1H-NMR GLYCOPROTEIN ANALYSIS: AN ADVANCED APPROACH FOR INFLAMMATORY DISEASES DIAGNOSIS

Rocio Fuertes Martín

UNIVERSITAT ROVIRA I VIRGILI
1H-NMR GLYCOPROTEIN ANALYSIS: AN ADVANCED APPROACH FOR INFLAMMATORY DISEASES DIAGNOSIS
Rocio Fuertes Martín

UNIVERSITAT ROVIRA i VIRGILI

BiosferTeslab
Advancing testing

DOCTORAL THESIS
Rocío Fuertes Martín

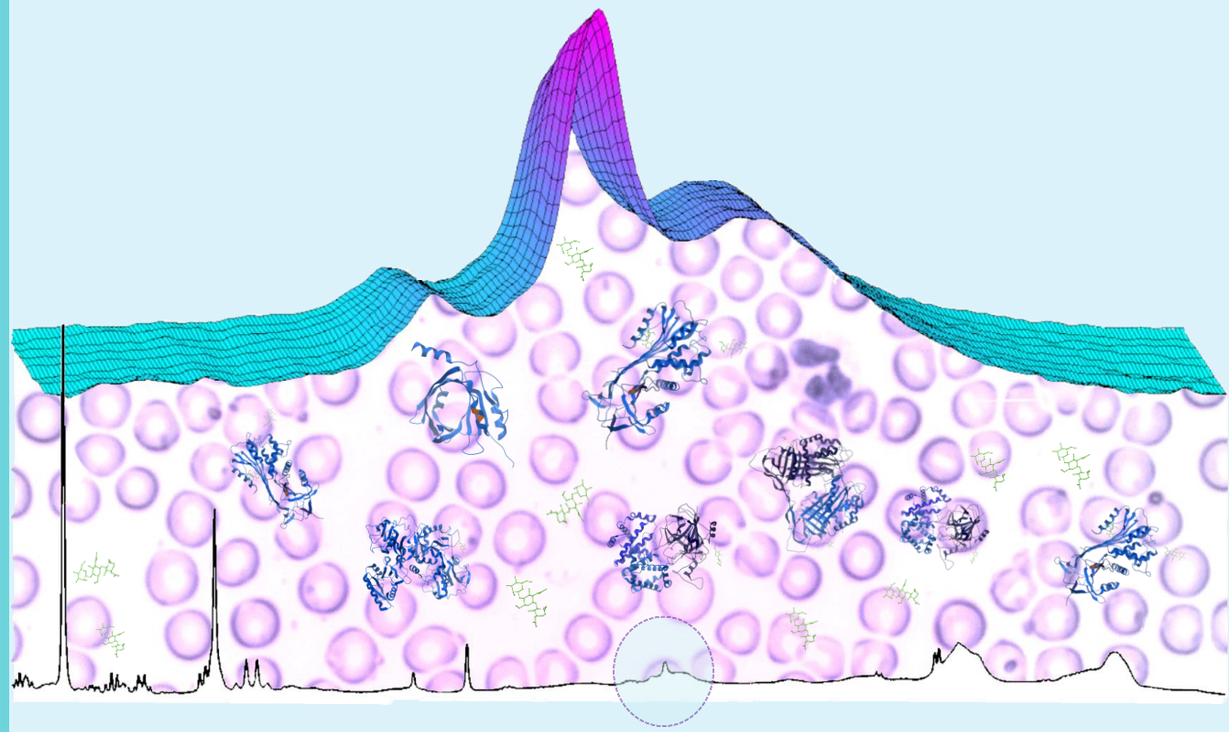
^{1}H-NMR GLYCOPROTEIN ANALYSIS: AN ADVANCED APPROACH FOR INFLAMMATORY DISEASES DIAGNOSIS

2019

UNIVERSITAT ROVIRA i VIRGILI

^{1}H-NMR glycoprotein analysis: An advanced approach for inflammatory diseases diagnosis

Rocío Fuertes Martín



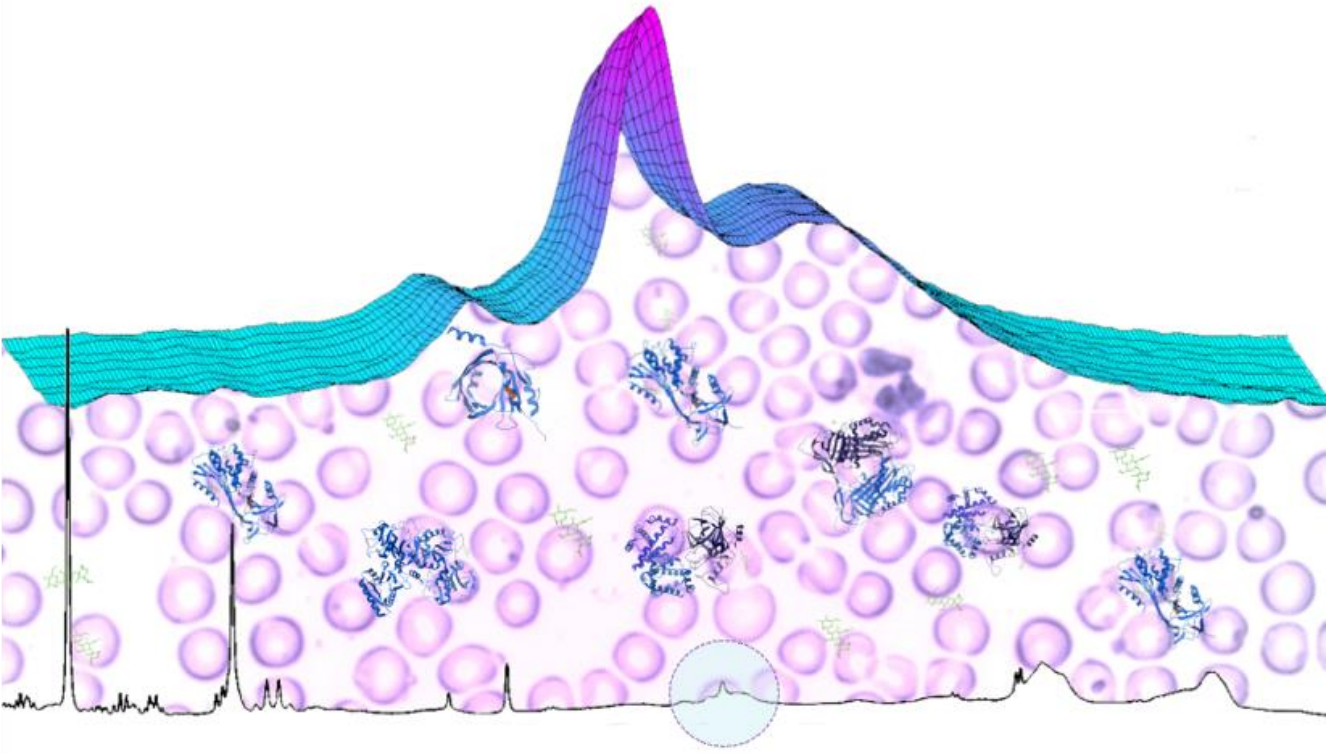
Doctoral Thesis

2019

UNIVERSITAT ROVIRA i VIRGILI

^{1}H-NMR glycoprotein analysis: An advanced approach for inflammatory diseases diagnosis

Rocío Fuertes Martín



Doctoral Thesis

2019

Tesis en versión reducida por acuerdo de confidencialidad

Generalitat de Catalunya Industrial Ph.D. plan (2014DI0029)

Rocío Fuertes Martín

^{1}H-NMR glycoprotein analysis: An Advanced approach for inflammatory diseases diagnosis

DOCTORAL THESIS

Supervised by Dr. Xavier Correig Blanchar

and

Dra. Núria Amigó Grau

Departament d'Enginyeria Electònica, Elèctrica i Automàtica (DEEA)

UNIVERSITAT ROVIRA i VIRGILI

Tarragona

2019

UNIVERSITAT ROVIRA i VIRGILI

Escola Tècnica Superior d'Enginyeria
Departament d'Enginyeria Electrònica, Elèctrica i Automàtica
Av. Països Catalans 26
Campus Sescelades
43007 Tarragona

I STATE that the present study, entitled "**^{1}H-NMR glycoprotein analysis: an Advanced approach for inflammatory diseases diagnosis**", presented by **Rocío Fuertes Martín** for the award of the degree of Doctor, has been carried out under my supervision at the Department of Electronic Engineering of this university.

Tarragona, April 29

Doctoral Thesis Supervisors

Xavier Correig Blanchar **Núria Amigó Grau**

Ilusión, inquietud, curiosidad. Con estos elementos en la maleta, y un viaje rumbo a Reus, comenzó la aventura de esta tesis. Una experiencia cargada de retos, aprendizaje, y momentos de todo tipo que me han hecho crecer personal y profesionalmente.

Toda aventura comienza con una oportunidad. Por ello, en primer lugar, quiero dar las gracias a las personas que me dieron la gran oportunidad de comenzar y que me han guiado a lo largo del camino. Gracias Xavier por haber confiado en mí estos años, por tus consejos, por tu entrega e ilusión en el proyecto, por tus ánimos en los momentos más difíciles. Núria, gracias por haberme transmitido tu pasión, tu alegría y tu fuerza, gracias por haber confiado en mí siempre. De vosotros he aprendido a extraer siempre el lado positivo y a ser optimista. No hubiera sido posible realizar esta tesis sin vuestro constante apoyo y sabiduría, he aprendido muchísimo. Gracias también a Roger por apostar por mí en los comienzos, has sido todo un ejemplo a seguir.

