Proteome Discoverer software v2.1.0.81 (Thermo) using Sequest HT search engine and SwissProt mouse with contaminants database. Search was run against targeted and decoy database to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in methionine and acetylation in protein N-terminus as dynamic modifications. Peptide mass tolerance was 10 ppm and MS/MS tolerance was 0.02 Da. Peptides with a q-value lower than 0.1 and a FDR < 1% were considered as positive identifications with a high confidence level. The Exponentially Modified Protein Abundance Index (emPAI), based on protein coverage by the peptide matches in a database search result, was used to relatively quantify the abundance of the proteins present in the mixture.

2.6. N-glycan preparation

mTf purified samples were reduced and digested with PNGase F to release *N*-glycans as explained elsewhere ¹⁹. Released *N*-glycans were purified by solid phase extraction (SPE) using Hypercarb cartridges (25 mg, 1 mL volume, Thermo Fisher Scientific) as described previously ¹³. The labeling was carried out by adding 10 μ L of reaction mixture (0.35 M aniline and 1 M NaCNBH₃ in DMSO with 30% HAc) to the dried glycans and incubating the mixture in a thermoshaker for 2 h at 70 °C. mTf glycans obtained from WT mice serum samples (non-immunized and CIA samples) were labelled with [¹²C₆]-aniline, while mTf glycans obtained from commercial control mouse serum were labelled with [¹³C₆]-aniline. After incubation, samples were cooled to room temperature and labelled glycans were precipitated with acetone as described in ²⁰. Subsequently, equimolar mixtures of mTf non-immunized or CIA samples and mTf control mouse serum samples were prepared. Finally, the obtained mixtures were evaporated to dryness by Speed Vac and, dried *N*-glycans were stored at -20 °C until analysis. Centrifugations were performed in a Mikro 220R centrifuge (Hettich Zentrifugen, Tuttlingen, Germany).

2.7. CapZIC-HILIC-MS

1200 Series capillary liquid chromatography system coupled to a 6220 oa-TOF LC/MS mass spectrometer with an orthogonal G1385-44300 interface (Agilent Technologies) were used to perform CapLC-MS experiments. ZIC-HILIC column packed with 3.5 mm particles, 150 x 0.3 mm L_T x ID (SeQuant, Umeå, Sweden) was used for chromatographic separations. Experiments were performed at room temperature with gradient elution at a flow rate of 4 μ L/min and injecting 0.25 μ L of each glycan sample (glycan concentration: ~100-50 pmol/ μ L). Eluting solvents were A: 1 mM NH₄Ac solution and B: acetonitrile. The following gradient conditions were used: solvent B from 90% to 80% (within 5 min) and from 80% to 65% (within 20 min) as linear gradient, followed by cleaning and equilibration steps of B: 65% \rightarrow 50% (within 5 min), 50% \rightarrow 0%

(within 5 min), 0% (over 15 min), 0% \rightarrow 90% (within 5 min) and 90% (over 10 min). Tuning and calibration of the mass spectrometer were carried out in accordance with the manufacturer's instructions. Measurement parameters and optimal operational conditions are detailed in ¹⁹. MassHunter Workstation software (Agilent Technologies) was used for CapLC-MS control, data acquisition and analysis.

2.8. Multivariate data analysis

CapLC-MS data collected for mTf purified from serum samples was processed to obtain the extracted ion chromatograms (EIC) of all glycans. The EIC of each glycan was obtained based on the m/z of the most abundant molecular ions observed for each glycan (usually adducts corresponding to the deprotonated glycan: [M-2H]⁻², [M-3H]⁻³ and [M- $4H^{-4}$). Peak areas of the N-glycan isomers were measured from the EICs and used to calculate its relative area applying a GRIL strategy with $[^{12}C_6]/[^{13}C_6]$ aniline (i.e. area of the glycan isomer in the WT mice sample, non-immunized or CIA, labelled with $[^{12}C_6]$ aniline, divided by the area of the same glycan isomer in the control mouse serum sample, labelled with $[^{13}C_6]$ -aniline). Relative areas of glycan isomers were used to build a matrix for multivariate data analysis. The relative areas of the matrix were autoscaled (mean centered and scaled to unit standard deviation). Principal component analysis (PCA) was performed to explore the data for different classes and detect the presence of outliers ²¹. Partial least squares discriminant analysis (PLS-DA) was applied afterwards to maximize class separation and identify which glycan isomers were the most meaningful to discriminate between classes taking into account the variable importance in the projection (VIP) scores ^{22,23}. Leave-one-out cross validation of the PLS-DA model was performed during calibration ²⁴. In the prediction step, the model was used to classify the validation sample. SOLO (Version 8.7, student edition, Eigenvector Research Inc., Wenatchee, WA, USA) was used for PCA, PLS-DA and VIP calculations. Nomenclature used for glycans correspond to their composition, in terms of number of hexoses (H), N-acetylglucosamines (N), fucoses (F) and sialic acids (S), followed by an index that indicates the isomer number.

3. Results and Discussion

3.1 Purification of mTf by immunoaffinity chromatography

Immunoaffinity chromatography (IAC) was used to isolate Tf from mice serum samples using a cyanogen-bromide Sepharose column with an immobilised antibody against human Tf. In order to evaluate the performance of mTf purification, SDS-PAGE as well as nanoLC-MS/MS analysis of the eluted IAC fractions were carried out. Firstly, the IAC column and the purification procedure reported previously for the analysis of human and mouse Tf glycopeptides were used ^{11,15}. Figure 1A shows the SDS-PAGE results obtained for mouse apo-transferrin standard (lane 1), and for the control mouse serum purified using our initial IAC conditions (lane 2). As can be observed, other serum proteins were non-specifically retained by the IAC column. Table 1 shows the list of proteins identified by nanoLC-MS/MS with emPAI values higher than 10%, considered the most abundant ones, in the eluted fraction, which corresponds to lane 2 of the SDS-PAGE gel in Figure 1A. Most of these proteins present *N*-glycosylation sites, some of them with a high percentage of glycosylation, which would contribute to the total amount of glycans of the sample analysed.

Table 1 Proteins identified by nanoLC-MS/MS with emPAI values higher than 10%, in contro
mouse serum sample purified by immunoaffinity chromatography (IAC), using the initial and the
optimized IAC conditions.

Protein name ^a	Accession number ^b	MW (kDa)	N-Glycosylation sites	emPAI ^c (%)					
Initial IAC conditions									
Serum albumin	P07724	68.6	0	93.8					
Serotransferrin	Q921I1	76.7	1	34.3					
Apolipoprotein A-I	Q00623	30.6	0	24.1					
Vitamin D-binding protein	P21614	53.6	1	23.2					
Pregnancy zone protein	Q61838	165.7	11	21.0					
Alpha-1- antitrypsin 1-5	Q00898	45.9	3	18.0					

Serine protease inhibitor A3K	P07759	46.9	4	16.3			
Alpha-1- antitrypsin 1-3	Q00896	45.8	3	14.5			
Immunoglobulin kappa constant	P01837	11.8	0	13.7			
Alpha-1- antitrypsin 1-4	Q00897	46.0	3	12.9			
Alpha-1- antitrypsin 1-2	P22599	45.9	3	11.6			
Serine protease inhibitor A3M	Q03734	47.0	3	10.1			
Optimized IAC conditions							
Serotransferrin	Q921I1	76.7	1	55.2			
Serum albumin	P07724	68.6	0	21.3			

^aProtein Name according to UniProt or to NCBI.

^bUniProtKB/Swiss-Prot or NCBI Accession Number.

^cExponentially Modified Protein Abundance Index (emPAI): relative quantitation of the proteins in a mixture based on protein coverage by the peptide matches in a database search result.