Biosfer Teslab, mi equipo, mis compañeros y amigos con los que he tenido el placer de trabajar día a día, gracias. Miguel, gracias por ayudarme a entrar en el mundo de la RMN, ha sido un placer haber trabajado contigo. Dani, no sé qué hubiera hecho sin tu ayuda con algún que otro script de Matlab, gracias por esto y por endulzarnos tantos desayunos en la oficina. Miriam, mi compañera de batalla, sabes que jamás olvidaré nuestras historias y anécdotas internacionales, casi dan para escribir otra tesis. Sara, gracias por haber traído ese aire fresco de alegría al equipo, por tu optimismo y tus ganas de hacer, que tanto

contagias. Enrique, gracias por tu apoyo y por compartir tus anécdotas, eres todo un aventurero. Gracias a todos por vuestros ánimos.

Gracias a los compañeros y amigos del Yanes Lab; Jordi, Miriam, Sandra, Joan, no olvidaré nuestras comidas italianas de los viernes. Especialmente a Sandra y Miriam, gracias siempre por vuestro apoyo y gran amistad. A los que me guiaron en mis primeros pasos en el COS, gracias. Miguel gracias por tu ayuda con el IconNMR. Gracias Mariona por tus consejos siempre.

Agradezco también a los que dejaron el camino sembrado con su conocimiento y que tuve la suerte de conocer; Rubén, Josep, Daniel. Y al resto de gente del departamento Noelia, Serena, Nico, Jesús, Pere, Alex, Lluc, Carla, Sonia, gracias por los momentos compartidos.

Gracias a todos los grandes profesionales y grandes personas que han confiado en mí, en el proyecto y con los que he tenido el placer de colaborar. Gracias Elisabeth Vilella, Lourdes Martorell y Gerard Muntané del IISPV por el gran interés en el proyecto desde casi el comienzo y por las reuniones tan amenas en el Institut Pere Mata.

Gracias también a María Insenser y Samuel Moncayo, realmente ha sido un placer poder colaborar con vosotros.

Gracias al grupo de la URLA, que ha sido esencial para el desarrollo de la tesis. Gracias por ser siempre tan cercanos. Dr. Luis Masana, gracias por haber confiado en mí, por tus grandes y sabios consejos. Joan Carles Vallvé, por escucharme siempre y por enseñarme, he aprendido tanto contigo que te estoy profundamente

agradecida. Daiana, Josefa Girona, Josep Ribalta, gracias por vuestros ánimos. Merche gracias por los buenos ratos que hemos pasado en las comidas de chicas y por los demás momentos compartidos.

No fue fácil dejar atrás mi ciudad, Granada, mis amigos, mi casa, mi familia para venir a desarrollar esta tesis. Ni han sido fáciles tantas despedidas en estos años, una nunca se acostumbra a despedirse. Por ello quiero dar las gracias a todas las personas que me han hecho sentir en Reus como en casa. Gracias a toda la gran familia del "English corner", en especial gracias a Ana, Regina, Mag, Adrià, Borja, Juanjo y Zaira. A los demás miembros de nuestro pequeño proyecto musical; María, Iván, Miguel y Noe; gracias por la buena música de los martes y por vuestros ánimos en mi última etapa. A mis amigos de siempre, que, aunque lejos, habéis estado ahí; Mireia, Laura, Azahara, Zaira, Gema, Fran, Alberto, Nuri, Laura, Inés, Laor. Gracias por vuestro apoyo.

Quiero dar las gracias en especial a la persona que ha estado a mi lado cada día durante estos dos últimos años, sobrellevando mis batallas y celebrando mis glorias. Gracias Óscar, por tu apoyo personal, por entender mis decisiones, por compartir los buenos momentos y por sacrificar otros.

Por último, gracias a mi familia. A mi abuela, mis tíos y primos que me han dado ánimos. Pero en especial a mis padres y hermano. No hubiera sobrevivido a algunos momentos de crisis si no es por ellos, el pilar fundamental de mi vida. Gracias Javi, mi hermano, por estar siempre ahí, por tus consejos, por ser motivo de muchas alegrías, por ser mi apoyo incondicional. Gracias papá y mamá por haberme

inculcado vuestros valores, por animarme a llevar a cabo cada proyecto, por inspirarme, por enseñarme a vivir día a día. Sois mis maestros, no hay suficientes palabras para daros las gracias por todo.

Gracias a todos por haber formado parte de esta aventura,

Rocío.

"The most beautiful experience we can have is the mysterious. It is the fundamental emotion that stands at the cradle of true art and true science"

Albert Einstein

ABSTRACT

Inflammation is a complex biological response of an organism to a stimulus that can be caused by various external factors such as pathogenic microorganisms, or endogenous factors such as immune processes, genetic alterations or cell changes. In many cases the inflammatory response is maintained over time causing a state of chronic inflammation, present in many pathological states including metabolic diseases such as obesity and diabetes, immune diseases such as rheumatoid arthritis (RA) or systemic lupus erythematosus (SLE), polycystic ovary syndrome (PCOS) and even cancer, among others.

During the inflammatory response there is a variation in the concentrations of certain proteins present in plasma called acute-phase proteins. Most acute-phase proteins present in human serum or plasma are glycoproteins. Protein glycosylation is an enzyme-mediated post-translational process consisting of the addition of glycans to the protein chain through O-glucosidic or N-glucosidic bonds. This process occurs mostly in hepatocytes in response to proinflammatory cytokines such as interleukin 6 (IL-6), which induces synthesis and secretion of positive acute-phase proteins such as α1-acid glycoprotein, haptoglobin, and α1-antitrypsin, among others. During the acute phase response not only the concentration of these proteins increases, but also the structure of the glycan is modified by increasing the branches and residues of monosaccharides that form it, such as the residues of N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc) or N-acetylneuraminic acid or sialic acid (Neu5Ac), among others.

There is a growing interest in the search for new biomarkers that can more accurately monitor inflammation in a more robust way than those currently used in clinical trials such as C-reactive protein (CRP), of known intra-individual variability. In recent years different analytical approaches have appeared to characterize circulating glycoproteins in serum or plasma in order to determine more accurate information about their involvement in the development of diseases. One of the analytical techniques that is emerging recently in the clinical scene is nuclear magnetic resonance (NMR). NMR is a high-performance technology, with minimal sample manipulation, highly reproducible and versatile for cost-effectively quantifying metabolites.

The objective of the thesis presented here has been to develop a method of quantification of glycoproteins in serum and plasma samples through ^{1}H-NMR. This method has allowed to obtain a profile of glycoproteins including the characterization of new variables not described until now. Apart from those already existing in the scientific community, such as GlycA or GlycB, our new parameters include GlycF, and the H/W ratios of GlycA and GlycB, which will be discussed throughout the thesis. The methodology developed has been used in three different studies. In the first, the profiling of glycoproteins in a cohort with rheumatoid arthritis (RA) versus a control group is studied. The results showed a significant increase in the area of GlycA and in the H/W ratios with respect to the control group. In addition, adding the parameters of 1H-NMR to those used in the traditional clinic (fibrinogen, CRP and erythrocyte sedimentation rate or ESR) improved the classification of patients with respect to the DAS28 disease activity index.

In the second work, glycoprotein profiling was studied in a cohort of obese and non-obese women with polycystic ovary syndrome (PCOS) and compared with control men and women. The results showed an increase in the areas of GlycA, GlycB and GlycF and also the H/W ratios in obese. Patients with PCOS also had higher H/W ratios compared to controls.

Finally, in the third work, the glycoprotein profile was characterized in patients with and without diabetes and/or atherogenic dyslipemia (AD) versus a control group. All glycoprotein parameters were significantly higher in the AD and diabetes groups than in the control group. In addition, as in the first work, by adding the new glycoprotein variables to the traditionally inflammatory marker CRP, the AUC increased considerably in the classification models between; a) the control group and the rest (0.68 to 0.84), b) patients with and without dyslipidaemia (0.54 to 0.86), and c) patients with and without diabetes (0.55 to 0.75).

The glycoprotein test using ^{1}H-RMN, which characterizes states of inflammation, is highly sophisticated and innovative from the methodological point of view. In addition, it represents an encouraging new paradigm for its future incorporation into clinical diagnosis.

RESUM

La inflamació és una resposta biològica complexa d'un organisme davant d'un estímul que pot ser provocat per diversos factors externs com microorganismes patògens, factors endògens com són ara processos immunes, alteracions genètiques o canvis cel·lulars. En molts casos la resposta inflamatòria es manté al llarg del temps provocant un estat d'inflamació crònica, present en molts estats patològics. Aquests inclouen malalties metabòliques com la obesitat, la síndrome de l'ovari poliquístic (SOP) i la diabetis, malalties immunitàries com l'artritis reumatoide (AR) o lupus eritematós sistèmic (LES), i fins i tot, el càncer.