Since the purpose of this study was focused on the analysis of glycans obtained only from mTf, a more selective purification method was necessary. First, a new IAC column was prepared modifying some steps of the previous protocol reported in ¹⁵, as explained in section 2.3.1. Briefly, the antibody to resin ratio as well as the incubation time were substantially increased and the antibody coupling steps were performed at 4°C as suggested by ¹⁶. However, mTf purification using the new IAC column also resulted in a high non-specific retention, similar to the one showed in Figure 1A. Therefore, the purification procedure was changed increasing the washing and elution volumes trying to favour both the clean-up and the elution of the target protein. As can be observed in Figure 1B, these modifications were enough to significantly improve mTf isolation from control mouse serum by IAC (compare lane 2 of Figures 1A and 1B). Nevertheless, some bands, corresponding to other proteins, were still present in the SDS-PAGE gel in contrast to mouse apo-transferrin standard (lane 1). To avoid this non-specific retention, 0.25 M of NaCl was added to the binding buffer. Lane 3 of the SDS-PAGE gel of Figure 1B shows the results obtained using these conditions. Although the number of bands decreased, an

intense band at around 68 kDa still remained, probably corresponding to albumin. The use of a depletion kit prior to IAC purification was then evaluated in order to eliminate albumin. Although mTf band seemed to be more intense when using a depletion kit (lane 4 of Figure 1B), the significant increase in costs and in time required to perform the purification was not worth it. In addition, albumin band was still present. Finally, we also tested a double immunopurification but this resulted in a very poor mTf detection for both depleted and non-depleted control mouse serum. Given the results obtained, the IAC procedure using a binding buffer containing 0.25 M NaCl, washing with 6-column volumes of washing buffer and eluting with 1-column volume of elution buffer, was established as the optimum (lane 3 of Figure 1B). Under these conditions, only transferrin and albumin were identified by nanoLC-MS/MS with emPAI values higher than 10% (see Table 1). As albumin is not a glycosylated protein, it should not contribute to the glycomic profile of the purified mTf. To corroborate this issue, the bar graph of Supplementary Figure 1 compares the glycomic profile obtained from the purified control mouse serum sample using the initial IAC conditions and the optimised protocol with the new IAC column by CapZIC-HILIC-MS. As can be observed, the profile using the initial procedure was quite different, probably because of the contribution of glycans that came from other glycosylated proteins.

B)

A)



Figure 1: (A) SDS-PAGE results obtained for mouse apo-transferrin standard (lane 1), and commercial control mouse serum purified by the initial IAC conditions (lane 2). (B) SDS-PAGE results obtained for mouse apo-transferrin standard (lane 1) and control mouse serum purified by the new IAC column: using larger washing and elution volumes (lane 2); adding 0.25 M NaCl to the binding buffer (lane 3) and using a depletion kit before IAC (lane 4).

3.2 Analysis of mTf glycan isomers in CIA samples

Once optimized mTf purification from serum samples, with the aim of detecting possible modifications on its glycosylation in mice with collagen-induced arthritis (CIA), a reference glycan profile of mTf was first established by analysing mTf glycans isolated from the control mouse serum. Table 2 lists the glycans and their corresponding isomers detected by CapZIC-HILIC-MS in the control mouse serum sample after IAC purification.

Glycan	Isomer	t _R (min)	Relative Area (%)ª	M_{theo}^{b}	Error (ppm)
H5N4S1	1	24.0	2.41	2024.7216	8.4
H5N4F1S1	1	24.8	1.53	2170.7796	9.5
H5N4S2	1	21.2	9.99	2331.8120	3.1
	2	21.9	45.3	2331.8120	2.6
H5N4F1S2	1	21.8	1.84	2477.8699	5.1
	2	22.5	13.4	2477.8699	4.3
H5N4S3	1	20.8	1.61	2638.9023	6.3
	2	21.6	10.5	2638.9023	5.2
H5N4F1S3	1	21.9	1.21	2784.9602	10.9
H6N5S2	1	23.3	1.19	2696.9441	9.8
H6N5S3	1	21.7	0.49	3004.0345	7.3
	2	22.3	3.45	3004.0345	5.2
	3	22.9	5.90	3004.0345	5.6
H6N5F1S3	1	24.2	0.67	3150.0924	9.7
	2	24.6	0.55	3150.0924	10.3

Table 2.- Glycans detected by CapZIC-HILIC-MS in the control mouse serum sample after IAC purification.

^aRelative area was calculated as the peak area of each glycan divided by the sum of the peak areas of all glycans detected. ^bTheoretical mass of the glycan labelled with [$^{12}C_6$]-aniline.