Durant la resposta inflamatòria es produeix una variació en les concentracions de certes proteïnes presents en el plasma anomenades proteïnes de fase aguda. La majoria de les proteïnes de fase aguda presents en sèrum o plasma humà són glicoproteïnes. La glicosilació de les proteïnes és un procés post-traduccional mediat per enzims que consisteix en l'addició de glicans a la cadena proteica a través d'enllaços O-glicosídics o N-glicosídics. Aquest procés s'esdevé majoritàriament en els hepatòcits en resposta a citocines proinflamatòries com la interleuquina 6 (IL-6), que indueix la síntesi i la secreció de proteïnes de fase aguda positives com són α1-acid glycoprotein, haptoglobin, and α1-antitrypsin, entre d'altres. Durant la resposta de fase aguda no només augmenta la concentració d'aquestes proteïnes, sinó que també es modifica l'estructura del glicà augmentant les ramificacions i els residus de monosacàrids que el formen com són la N-acetilglucosamina (GlcNAc), la N-acetilgalactosamina

(GalNAc), l'àcid N-acetilneuramínic o àcid siàlic (Neu5Ac), entre d'altres.

Hi ha un interès creixent en la recerca de nous biomarcadors que puguin monitoritzar la inflamació d'una manera més robusta que els actualment utilitzats en clínica, com la proteïna C reactiva (PCR), la qual presenta una elevada variabilitat intraindividual. Durant els últims anys han aparegut diferents aproximacions analítiques per caracteritzar les glicoproteïnes circulants en sèrum o plasma de manera que es pugui determinar informació més exacta sobre la seva implicació en el desenvolupament de malalties. Una de les tècniques analítiques que està emergint recentment a l'escena clínica és la ressonància magnètica nuclear (RMN). La RMN es caracteritza per ser una tecnologia d'alt rendiment, amb una mínima manipulació de la mostra, altament reproduïble i versàtil, que permet quantificar metabòlits de forma eficient.

L'objectiu de la tesi aquí presentada ha estat desenvolupar un mètode de quantificació de glicoproteïnes en mostres de sèrum i plasma a través de ^{1}H-RMN. Aquest mètode ha permès obtenir un perfilat de glicoproteïnes incloent la caracterització de noves variables no descrites fins ara. A part de les ja existents a la comunitat científica, com GlycA o GlycB, els nostres nous paràmetres inclouen GlycF, i els ràtios H / W de GlycA i GlycB, que es discutiran al llarg de la tesi. La metodologia desenvolupada ha estat emprada en tres estudis diferents. En el primer, s'estudia el perfilat de glicoproteïnes en una cohort amb artritis reumatoide (AR) en relació a un grup control. Els resultats van presentar un augment significatiu en l'àrea de GlycA i en els ràtios H /

W respecte al grup control. A més, afegint els paràmetres de ^{1}H-RMN als utilitzats tradicionalment en la pràctica clínica (fibrinogen, PCR i taxa de sedimentació eritrocítica o ESR) va millorar la classificació dels pacients pel que fa a l'índex d'activitat de la malaltia DAS28.

En el segon treball, es va estudiar el perfilat de glicoproteïnes en una cohort de dones obeses i no obeses amb síndrome d'ovari poliquístic (SOP) i aquest es va comparar amb el d'homes i dones controls. Els resultats van mostrar un augment de les àrees de GlycA, GlycB i GlycF i també de les ràtios H / W en obesos. Les pacients amb SOP a més a més van presentar majors ràtios H / W en comparació amb els controls.

Finalment, en el tercer treball es va caracteritzar el perfil de glicoproteïnes en pacients amb i sense diabetis i / o dislipèmia aterogènica (DA) davant d'un grup control. Tots els paràmetres de glicoproteïnes van ser significativament més alts en els grups amb DA i diabetis respecte al grup control. Afegint les noves variables de glicoproteïnes al marcador tradicionalment utilitzat de la proteïna C reactiva (PCR) de la inflamació, les AUC van augmentar considerablement en els següents models de classificació a)el grup control i la resta (0.68-0.84), b) els pacients amb i sense dislipidèmia (0.54 de 0.86) i c) els pacients amb i sense diabetis (0.55-0.75).

El test de glicoproteïnes mitjançant ^{1}H-RMN, que caracteritza estats d'inflamació, és des del punt de vista metodològic altament sofisticat i innovador. Addicionalment, representa un nou paradigma encoratjador per a la seva futura incorporació al diagnòstic clínic.

RESUMEN

La inflamación es una respuesta biológica compleja de un organismo frente a un estímulo que puede ser causado por varios factores externos como microorganismos patógenos, o factores endógenos como procesos inmunes, alteraciones genéticas o cambios celulares. En muchos casos la respuesta inflamatoria se mantiene a lo largo del tiempo provocando un estado de inflamación crónica, presente en muchos estados patológicos incluyendo enfermedades metabólicas como son obesidad y diabetes, enfermedades inmunitarias como artritis reumatoide (AR) o lupus eritematoso sistémico (LES), síndrome de ovario poliquístico (SOP) e incluso cáncer, entre otras.

Durante la respuesta inflamatoria se produce una variación en las concentraciones de ciertas proteínas presentes en plasma llamadas proteínas de fase aguda. La mayoría de las proteínas de fase aguda presentes en suero o plasma humano son glicoproteínas. La glicosilación de las proteínas es un proceso post-traduccional mediado por enzimas que consiste en la adición de glicanos a la cadena proteica a través de enlaces O-glucosídicos o N-glucosídicos. Este proceso ocurre mayoritariamente en los hepatocitos en respuesta a citocinas proinflamatorias como la interleuquina 6 (IL-6), que induce la síntesis y secreción de proteínas de fase aguda positivas como son α1-acid glycoprotein, haptoglobin, and α1-antitrypsin, entre otras. Durante la respuesta de fase aguda no solo aumenta la concentración de estas proteínas, sino que también se modifica la estructura del glicano aumentando las ramificaciones y los residuos de monosacáridos que lo forman como son los residuos de N-acetilglucosamina (GlcNAc), N-

acetilgalactosamina (GalNAc) o ácido N-acetilneuramínico o ácido siálico (Neu5Ac), entre otros.

Existe un interés creciente en la búsqueda de nuevos biomarcadores que puedan monitorizar de una forma más exacta la inflamación de una manera más robusta que los actualmente utilizados en clínica como la proteína C reactiva (PCR), de conocida variabilidad intraindividual. Durante los últimos años han aparecido diferentes aproximaciones analíticas para caracterizar las glicoproteínas circulantes en suero o plasma de manera que se pueda determinar información más exacta sobre su implicación en el desarrollo de enfermedades. Una de las técnicas analíticas que está emergiendo recientemente en la escena clínica es la resonancia magnética nuclear (RMN). La RMN se caracteriza por ser una tecnología de alto rendimiento, con una mínima manipulación de la muestra, altamente reproducible y versátil para cuantificar metabolitos de forma rentable.

El objetivo de la tesis aquí presentada ha sido desarrollar un método de cuantificación de glicoproteínas en muestras de suero y plasma a través de ^{1}H-NMR. Este método ha permitido obtener un perfilado de glicoproteínas incluyendo la caracterización de nuevas variables no descritas hasta ahora. A parte de las ya existentes en la comunidad científica, como GlycA o GlycB, nuestros nuevos parámetros incluyen GlycF, y las ratios H/W de GlycA y GlycB, que se discutirán a lo largo de la tesis. La metodología desarrollada ha sido empleada en tres estudios diferentes. En el primero, se estudia el perfilado de glicoproteínas en una cohorte con artritis reumatoide (AR) frente a un grupo control. Los resultados presentaron un aumento

significativo en el área de GlycA y en las ratios H/W con respecto al grupo control. Además, añadiendo los parámetros de ^{1}H-NMR a los utilizados en la clínica tradicional (fibrinógeno, PCR y tasa de sedimentación eritrocítica o ESR) mejoró la clasificación de los pacientes con respecto al índice de actividad de la enfermedad DAS28.

En el segundo trabajo, se estudió el perfilado de glicoproteínas en una cohorte de mujeres obesas y no obesas con síndrome de ovario poliquístico (SOP) y se comparó con hombres y mujeres control. Los resultados mostraron un aumento de las áreas de GlycA, GlycB y GlycF y también de los ratios H/W en obesos. Las pacientes con SOP además presentaron mayores ratios H/W en comparación con los controles.

Por último, en el tercer trabajo se caracterizó el perfil de glicoproteínas en pacientes con y sin diabetes y/o dislipemia aterogénica (DA) frente a un grupo control. Todos los parámetros de glicoproteínas fueron significativamente mayores en los grupos con DA y diabetes con respecto al grupo control. Además, de nuevo, al añadir las nuevas variables de glicoproteínas al marcador tradicionalmente utilizado de la proteína C reactiva (PCR) de la inflamación, las AUC aumentaron considerablemente en los modelos de clasificación entre; a) el grupo control y el resto (0.68 a 0.84), b) los pacientes con y sin dislipidemia (0.54 a 0.86) y c) los pacientes con y sin diabetes (0.55 a 0.75).

El test de glicoproteínas mediante ^{1}H-RMN, que caracteriza estados de inflamación, es desde el punto de vista metodológico altamente sofisticado e innovador. Además, representa un nuevo

paradigma alentador para su futura incorporación al diagnóstico clínico.