As can be observed, most glycans correspond to biantennary structures, being H5N4S2 the most abundant one. Some of them were also fucosylated although they were less intense than the non-fucosylated counterpart. These results agreed with those reported in ¹¹ for the analysis of mTf at the glycopeptide level. However, tetrantennary structures were not detected in this work, probably due to their low abundance. On the other hand, all sialic acids present in mTf glycans were N-glycolylneuramic acids

(NeuGc), as reported by ^{11,25}. By way of an example, Figure 2 shows the extracted ion chromatograms (EIC) of the most abundant mTf glycans detected in the control mouse serum sample. Moreover, we compared the glycan profile of this commercially available control mouse serum with respect to a pool of non-immunized WT mice serum samples. As can be seen in supplementary Figure 2, both samples showed the same type of glycans and with similar abundances, which demonstrated this control mouse serum could be reliably used in this study as reference for the analysis of mTf glycans of all non-immunized and CIA mice samples.



Figure 2: Extracted ion chromatograms (EICs) of the most abundant mTf glycan isomers detected by CapZIC-HILIC-MS in the control mouse serum after IAC purification under the optimized conditions.

To unequivocally determine the under or overexpression of mTf glycans in the pathological samples, a glycan relative isotope labeling (GRIL) approach with $[^{12}C_6]/[^{13}C_6]$ -aniline previously established ²⁰ was used. Firstly, the reproducibility of the overall methodology (taking into account the GRIL method but also the sample treatment, including IAC purification) was evaluated. For this purpose, three independent control mouse serum samples (n = 3) were IAC purified and labeled with [$^{12}C_6$]AN and another three with [$^{13}C_6$]AN. After derivatization, equimolar mixtures of mTf-glycan-[$^{12}C_6$]AN and mTf-glycan-[$^{13}C_6$]AN were prepared and analysed by CapZIC-HILIC-MS by triplicate. Experimental ratios (area of the glycan labeled with [$^{12}C_6$]AN divided by the area of the same glycan isomer labeled with [$^{13}C_6$]AN) for all mTf glycans were close to 1 and relative standard deviations $\leq 6\%$, demonstrating the reproducibility of the established methodology. Afterwards, mTf glycans isolated from each non-immunized or CIA WT sample (labeled with [$^{12}C_6$]-aniline) were mixed (1:1) with mTf glycans obtained from the control mouse serum (labeled with [$^{13}C_6$]-aniline), and analysed by CapZIC-HILIC-MS. Four non-immunized and five CIA WT mice serum samples at different clinical scores were analysed. The information about the clinical score, gender and age of each mouse serum sample is detailed in Materials and methods section.

Experimental ratios (area of each glycan isomer in non-immunized or CIA WT samples divided by the area of the same glycan isomer in the control mouse sample) for all mTf glycan isomers detected in the different samples were calculated, and used for data interpretation in the multivariate data analysis approach. First, PCA was used to explore the data for the unsupervised identification of trends and detection of outliers. Four non-immunized (WT1, WT2, WT6 and WT7) and four CIA (WT1-CIA, WT3-CIA, WT6-CIA and WT7-CIA) WT samples were used to build the model. The scores plot for the first two principal components (a total of 77% of variance explained by the sum of PC 1 and PC 2) is shown in Supplementary Figure 3A. Despite the variability between the analysed samples, two sample groups could be observed. The loadings plot (supplementary Figure 3B) revealed that triantennary glycans were more related to CIA WT mice (they appeared in the right upper quadrant, the same region as most of CIA samples in the scores plot). In addition, outlier samples were not detected in the outlier