LIST OF ABBREVIATIONS

^{1}H-NMR	Proton Nuclear Magnetic Resonance
α1-AT	α-1-antitrypsin
a1-GP or **AGP**	α-1 glycoprotein
AACT	α-1-antichymotrypsin
ACs	Multicenter AIDS Cohort Study
ACPA	Anti-Citrullinated Peptide Antibodies
AD	Alzheimer disease
ADM:	Adalimumab
AGNES	ANgiography and GEnes Study
AGP	alpha-1-acid glycoprotein
AIM-HIGH	Atherothrombosis Intervention in Metabolic Syndrome with Low HDL/High Triglycerides and Impact on Global Health Outcomes
ARS	Acute radiation sequelae
ART	Antiretroviral therapy
AUC	Area under the curve
BMI	Body mass index
BMT	bone marrow transplantation
CAC	Coronary artery calcium
CAD	Coronary artery disease
CATHGEN	CATHeterization GENetics
CCL2	Chemokine ligand 2
CD	Crohn's disease
CHC	Chronic hepatitis C
CHF	Congestive heart failure
ChrIRD	Chronic inflammatory-related severe hospitalization and death
CKD	Chronic Kidney Disease
COPD	Chronic obstructive pulmonary disease
CPMG	Carr-Purcell-Meiboom-Gill
CRC	Colorectal cancer
CRP	C reactive protein
CTLS	Constrained Total-line shape

CVD	Cardiovascular disease
CVE	Cardiovascular event
DM	Diabetes Mellitus
DILGOM07	DIetary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome 2007 study
EBC	Early breast cancer
EC	Esophageal cancer
ECLIA	Electrochemiluminescence immunoassay
eGFR	Estimated glomerular filtration rate
ELSA-Brasil	The Brazilian Longitudinal Study of Adult Health
ELISAs	Enzyme-linked immunosorbent assays
ERN	Extended-release niacin
ESR	Erythrocyte sedimentation rate
FD	Functional Dyspepsia
FINRISK	Large Finnish population survey on risk factors on chronic, noncommunicable diseases
GA	Genetic Algorithms
GlcNAc	N-acetylglucosamine
GalNAc	N-acetylgalactosamine
GFR	Glomerular filtration rate
HAT	Human African Trypanosomiasis
HBP	High blood pressure
HCC	Hepatocellular carcinoma
HCV-infection	Hepatitis C virus infection
HDL	High density lipoprotein
HIV-1	Human immunodeficiency type 1
HNSCC	Head and neck squamous cell carcinoma
HOMAir	Insulin resistance
Hp	Haptoglobin
Hs-CRP	High-sensitive CRP
IBD	Inflammamtory bowel disease
ICH	Intracerebral hemorrhage
IFG	Impaired fasting glycemia
IFX	Infliximab

IgX	Immunoglobulin X
IRAS	Insulin Resistance Atherosclerosis Study
IS	Ischemic stroke
J-Res	J-resolved spectroscopy
JUPITER	Justification for the Use of Statins in Primary Prevention: An Intervention Trial Evaluating Rosuvastatin trial
KD	Kawasaki disease
LBC	Late stage breast cancer
LCAT	Cholesterol acyltransferase
LDL	Low-density lipoprotein
LED	Longitudinal eddy current delay
Lp-PLA$_2$	Lipoprotein-associated phospholipase A_2
LPS	Lipopolysaccharide
LS7	Life's simple 7 score
LV	Latent Variable
MBC	Metastatic breast cancer
MCI	Mild cognitive impairment
MESA	Multi-Ethnic Study of Atherosclerosis
MetS	Metabolic syndrome
METSIM	Metabolic Syndrome in Men study.
MG	Monoclonal gammopathy
MI	Myocardial infarction
MRS	Magnetic Resonance Sepctroscopy
NAC	N-acetyl glycoprotein (NAC1: GlycA/ NAC2: GlycB)
NAG	N-Acetyl glucosamine
NAGs	N-acetyl glycoproteins
NANA	N-acetyl neuraminic acid
NOESY	One-dimensional ^{1}H Nuclear Overhauser Effect Spectroscopy
NSCLC	Non-small cell lung cancer
PCOS	Polycystic ovary syndrome
PLS-DA	Partial Least Squares-Discriminant Analysis

PREVEND	Prevention of Renal and Vascular End-Stage Disease
PUFA	Polyunsaturated fatty acids
RA	Reumathoid artritis
RIA	Radioimmunoassays
ROC-curve	Receiver Operating Characteristic-curve
SCD	Sickle cell disease
SCD14	Soluble CD14
SCD163	Soluble CD163
SI	Insulin sensitivity index
SLE	Systemic lupus erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SU.VI.MAX	Supplémentation en Vitamines et Minéraux Antioxydants
T2DM	Type 2 Diabetes Mellitus
TA	Takayasu arteritis
TG	Tryglycerides
TF	Transferrin
UC	Ulcerative colitis
UST	Ustekinumab
UTUC	Upper tract urothelial carcinomas
VDM	Vedolizumab
VLDL	Very low-density lipoprotein
WBC	White blood cell count
WHS	Women's Health Study

LIST OF PUBLICATIONS

Fuertes-Martín R, Núria Amigó Grau, Joan Carles Vallvé, Xavier Correig Blanchar**.** *Human serum/plasma glycoprotein analysis by ^{1}H-NMR, an emerging method of inflammatory assessment.* [Submitted].

Fuertes-Martín R, Núria Amigó Grau, Ana Malo, Núria Plana, Daiana Ibarretxe, Josefa Girona, Xavier Correig Blanchar, Lluís Masana. *Glycoprotein Profile measured by ^{1}H-Nuclear Magnetic Resonance in patients with Diabetes. A new robust method to assess inflammation.* [Submitted]

Fuertes-Martín R, Samuel Moncayo, Maria Insenser, M. Ángeles Martínez-García, Núria Amigó Grau, Xavier Correig Blanchar and Héctor F. Escobar-Morreale. Influence of sexual steroids and obesity on glycoprotein profiles: new markers of low-grade chronic inflammation. [Submitted]

Fuertes-Martín, R.; Taverner, D.; Vallvé, J.-C.; Paredes, S.; Masana, L.; Correig Blanchar, X.; Amigó Grau, N. *Characterization of ^{1} H NMR Plasma Glycoproteins as a New Strategy To Identify Inflammatory Patterns in Rheumatoid Arthritis*. J. Proteome Res. 2018, 17 (11), 3730–3739. DOI: 10.1021/acs.jproteome.8b00411

Miranda, J.; Simões, R. V.; Paules, C.; Cañueto, D.; Pardo-Cea, M. A.; García-Martín, M. L.; Crovetto, F.; **Fuertes-Martin, R**.; Domenech, M.; Gómez-Roig, M. D.; et al. *Metabolic Profiling and Targeted Lipidomics Reveals a Disturbed Lipid Profile in Mothers and Fetuses with Intrauterine Growth Restriction*. Sci. Rep. 2018, 8 (1), 13614. DOI: 10.1038/s41598-018-31832-5

Barrilero, R.; Ramírez, N.; Vallvé, J. C.; Taverner, D.; **Fuertes, R**.; Amigó, N.; Correig, X. *Unravelling and Quantifying the "NMR-Invisible" Metabolites Interacting with Human Serum Albumin by Binding Competition and T2 Relaxation-Based Decomposition Analysis*. J. Proteome Res. 2017, *16* (5), 1847–1856. DOI: 10.1021/acs.jproteome.6b00814

LIST OF CONGRESSES

G. Llauradó, **R. Fuertes,** N. Amigó, A. Cano, L. Albert, I. Mazarico, S. Fernández-Veledo, J. Vendrell, X. Correig, J.-M. González-Clemente. *Type 1 diabetes: the association between ^{1}H-NMR plasma glycoproteins and subclinical atherosclerosis*. 55th Annual Meeting of the European Association for the Study of Diabetes (EASD). Barcelona, Spain 10th- 20th September 2019. Poster presentation.

Rocío Fuertes Martín, Núria Amigó Grau, Joan Carles Vallvé, Laura Brugnara, Anna Novials, Lluís Masana, Xavier Correig Blanchar. *Characterization of Glycoprotein Profiles of type-1 and type-2 Diabetes Mellitus and Atherogenic Dyslipidaemia patients by ^{1}H-Nuclear Magnetic Resonance Spectroscopy.* 86th European Atherosclerosis Society (EAS) Congress, Lisbon 5th-8th May 2018. Poster presentation.

Rocío Fuertes. *Caracterización metabolómica de pacientes con artritis reumatoide (AR) a través de resonancia magnética nuclear de proton (^{1}H-RMN).* II Seminary IISPV, 6th July 2017. Oral communication.

Rocío Fuertes Martín, Núria Amigó Grau, Joan Carles Vallvé, Silvia Paredes, Dèlia Taverner, Lluis Masana, Xavier Correig Blanchar. *Caracterizacion del perfil inflamatorio en pacientes con artritis reumatoide (AR) frente a población sana mediante ^{1}H-RMN*. XXX National congress of Spanish society of arteriosclerosis (SEA), Cádiz 31th May-2th Juny 2017. Oral comunication.

Rocío Fuertes Martín, Núria Amigó Grau, Joan Carles Vallvé, Silvia Paredes, Dèlia Taverner, Lluis Masana, Xavier Correig Blanchar. *Characterization of glycoprotein and lipoprotein profiles of rheumatoid arthritis (RA) patients by ^{1}H-nuclear magnetic resonance spectroscopy (^{1}H-NMR).* 85th European Atherosclerosis Society (EAS) Congress, Praga (Czech Republic), 23th-26th April 2017. Poster presentation.

Rocío Fuertes. *BIOSFER TESLAB SL, from Research to Clinical Practice. Characterization of Rheumatoid Arthritis (RA) patients by 1H-Nuclear Magnetic Resonance Spectroscopy (^{1}H-NMR).* Retreat Científic Institut d'investigació Sanitària Pere Virgili (IISPV), Tortosa 16th-17th November 2016. Oral Communication.