detection plot (Q Residuals vs. Hotelling T²) of Supplementary Figure 3C. However, PCA did not show the importance of each glycan isomer to differentiate between the two groups. Therefore, a partial least squares discriminant analysis (PLS-DA) model was built in order to improve class separation and identify the glycan isomers that could be potentially used as biomarkers of CIA in mice. The same mice samples than for the PCA model were used as the calibration set classifying them in non-immunized (WT) and CIA (WT-CIA) mice. As can be seen in the scores plot of Figure 3A, two latent variables (LVs) allowed discrimination between the two groups (37 and 34% of the X and Y variances explained, respectively). The loadings plot (Figure 3B) showed again that triantennary glycans, such as H6N5S2, H6N5S3 2 and H6N5S3 3, were more related to CIA. To complete this qualitative information, the VIP scores (Figure 3C) allowed us to quantify the influence of the different glycan isomers on the separation between WT and WT-CIA samples. As can be observed in Figure 3C, five out of fifteen glycan isomers were important (VIP>1) to distinguish between non-immunized and pathological samples. Among these glycan isomers, three of them corresponded to triantennary structures (H6N5S2, H6N5S3 2 and H6N5S3 3) and only one was fucosylated (H5N4F1S1). It is also worth mentioning that most of them were highly sialylated structures (H5N4S3 2, H6N5S3 2 and H6N5S3 3). These results are in accordance with the previously described increase of mTf triantennary glycopeptide glycoforms in CIA samples ¹¹ and also with the reported increase of highly sialylated Tf isoforms in human patients with rheumatoid arthritis (RA)⁶. In addition, when characterizing human alfaacid-glycoprotein (AGP) glycan isomers using exoglycosydase digestion and tandem mass spectrometry 19,26 , we demonstrated that isomers with higher proportion of $\alpha 2-6$ linked sialic acids were more retained in the ZIC-HILIC column. In this study, as most mTf glycan isomers considered biomarker candidates of CIA correspond to the most retained isomers within a glycan, it can be assumed that they present higher proportion of α 2-6 linked sialic acids and hence that this linkage-type could be related to CIA diagnosis.

In order to validate the PLS-DA model, WT4-CIA sample was also analysed and data processed following the same procedure described above. As can be observed in the scores plot of Figure 3A, class prediction was excellent demonstrating the good performance of the model to discriminate between non-immunized and CIA WT mice samples.



Figure 3: A) Scores plot, B) Loadings plot and C) VIP scores of the PLS-DA model applied to the ratios (area of the glycan isomer in the non-immunized or CIA sample vs. area of the same

glycan isomer in the control mouse serum sample) obtained for each mTf glycan isomer in the non-immunized and CIA WT mice samples analyzed. Calibration set: four non-immunized samples (WT1, WT2, WT6 and WT7) and four CIA samples (WT1-CIA, WT3-CIA, WT6-CIA and WT7-CIA). Validation sample: WT4-CIA.

Some of the mice serum samples came from the same mouse before and after the induction of CIA (WT1, WT6 and WT7). Figure 4A and B shows, as an example, the extracted ion chromatograms (EICs) of H5N4S3 and H6N5S3 glycans, respectively, detected in WT1 mouse, before (non-immunized) and after the induction of CIA, and compared to the control mouse serum (both glycans identified by PLS-DA as potential biomarkers of CIA). Having as a reference the peak signal of each isomer in the control mouse serum (EICs in black), both isomers of H5N4S3 as well as isomers 2 and 3 of H6N5S3 were upregulated in the CIA sample with respect to the non-immunized one. By way of an example, the MS spectrum of isomer 2 of H6N5S3 as well as the [M-3H]³⁻ ion spectra of isomers 2 and 3 of the same glycan detected in WT1-CIA sample are shown in Supplementary Figure 4. As can be observed, both chromatographic peaks exhibit the same isotopic envelope and identical accurate masses either in $[^{12}C_6]AN$ and $[^{13}C_6]AN$, corroborating that they correspond to isobaric isomers. Moreover, in order to evaluate if the overexpression of these glycan isomers increase with the clinical score, the three pairs of serum samples, in which the same mouse was analysed before and after the induction of CIA (WT1, WT6 and WT7), were used. In this regard, Figure 4C shows the ratio of the CIA samples (Area CIA / Area control) divided by the ratio of the non-immunized ones (Area non-immunized / Area control) obtained for the most meaningful mTf glycan isomers. As can be observed, all glycan isomers were more overexpressed in WT6 and WT7 (clinical score of 8) than in WT1 (clinical score of 5), with the exception of H5N4S3 2 in WT6. These results suggested that the upregulation of these glycan isomers also seemed to be related with the CIA clinical score in mice, especially H6N5S2 and

H6N5S3_2, and therefore they could be used to study the response to novel treatments as well as to monitor this pathology.