Rocío Fuertes Martín, Núria Amigó Grau, Joan Carles Vallvé, Silvia Paredes, Dèlia Taverner, Lluis Masana, Xavier Correig Blanchar. *Charaterization of rheumathoid arthritis (RA) patiens by ^{1}H-Nuclear Magnetic Resonance spectroscopy (^{1}H-NMR).* CSC-Copenhagen 2016, 18th April – 6th May. Oral communication.

Núria Amigó, **Rocío Fuertes**, Carmen Cabré, Maria Vinaixa, Marta Romeu, Mónica Muñoz, Montse Giralt, Jordi Soler, Josep Aguilera, Maria Teresa Compte, Xavier Correig, Alberto Martinez-Vea. *Pro-aterogenic lipoprotein profile associated with white matter lesions in chronic kidney disease patients: a ^{1}H-NMR metabolomic approach.* 84th European Atherosclerosis Society (EAS) Congress, Innsbruck (Austria), 29th May-1th Juny 2016. Science at a Glance presentation - awarded as best poster award.

Núria Amigó, **Rocío Fuertes**, Carmen Cabré, Maria Vinaixa, Marta Romeu, Mónica Muñoz, Montse Giralt, Jordi Soler, Josep Aguilera, Maria Teresa Compte, Xavier Correig, Alberto Martinez-Vea. *Metabolomic approach to white matter lesions in chronic hemodialysis patients identifies a novel metabolic profiles associated with these lesions*. 53rd European Renal Association- European Dialysis and Transplant Association (ERA-EDTA) Congress, Viena (Austria) 21th-24th May 2016. Poster presentation.

N. Amigó, **R. Fuertes**, C. Cabré, M. Vinaixa, M. Romeu, M. Muñoz, M. Giralt, J. Soler, J. Aguilera, MT. Compte, X. Correig, A. Martínez-Vea. *Aproximación metabolómica con ^{1}H-RMN para la caracterización de pacientes con enfermedad renal crónica en hemodiálisis que presentan lesiones en la sustancia blanca cerebral.* Spanish Society of Nephrology (SEN). Oviedo, 8th- 1th October 2016. Poster presentation.

Rocío Fuertes Martín, Núria Amigó Grau, Joan Carles Vallvé, Silvia Paredes, Dèlia Taverner, Lluis Masana, Xavier Correig Blanchar. Caracterización metabolómica de pacientes con artritis reumatoide (ar) a través de ^{1}H-Resonancia Magnética Nuclear (^{1}H –RMN). XXIX National congress of Spanish society of arteriosclerosis (SEA), Granada 18th – 20th May 2016. Oral communication.

TABLE OF CONTENTS

CHAPTER 1

1. Introduction

1.1. Background

Inflammation is a complex biological response of an organism to a damaging stimulus which can be caused by various factors such as a pathogenic microorganism or cell failure. The process takes place in all pathological conditions[1] and involves a cellular signalling pathway to remove that stimulus and to initiate the process of healing. The acute phase of inflammation is an innate immune reaction in which hemodynamic changes, altered vascular permeability, and leukocyte modifications are produced in order to repair the damage. When the inflammation lasts for a long time (weeks or months), it is called chronic inflammation and is the basis of many diseases such as autoimmune diseases, obesity and even cancer, among others.[2] Both acute and chronic inflammation are very complex and our knowledge of how to diagnose, understand, and control the inflammatory response remains limited. Every inflammatory process is accompanied by numerous changes at the site of inflammation, as well as numerous systemic physiological and biochemical changes,[3] but in the last decades increasing attention has been paid to changes in glycosylation.[4,5]

Glycosylation is the most common post-translational modification affecting the functions of proteins, such as protein stability, enzymatic activity and protein-protein interactions. Glycoproteins play key roles in cell communication, signalling and cell adhesion.[6] Most of the human serum or plasma proteins, except for albumin, are glycoproteins belonging to the family of acute-phase proteins, which rise or fall in response to acute and chronic

inflammatory stimuli.[1,3] An example of a mediator of the acute-phase response is interleukin-6 (IL-6), which induces hepatic synthesis and secretion of positive acute-phase serum proteins such as α1-acid glycoprotein, haptoglobin, and α1-antitrypsin and reduces levels of negative acute-phase proteins such as transferrin.[3,7] The increase on the protein complex carbohydrate levels in the plasma of people suffering from different disease states is described from years ago.[8] Individual variation in glycosylation is potentially important for personalized disease risk, disease course and response to therapy. It is known that protein glycosylation can differ between persons, but is remarkably stable per individual.[1] It is only when the homeostasis of a person changes, by lifestyle or pathological conditions, that the glycosylation will change notably.[9] Evidence in support of this hypothesis is accumulating, but further studies are needed to enable understanding of the role of changes in protein glycosylation in disease.[10] Thus, it is important to be able to profile the amount and types of the expressed glycoproteins.

Determination of the concentration change of such inflammatory markers may allow screening of individuals and detect disease at early stages as well as facilitate the prognosis prediction and monitoring of treatment response. Besides, body fluids provide an attractive specimen source because body fluids are generally readily accessible and available in reasonable quantities for clinical analysis. It is therefore apparent that a general method for the quantitative analysis of the proteins contained in body fluids in health and disease would be of great diagnostic and clinical importance.

A key problem with the proteomic analysis of serum and many other body fluids is the peculiar protein composition. The protein composition is dominated by a few proteins that are extraordinarily abundant, with albumin alone representing 50% of the total plasma proteins. Due to the abundance of these major proteins, the large number of protein species of lower abundance are obscured or inaccessible by traditional proteomics analysis methods such as two-dimensional electrophoresis (2DE).[11] Thus, there exists a need for methods of high throughput and quantitative analysis of blood, serum or plasma glycoproteins and glycoprotein profiling for diagnostic purposes.

In recent years, Nuclear Magnetic Resonance (NMR) has been used in many areas of science as an analytical technique of high performance. The role of NMR as a robust and reproducible tool to analyse serum lipoproteins for cardiovascular risk prediction[12,13] makes it especially attractive for glycoprotein analysis. In fact, a new marker, obtained by NMR, called GlycA has emerged in the clinical research, reflecting the concentration of at least five circulating glycoprotein concentrations: predominantly alpha-1 antitrypsin (AAT), alpha-1-acid glycoprotein (AGP), haptoglobin (HP), transferrin (TF), and alpha-1-antichymotrypsin (AACT).[14] In addition, it has been suggested in several studies that this marker is a better indicator of the detection, prognosis and therapeutic monitoring of tissue lesion marked by systemic inflammatory processes, as well as cardiovascular risk and type 2 diabetes, than the classic parameter currently used for chronic inflammation such as the highly sensitive C-reactive protein (hsCRP).[15–17] The heterogeneous composition of

GlycA represents a challenge for further research towards investigating and developing molecular intervention strategies.[14,18]

The goal of implementing these biomarkers as standard routine in clinical tests could result in an improvement in the diagnosis and prognosis of some diseases and, therefore, a decrease in the morbidity and mortality.

1.2. Fundamentals of Glycoproteins

A glycoprotein is a molecular entity composed of a polypeptide chain and one or more glycans. A glycan, in turn, is a complex oligosaccharide composed of 10–15 monosaccharide residues. Can be covalently attached to proteins to make glycoproteins, or lipids to make glycolipids.[10] Glycans are assembled from monosaccharide residues through a carefully regulated enzyme-directed process of glycosylation. At the cellular level, the protein component of all glycoproteins is synthesized in the endoplasmic reticulum (ER). They then make their way through multiple stacks of the Golgi apparatus, finally being distributed to various destinations from the trans-Golgi network. This is where the addition of glycans to the polypeptide chain of the protein occurs by complex dynamic interactions between hundreds of enzymes such as glycosyltransferase reactions, enzymes that transfer activated forms of monosaccharides from nucleotide sugars and lipid-linked sugar intermediates to acceptors including proteins, lipids, and growing glycan chains.[19] Some of these activated forms of monosaccharides are mannose, fucose, galactose, N-

Acetylglucosamine (GlcNAc), N-Acetylgalactosamine (GalNAc), N-acetylneuraminic acid (Neu5Ac) or sialic acid, among others.[20] Basically, the monosaccharides or oligosaccharide chains are attached to protein through either N-glycosidic or O-glycosidic bonds. **Figure C1.1** shows the structures of N- and O-glycans and an example of a glycated protein.

N-Glycosidic bonds (N-glycans) are formed between the amide group of an asparagine (Asn) chain, and the anomeric carbon of N-acetylglucosamine (GlcNAc). The enzyme catalyzing formation of this bond recognizes Asn residues in the protein having the sequence Asn-X-Thr (threonine), or Asn-X-Ser (serine), where X represents any amino acid except proline.[21] All eukaryotic N-glycans share a common core sequence: two GlcNAc residues and three mannose (Man) residues. From this core they can branch in three ways: (1) oligomannose, in which only Man residues extend the core; (2) complex, in which "antennae" initiated by GlcNAc extend the core; and (3) hybrid, in which one arm is composed of Man and the others arms are composed of other monosaccharides such as galactose (Gal), glucose (Glc), xylose (Xyl) or Fucose (Fuc).

O-linked glycans are attached to the oxygen of serine (Ser), threonine (Thr), tyrosine, hydroxylysine or hydroxyproline side chains. Normally these glycoproteins are initiated by GalNAc attached to the hydroxyl of Ser or Thr residues. The sugars found in O-GalNAc glycans include GalNAc, galactose, GlcNAc, fucose, and Neu5Ac, whereas Man, Glc, or Xyl residues are not represented.[22]

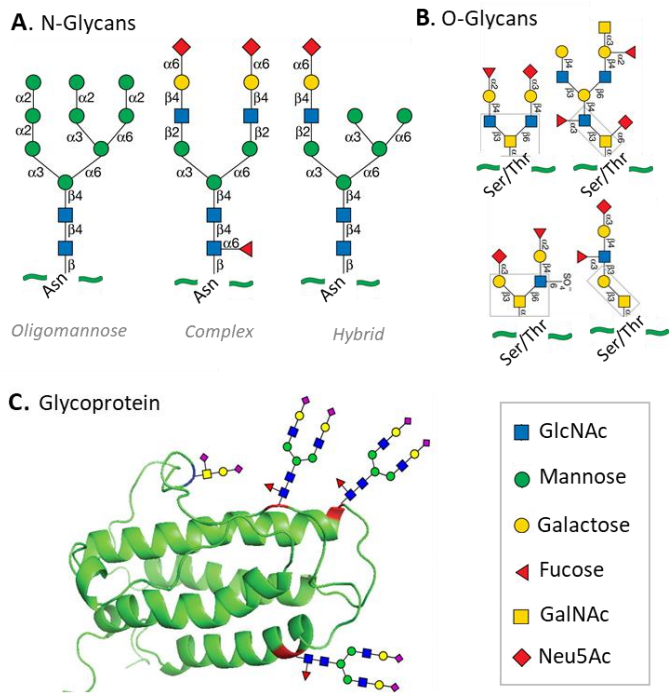


Figure C1.1. A, N-glycans structures; B, O-Glycans structures; C, Glycoprotein example. Adapted from *Essentials of Glycobiology*, 3erd edition [5] and from John A. Brailsford and Samuel J. Danishefsky.[23]

N-glycans are the major and best studied class of carbohydrate protein modifications. However, it should be noted that both types of glycans participate in numerous molecular processes, including protein folding, cell-adhesion, molecular trafficking, signal transduction, modulation of receptor activity and others.[6] They play a major role in all fundamental functions of the multicellular organism, including the immune system.[24]

Because glycans are involved in many biological processes, it is not surprising that alterations in both composition and concentration

can be recognized as causes of an increasing number of diseases, some of which are described in this thesis.

Different techniques have been used to measured individual glycoproteins such as enzyme linked immunosorbent assays (ELISAs), electrochemiluminescence immunoassay (ECLIA), luminex based assays, radioimmunoassays (RIA) and nephelometric assays, among others,[25] that quantify the amount of protein present in biological samples but most of these techniques measure a specific individual glycoprotein. While quantifying protein levels remains the mainstay for measurement of inflammatory glycoprotein levels, measuring the glycan portion of inflammatory proteins is becoming increasingly useful for diagnostic purpose. This can be accomplished using a high throughput technology such as mass spectroscopy (MS) or nuclear magnetic resonance spectroscopy (NMR).

1.3. Nuclear magnetic resonance spectroscopy

NMR spectroscopy is one of the most common techniques in metabolomic studies along with mass spectroscopy (MS). However, the high reproducibility and the non-destructive and non-invasive characteristics of NMR make it potentially attractive for translational medicine. MS techniques are also becoming more common place in clinical laboratories[25] but the effective analysis of protein-derived circulating glycans is still difficult to accomplish due to the high complexity that is caused by variations in glycan linkage and branching. Furthermore, quantification by MS is not always reliable

and for some samples there can be overlap from isobaric glycans (discrete isomeric glycan structures that possess the same mass).[26] MS data can be very complex and interpretation requires expertise. Furthermore, in MS, matrix effects, ionization suppression and enhancement, can be affected by the presence of other chemical species and cause inconsistent results in the analysis.[27]

Compared to MS, it stands out that NMR sample preparation is minimal. In NMR a high number of metabolites can be detected simultaneously in a short period of time of only a few minutes, and the sample can be retrieved and stored for a long time and reused.[28]

For these reasons, NMR is found in a myriad of applications in many different areas of science. A clear example of this is its implementation in the analysis of lipoproteins for the prediction of cardiovascular risk.[12,29,30] Recently, there is an increasing interest in the quantification of blood glycoproteins through NMR since most acute-phase proteins that rise or fall in response to acute and chronic inflammatory stimuli are glycoproteins,[1,3] and are thought to be potential markers of inflammation. Specifically, the signal produced by the -COCH3 acetyl groups of some of the GlcNAc, GalNAc and Neu5Ac monosaccharides linked to protein mentioned above has been studied.

1.3.1. Fundamentals of NMR

An NMR spectrometer essentially consists of a magnet, a radio frequency emitter and a radio frequency detector. The nuclei present in a sample are positively charged and have a rotating movement on an

axis behaving like small magnets. The initial arrangement of the nuclei and electron spins within the sample will be the most energetically favourable. When the sample is introduced into the spectrometer, it is placed between the two poles of a magnet and it is subjected to the radiofrequency field (RF) of the emitter, particularized at the proton resonance frequency. The sample is capable of absorbing RF energy (this is what is called entering resonance). The nuclei with positive spin are oriented in the same direction of the field, in a state of minimum energy called spin state α, while nuclei with negative spin are oriented in the opposite direction to that of the magnetic field, in a state of greater energy called spin state β. The energy difference between the two spin states α and β, depends on the applied magnetic field strength. The greater the magnetic field, the greater the energy difference between the two spin states. Since the short RF pulse covers a wide frequency range, the individual protons absorb the frequency radiation necessary to enter resonance (change of spin state).

The integrated surface (Area) of the absorption signal is proportional to the energy absorbed. This, in turn, is proportional to the number of nuclei that have passed from the fundamental state to the excited state or, more precisely, to the difference between the number of nuclei that pass from the fundamental to the and vice versa.

The magnetic nuclei are part of a set of molecules that constitute the sample; the entire molecular system is called a reticulum independently of the physical state of the sample. In general, NMR spectra are determined in dissolution. In liquids, in relation to a nucleus taken as a reference, there are movements of the other molecules

(translations) and movements of other atoms of the molecule of which the nucleus is a part (vibrations and rotations). A nucleus in the upper spin state can relax towards a lower spin state and the eliminated energy will be poured into the network as supplementary energy of translation (intermolecular) and vibration-rotation (intramolecular). The heat capacity of the reticulum is so great with respect to the spin that the dissolution will not heat up. The cooling system is called spin-network relaxation. The spin-network relaxation (T1) allows a nucleus to change spin without absorbing or emitting energy being possible in both directions, from the fundamental to the excited and vice versa. In this way, the system will reach a new equilibrium and the NMR signal will remain stationary.

Another important relaxation is known as spin-spin relaxation (T2). The field associated with an animated nucleus of a precessional movement can be broken down into two components: one static and parallel to the direction of the main field B_0 and the other rotating with the precessional frequency in a plane perpendicular to the direction of the main field. The static component does not intervene in relaxation. A spin-spin relaxation time T2 is defined for the set of these two processes. T2 is the mean time spent by a nucleus in a given spin state and is the time that conditions the natural width of the signal. The Heissenberg uncertainty principle demonstrates that the natural width of a band is inversely proportional to the mean time the system remains in the excited state. The modification of B_0 field from one nucleus to another is another factor that contributes to the widening of the signals. For example, in a solid state the atoms are very close, the T2 will be very small and it will result in wider signals. However, in the liquid

state, the atoms will be very far away, the T2 will be large and it will result in narrower signals.

When the NMR spectrometer detects these signals, it captures the signal known as the free induction decay (FID) and contains the sum of the relaxation of all the excited spins. The computer collects the intensity with respect to time and converts these data into intensity with respect to frequency, this is what is known as the Fourier Transform (FT-RMN). In this way we obtain the graph of the frequencies versus intensity, which is the so-called NMR spectrum. **Figure C1.2** shows a summarized spectrum acquisition workflow.

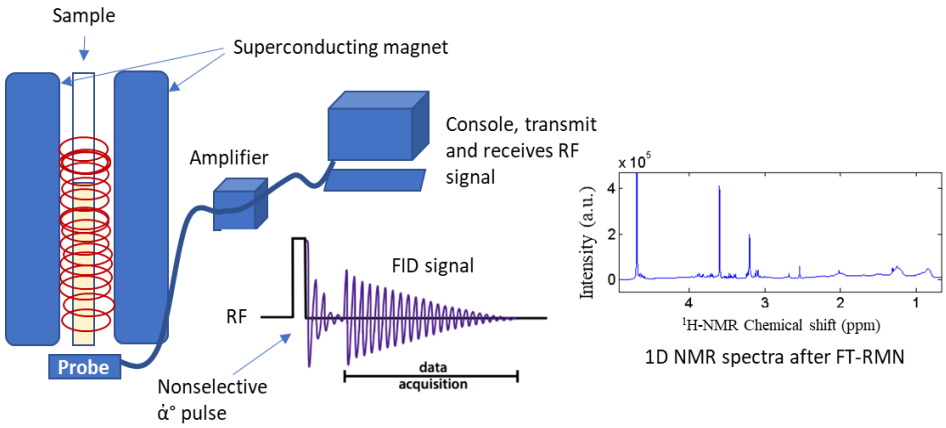


Figure C1.2. Summarized spectrum acquisition workflow

1.3.2. ^{1}H-NMR profiling of biological matrices

Recent decades of advances in NMR have made it a very powerful tool for research, increasingly clinically relevant. Despite its sensitivity limitations in relation to mass spectrometry techniques,