A) H5N4S3



Figure 4: Extracted ion chromatograms (EICs) obtained for (A) H5N4S3 and (B) H6N5S3 glycans in non-immunized (i) and CIA (ii) WT1 mouse samples, both with respect to control mouse serum. EICs of glycans in the non-immunized sample are shown in green, in the CIA

sample in red and in the control mouse serum sample in black. C) Bar graph of the ratio of the CIA samples (Area CIA / Area control) divided by the ratio of the non-immunized ones (Area non-immunized / Area control) obtained in WT1 (score 5), WT6 (score 8) and WT7 (score 8) mice samples.

Concluding remarks

Purification of Tf from mice serum samples was achieved by IAC using a CNBrsepharose column activated with a polyclonal anti-human Tf antibody. The activation and coupling protocol as well as the IAC purification procedure were optimized in order to obtain a selective isolation of Tf from mice serum. Under the optimal conditions, SDS-PAGE and nanoLC-MS/MS results confirmed Tf was the major protein present in the eluted fraction and no other glycosylated proteins would interfere in the subsequent glycan analysis. Once established a reference glycomic profile of mTf, glycan isomers from non-immunized and CIA serum samples were analyzed by CapZIC-HILIC-MS using the GRIL strategy with $[^{12}C_6]/[^{13}C_6]$ -aniline. Multivariate data analysis tools showed that five out of fifteen mTf glycan isomers were important to distinguish CIA from control mice, three of them corresponding to triantennary structures (H6N5S2. H6N5S3 2 and H6N5S3 3), one to a highly sialylated biantennary glycan (H5N4S3 2) and only one fucosylated (H5N4F1S1). In addition, these glycan isomers, especially H6N5S2 and H6N5S3 2, seemed to be related with the progression of CIA, being more overexpressed in samples with higher clinical scores. Although a larger cohort of samples should be analysed in the future to corroborate the biomarker value of these glycans, this study could be considered an starting point to find novel glycan-based biomarkers for CIA diagnosis. Moreover, the presented methodology could be also implemented to identify modifications in Tf glycosylation associated with rheumatoid arthritis or other inflammatory processes.

Associated Content

Supporting Information

Figure S-1. Bar graph of the glycans detected in the control mouse serum using the initial and the optimized IAC conditions. **Figure S-2.** Bar graph of the glycans obtained from a pool of non-immunized WT and from the control mouse serum using the optimized IAC conditions. **Figure S-3.** PCA model applied to the ratios obtained for each mTf glycan isomer in all mice samples analyzed. **Figure S-4.** MS spectra of isomer 2 of H6N5S3 and [M-3H]³⁻ ion spectra of isomers 2 and 3 of the same glycan in the WT1-CIA sample.

Author Information

Corresponding Author

*E-mail: vsanz@ub.edu (V. Sanz-Nebot, PhD). Phone: +34 934021283.

Author Contributions

Designed the study (M.M-A., E.G., J.S. and V.S-N.), collected mice serum samples (J.S.), carried out serum purification (M.M-A), performed GRIL - mass spectrometry and data analysis (M.M-A), wrote the manuscript (M.M-A., E.G. and V.S-N). All authors reviewed the manuscript.

Notes

The authors declare no competing financial interest.

Acknowledgments

This work was supported by a grant from the Spanish Ministry of Economy and Competitiveness (RTI2018-097411-B-I00) and the Cathedra UB Rector Francisco Buscarons Ubeda (Forensic Chemistry and Chemical Engineering). Montserrat ManceraArteu acknowledges the University of Barcelona for an ADR fellowship. The authors thank Dr. Marina Gay, from the Mass spectrometry and Proteomics Unit of the Institute for Research in Biomedicine (IRB Barcelona), for her helpful assistance with data interpretation in the nanoLC-MS/MS analysis.