NMR has an unprecedented number of advantages. Compared to traditional methods of global data analysis based on relative comparisons, the absolute quantification offered by NMR offers numerous advantages, including the reduction of possible errors derived from factors such as a wide spectral background, strong solvent signals, and peak misalignments. The ability to analyse intact samples without the need for sample preparation or separation, or minimal sample preparation, represents an important feature of NMR.[31]

However, while biospecimens such as urine and mixtures of metabolites extracted from cells and tissues enjoy the benefit of analysis using intact samples, as they generally lack macromolecules, biospecimens such as human blood serum or plasma, which have high amounts of macromolecules, are much more difficult. The unequivocal assignment of unknown peaks or signals, especially for low-concentration metabolites, remains a challenge because most peaks of small metabolites in the NMR spectrum overlap partially or completely with large signals of more abundant metabolites and pose problems for their identification. However, there are experimental and comprehensive protocols that allow for easy reproduction of NMR. These protocols in turn allow non-expert NMR users to easily identify and quantify blood metabolites, thus enabling the effective use of this tool.[32]

1.3.3. ^{1}H-NMR profiling of biofluids: Serum/plasma

Biofluids are commonly used in metabolomics studies because they contain hundreds to thousands metabolites and samples can be obtained in a non-invasive (e.g., saliva, urine) or minimally invasive manner (e.g., blood plasma or serum).[33] Moreover, sample preparation for NMR experiments on biofluids only requires the addition of phosphate buffer in a small volume of deuterated solvent, and the addition of an internal standard for chemical shift reference and quantitative normalization. Commonly used internal standards are 4.4-dimethyl-4-silapentate-1-sulfonic acid (DSS), 3-(trimethylsilyl) 2,2,3,3-tetradeutero-propionic acid (TSP) that has been used as the preferred reference for calibrating and quantifying aqueous NMR samples and tetramethylsilane (TMS) for organic solvents.

Among the nuclei that can be analysed with NMR, proton (^{1}H-NMR) is the most used in metabolomics because of its high sensitivity, fast relaxation, natural abundance, and its nearly ubiquitous presence in organic metabolites.[33] ^{1}H-NMR spectra of biofluids consist of a conglomerate of severely overlapped signals from a vast number of compounds at very different concentrations, making a reliable identification and quantification a challenging task. Additionally, the spectral complexity is amplified by chemical exchange processes. For instance, pH, ionic strength, and metal ion composition affect specific groups of metabolites, causing chemical shifts variation between samples.[34]

Blood is the primary body fluid connected to systemic metabolism and it is considered the most representative biofluid of the

homeostatic state of the body. However, untreated blood is not a suitable biofluid for lipoprotein NMR measurements largely due to the presence of cells and their sedimentation gives rise to problems of lack of homogeneity in the magnetic field and line broadening. This is why the collection of blood plasma or serum is the first step in most clinical, metabolic and nutritional studies.[13,35] Blood contains molecules of various size and mobility: proteins, lipids, lipoproteins, cholesterol, low-molecular-weight metabolites and ions. Blood plasma is blood without cells, while blood serum is blood plasma without the protein components that constitute the clot. Protocols for serum/plasma sample preparation have been previously described.[36] For some NMR applications serum may be preferred to plasma. Anticoagulant problems such as fibrinogen precipitation on freezing and thawing are avoided if serum is used. **Figure C1.3** shows an example of an ^{1}H-NMR spectra of human serum and its most characteristic identified peaks.

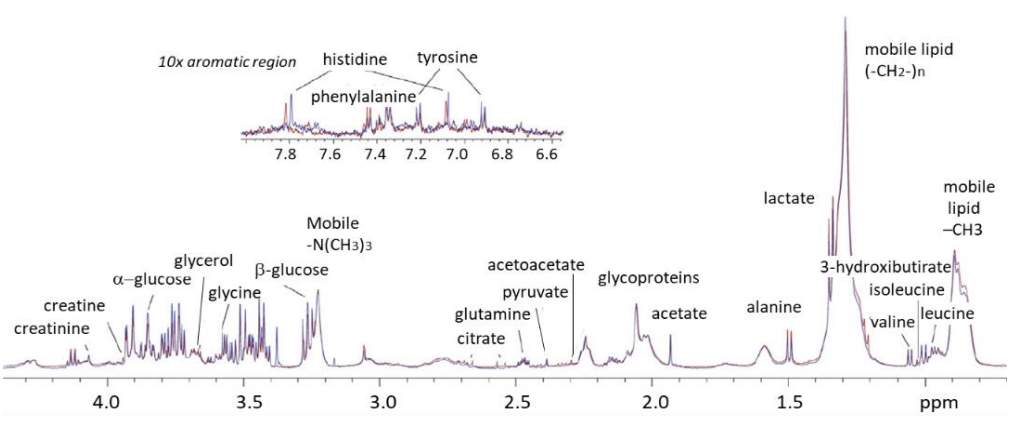


Figure C1.3. ^{1}H-NMR spectra of human serum (real human serum is in blue and human serum mimic in red). and its most characteristic identified peaks. Adapted from M. Tiainen et al.[37]

One major advantage of using NMR spectroscopy to study plasma and serum is that measurements can often be made with minimal sample preparation (usually with only the addition of 5–10% D_2O for locking, which is a process to control magnetic field in the sample so that the resonance frequencies do not drift; and a suitable standard in the sample or in an inner capillary for referencing or quantification purposes). Moreover, a holistic analytical profile can be obtained on the whole biological sample.

For serum/plasma analysis, special consideration must be given to internal standards. DSS is not suitable because it has signals arising at 3.1 ppm (a triplet), 1.9 ppm (a pentet), and 0.8 ppm (another triplet) all of which correspond to methylene protons of the alkylic chain. TSP does not add any other signals apart from the sharp methyl resonance at 0 ppm but TSP must not be used because it interacts with proteins.[38]

One general way to use TSP as a reference substance is to introduce a co-axial inner capillary tube with the compound dissolved in D_2O. This system prevents the TSP from interacting with the blood proteins and prevents the sample from being modified by deuterated solvents. Because of the stability of modern NMR spectrometers, 'Digital ERETIC' (Digital Electronic REference To access In vivo Concentrations, Bruker®) is an interesting method for quantification in serum. It does not involve any reference substance but adds a digital

signal at a chosen frequency to the final corrected spectrum.[39] This signal is previously calibrated against a reference sample and corrected taking into account the variation of the 90° pulse length, the number of scans and the receiver gain, always after the tuning and matching have been automatically adjusted.[40] As will be seen below, this thesis focuses on the analysis of the glycoprotein signal between 2.15-1.90 ppm of the ^{1}H-NMR spectra.

1.3.4. ^{1}H-NMR in clinical field, a general overview

^{1}H-NMR profiling has led to a deeper understanding of disease pathogenesis and the identification of new biomarkers for disease diagnosis or treatment monitoring.[41] There are some examples of the first steps in applying NMR to clinical practice. One of them is the creation of advanced lipoprotein tests where ^{1}H-NMR serum or plasma measurements have been applied to determine lipoprotein subclass composition, size and functionality which has made it possible to predict, detect and monitor cardiometabolic diseases such as cardio vascular diseases (CVDs).[13] Another example of the usefulness of NMR in the clinical field is the NMR-Based Screening for Inborn Errors of Metabolism, which are rare genetic disorders in which the body is not capable of converting food into energy. ^{1}H NMR can detect numerous metabolites in urine along with the enzyme defects that cause the inborn errors of metabolism.[27]

The quantification of serum/plasma NMR signals reflecting glycoprotein acetylation has been shown to be relate to the inflammatory status of patients.[42] Increased circulating glycoproteins have been linked to

several pathologies such as rheumatoid arthritis and its severity as well as to cardiovascular and cancer mortality among others.[41] These findings suggest a potential utility of the NMR-based profiling for glycoprotein quantification as possible markers of systemic inflammation that could provide complementary information in clinical risk assessment. It is therefore a very encouraging challenge to translate this information into clinical practice.

The use of advanced statistics, chemometrics and multivariate algorithms allows transforming a high data set into metabolic profiles, and ultimately into clinical information.

1.4. Thesis motivation

The doctoral thesis presented in this document is the result of the research conducted mainly in the Department of Electronic, Electrical and Automation Engineering at the Universitat Rovira i Virgili (URV) and the spin-off company Biosfer Teslab as a result of the 2014DI0029 industrial doctoral project from the *Generalitat de Catalunya*. Other research institutions have been involved in this research, such as the Metabolomics Platform, a joint research facility created by URV, the Research Unit of Lipids and Atherosclerosis (URLA) of Sant Joan University Hospital from Reus (HUSJR), the CIBER of Diabetes and Metabolic Diseases (CIBERDEM) and the Institut d'Investigació Sanitària Pere Virgili (IISPV). **Figure C1.4** shows the workflow of this thesis.

These collaborations have provided a fruitful environment for the development of this project combining the interaction between three work areas:

1) **Clinical area**, represented mainly by the URLA and other research groups at CIBERDEM and IISPV. These groups have made possible to collect samples of the inflammatory and cardiovascular diseases that are the subject of this thesis.

2) **Technological area** led by the university's Metabolomic Platform whose main objective is to provide technical support to biomedical and clinical research groups in the field of metabolomics. Metabolomic Platform has provided technical support to carry out the analysis of samples by ^{1}H-NMR and the necessary data processing algorithms and statistical and multivariate analysis tools for develop the glycoprotein analysis method.