References

- [1] O. Gornik, G. Lauc, Glycosylation of serum proteins in inflammatory diseases, Dis. Markers. 25 (2008) 267–278. https://doi.org/10.1155/2008/493289.
- [2] C. Reily, T.J. Stewart, M.B. Renfrow, J. Novak, Glycosylation in health and disease, Nat. Rev. Nephrol. 15 (2019) 346–366. https://doi.org/10.1038/s41581-019-0129-4.
- [3] S. Albrecht, L. Unwin, M. Muniyappa, P.M. Rudd, Glycosylation as a marker for inflammatory arthritis, Cancer Biomarkers. 14 (2014) 17–28. https://doi.org/10.3233/CBM-130373.
- [4] V. Majithia, S.A. Geraci, Rheumatoid arthritis: Diagnosis and management, Am. J. Med. 120 (2007) 936–939. https://doi.org/10.1016/j.amjmed.2007.04.005.
- [5] R.A. Feelders, G. Vreugdenhi, G. de Jong, A.J.G. Swaak, H.G. van Eijk, Transferrin microheterogeneity in rheumatoid arthritis, Rheumatol. I. 12 (1992) 195–199. https://doi.org/10.1007/bf00302152.
- [6] M. Gudowska, E. Gruszewska, A. Wrona, E. Gindzienska-Sieskiewicz, I. Domyslawska, K. Lipartowska-Klimuk, B. Cylwik, S. Sierakowski, L. Chrostek, The Profile of Serum Transferrin Isoforms in Rheumatoid Arthritis, J. Clin. Rheumatol. 25 (2019) 159–162. https://doi.org/10.1097/rhu.00000000000808.
- [7] D.D. Brand, K.A. Latham, E.F. Rosloniec, Collagen-induced arthritis, Nat. Protoc. 2 (2007) 1269–1275. https://doi.org/10.1038/nprot.2007.173.
- [8] S.Y. Gwon, K.J. Rhee, H.J. Sung, Gene and protein expression profiles in a mouse model of collagen-induced arthritis, Int. J. Med. Sci. 15 (2018) 77–85. https://doi.org/10.7150/ijms.22345.
- [9] J.J. Inglis, G. Criado, M. Medghalchi, M. Andrews, A. Sandison, M. Feldmann, R.O. Williams, Collagen-induced arthritis in C57BL/6 mice is associated with a robust and sustained T-cell response to type II collagen, Arthritis Res. Ther. 9 (2007) 1–8. https://doi.org/10.1186/ar2319.
- [10] A. Rosal-Vela, S. García-Rodríguez, J. Postigo, M. Iglesias, V. Longobardo, A. Lario, J. Merino, R. Merino, M. Zubiaur, J. Sancho, Distinct serum proteome profiles associated with collagen-induced arthritis and complete Freund's adjuvant-induced inflammation in CD38-/- mice: The discriminative power of protein species or proteoforms, Proteomics. 15 (2015) 3382–3393. https://doi.org/10.1002/pmic.201400536.
- [11] A. Rosal-Vela, A. Barroso, E. Giménez, S. García-Rodríguez, V. Longobardo, J. Postigo, M. Iglesias, A. Lario, J. Merino, R. Merino, M. Zubiaur, V. Sanz-Nebot, J. Sancho, Identification of multiple transferrin species in the spleen and serum from mice with collagen-induced arthritis which may reflect changes in transferrin glycosylation associated with disease activity: The role of CD38, J. Proteomics. 134 (2016) 127–137. https://doi.org/10.1016/j.jprot.2015.11.023.
- [12] E. Giménez, M. Balmaña, J. Figueras, E. Fort, C. de Bolós, V. Sanz-Nebot, R. Peracaula, A. Rizzi, Quantitative analysis of N-glycans from human alfa-acid-glycoprotein using stable isotope labeling and zwitterionic hydrophilic interaction capillary liquid chromatography electrospray mass spectrometry as tool for pancreatic disease diagnosis, Anal. Chim. Acta. 866 (2015) 59–68.

https://doi.org/10.1016/j.aca.2015.02.008.

- [13] M. Mancera-Arteu, E. Giménez, M. Balmaña, S. Barrabés, M. Albiol-Quer, E. Fort, R. Peracaula, V. Sanz-Nebot, Multivariate data analysis for the detection of human alpha-acid glycoprotein aberrant glycosylation in pancreatic ductal adenocarcinoma, J. Proteomics. 195 (2019) 76–87. https://doi.org/10.1016/j.jprot.2019.01.006.
- [14] J. Postigo, M. Iglesias, D. Cerezo-Wallis, A. Rosal-Vela, S. García-Rodríguez, M. Zubiaur, J. Sancho, R. Merino, J. Merino, Mice deficient in CD38 develop an attenuated form of collagen type II-induced arthritis, PLoS One. 7 (2012) 1–9. https://doi.org/10.1371/journal.pone.0033534.
- [15] A. Barroso, E. Giménez, F. Benavente, J. Barbosa, V. Sanz-Nebot, Analysis of human transferrin glycopeptides by capillary electrophoresis and capillary liquid chromatography-mass spectrometry. Application to diagnosis of alcohol dependence, Anal. Chim. Acta. 804 (2013) 167–175. https://doi.org/10.1016/j.aca.2013.09.044.
- [16] J.M. Kavran, D.J. Leahy, Coupling Antibody to Cyanogen Bromide-Activated Sepharose, Methods Enzymol. 541 (2014) 27–34. https://doi.org/10.1016/B978-0-12-420119-4.00003-3.
- [17] E. Giménez, R. Ramos-Hernan, F. Benavente, J. Barbosa, V. Sanz-Nebot, Analysis of recombinant human erythropoietin glycopeptides by capillary electrophoresis electrospray-time of flight-mass spectrometry, Anal. Chim. Acta. 709 (2012) 81–90. https://doi.org/10.1016/j.aca.2011.10.028.
- [18] Product Information ProteoExtract Albumin / IgG Removal Kit, Calbiochem. (2004) 1–11.
- [19] M. Mancera-Arteu, E. Giménez, J. Barbosa, V. Sanz-Nebot, Identification and characterization of isomeric N-glycans of human alfa-acid-glycoprotein by stable isotope labelling and ZIC-HILIC-MS in combination with exoglycosidase digestion, Anal. Chim. Acta. 940 (2016) 92–103. https://doi.org/10.1016/j.aca.2016.07.043.
- [20] E. Giménez, V. Sanz-Nebot, A. Rizzi, Relative quantitation of glycosylation variants by stable isotope labeling of enzymatically released N-glycans using [12C]/[13C] aniline and ZIC-HILIC-ESI-TOF-MS, Anal. Bioanal. Chem. 405 (2013) 7307–7319. https://doi.org/10.1007/s00216-013-7178-5.
- I.T. Joliffe, B.J. Morgna, Principal Component Analysis and exploratory factor analysis, Stat. Methods Med. Res. 1 (1992) 69–95. https://doi.org/10.1177/096228029200100105.
- [22] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: A basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.
- [23] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemom. 17 (2003) 166–173. https://doi.org/10.1002/cem.785.
- [24] G.C. Cawley, N.L.C. Talbot, Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers, Pattern Recognit. 36 (2003) 2585–2592. https://doi.org/10.1016/S0031-3203(03)00136-5.
- [25] B. Coddeville, E. Regoeczi, G. Strecker, Y. Plancke, G. Spik, Structural analysis

of trisialylated biantennary glycans isolated from mouse serum transferrin. Characterization of the sequence Neu5Gc(α 2-3)Gal(β 1-3)[Neu5Gc(α 2-6)]GlcNAc(β 1-2)Man, Biochim. Biophys. Acta. 1475 (2000) 321–328. https://doi.org/10.1016/S0304-4165(00)00083-0.

[26] M. Mancera-Arteu, E. Giménez, J. Barbosa, R. Peracaula, V. Sanz-Nebot, Zwitterionic-hydrophilic interaction capillary liquid chromatography coupled to tandem mass spectrometry for the characterization of human alpha-acidglycoprotein N-glycan isomers, Anal. Chim. Acta. 991 (2017) 76–88. https://doi.org/10.1016/j.aca.2017.07.068.



Figure 2