3) **The industrialization and business development area**. The company Biosfer Teslab, founded in December 2013 in Reus (Tarragona), operates in the field of in vitro diagnostics to provide analytical services to study and monitor alterations in lipid metabolism and its associated cardiovascular risk. and, bets on R&D as the main mechanism for its future survival. The knowledge and advice acquired in Biosfer, has allowed, on the one hand, to perfect the development of the methodology and, on the other hand, the future industrialization of the method. It is remarkable that the creation of Biosfer Teslab comes from

the results obtained by the research group in which the present thesis has been developed and which were based on the development of the patented and validated Liposcale test, an advanced lipoprotein test based on diffusion ordered NMR spectroscopy (DOSY-NMR) that measures the size, the lipid content and the number of the particles of the main types of lipoprotein (VLDL, LDL and HDL) and the concentration of particles of nine subtypes (large, medium and small of each main type).[43] The purpose of the present thesis is to provide more functionality to this test, incorporating the results of the glycoprotein profiling, as a markers of systemic inflammation.

In these years, Biosfer has achieved the industrial development of the Liposcale test surpassing all regulatory requirements to achieve CE certification and by the Spanish Agency of Medicine and Medical Devices (AEMPS) for manufacture of medical software for *in-vitro* diagnostic applications. It has also established solid research collaborations and has obtained good evaluations in different innovation and development projects. Reinforcing the innovative value of the company, since August 2014, Products and Technology (P&T) (a subsidiary of the Catalan pharmaceutical company Laboratorios Rubió) had become part of Biosfer's capital stock, enabling the industrial development and the introduction of new diagnostic tests in the clinic through a more consolidated business infrastructure, knowledge of the market and an extensive commercial network.

CHAPTER 1

It is necessary to point out two main advantages of the realization of this thesis within this ambit. The first is that because the Liposcale test has a similar methodology based on ^{1}H-NMR, the development of glycoprotein characterization can be done on the same NMR experiment used for Liposcale test. This means considerable experimental savings as well as a second perfectly qualified product (glycoprotein test) expanding the company's market field. Secondly, the development of this thesis within the Biosfer Teslab has a strong direct impact on society, as translational research takes place, bringing incalculable added value to the health sector and contributing to an improvement in the diagnosis and prognosis of inflammatory and cardiovascular diseases that are so frequent in society.

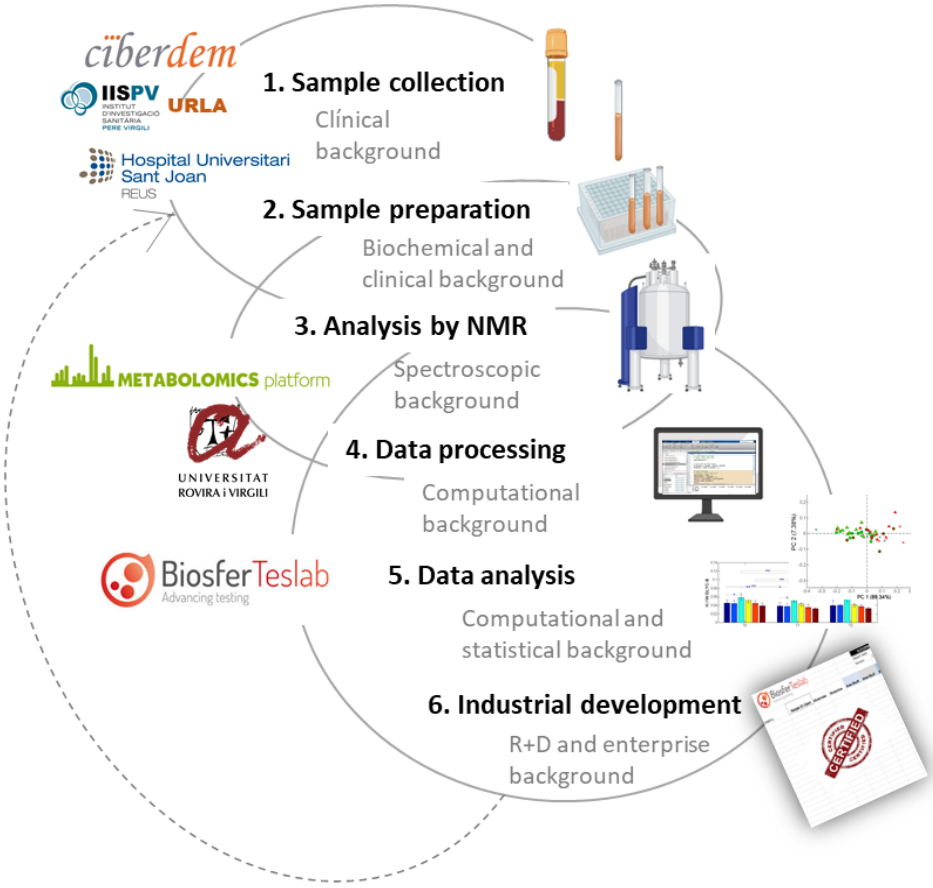


Figure C1.4. Workflow of this thesis including the mentioned collaborations.

1.5. Hypothesis and objectives

As discussed in the previous points, recently there is a growing interest in the study of new inflammatory markers that may help in the prognosis and diagnosis of some diseases. Most of the acute phase proteins that rise when there is an inflammatory process are

glycoproteins. Development of high-throughput methodologies to identify and quantify glycoproteins, as NMR based on ones, has become a present challenge in clinical research to provide clinicians with better tools for clinical diagnostic. In fact, several studies have recently been published in the literature showing the advantages of the NMR-GlycA marker over other traditional inflammatory markers such as CRP.[15–17] These advantages include lower intra-individual variability[15] and improved discriminatory capacity for CVD risk models[44,45], among others.

Due to all this, we hypothesize that the quantification of glycoproteins through ^{1}H-NMR, will allow to measure in a more robust way the degree of inflammation that with the habitual markers that are used in the current clinic.

The aim of this thesis has been to develop and evaluate a novel method based on 1D ^{1}H-NMR spectroscopy to quantify the serum/plasma circulating glycoproteins. Therefore, the goals related to this purpose are the following:

- Methodological and computational design to robustly deconvolute the resonance region of glycoproteins in the ^{1}H-NMR spectrum and characterization of the resulting functions included in the developed method.
- Characterization of glycoproteins in clinical sets of patients with rheumatoid arthritis, obesity, polycystic ovary syndrome, diabetes and atherogenic dyslipidaemia, among others.

- To evaluate the added value of glycoprotein profiling in the clinical field through the application of univariate statistical analysis techniques and multivariate models.

1.6. Organization of the document

This chapter has provided a general overview of the background and motivation for the realization of this thesis. As has been highlighted previously, interest in the search for new biomarkers that improve the diagnosis and prognosis of inflammatory diseases is constantly growing. Glycoproteins are good inflammatory biomarkers candidates that have been related to inflammatory diseases such as rheumatoid arthritis and even cardiovascular diseases. For this purpose, NMR spectroscopy appears to be a suitable analytical tool to develop novel methodologies, which will aid in the assessment of inflammatory diseases. The perfect cooperation between a clinical group in close contact with patients, a research group at the forefront of technological development and a spin-off company focused on reducing time between industrial research and the market access, makes potentially feasible the idea of enabling new technologies to be used in clinical environment

Chapter 2 describes the different methodological approaches that have been carried out during the development the method that ensures a robust characterization of the glycoprotein peak in the ^{1}H-NMR spectra. It is divided into several sections dealing with 1D ^{1}H-NMR spectral data pre-processing, the development of the

deconvolving model by using lineshape fitting analysis, the study and characterization of functions (GlycB, GlycF and GlycA) and derived-variables that compose the signal (areas and shape ratio H/W of each function) and finally, a section devoted to the study of the stability of the glycoprotein characterisation method as a function of time and different forms of sample storage.

Chapter 3 presents the state of the art of the growing interest in ^{1}H-NMR circulating glycoproteins determination and their application in the clinical field. The first part of the review describes the role of protein glycation and its relevance in the clinical field. The second part of the review focuses on the analysis of glycoproteins through NMR spectroscopy in different diseases such as cancer, metabolic disorders, CVD risk, human immunodeficiency virus (HIV)-infection, chronic inflammatory diseases and psychological health, among others.

Chapter 4 to 6 contain the results derived from the use on NMR glycoprotein characterization in different cohorts. **Chapter 4** describes our first attempt to characterize ^{1}H-NMR plasma glycoproteins in a rheumatoid arthritis (RA) cohort. We studied the use of NMR derived glycoprotein profile as an inflammatory biomarker. We model the activity of RA to identify patterns indicating the severity of the disease based on DAS28 index, a measure of the disease activity of patients with RA. These results have been published in the *Journal of Proteome Research.*

Chapter 5 describes the influence of sexual steroids and obesity on ^{1}H-NMR glycoprotein profiles. We extend the previous

NMR glycoprotein profile by adding a new variable called GlycF. We studied this new glycoprotein profile in nonobese and obese women with polycystic ovary syndrome (PCOS) and compared them with male and female control subjects. These results have been submitted to the *Journal of Proteome Research*. A second study carried out in the same population about the effect of post-prandial nutritional intervention consisted on protein, glucose and lipid intake on the NMR glycoprotein profile is still under preparation.

Chapter 6 is a study focused on the description of plasmatic ^{1}H-NMR glycoprotein profile of patients with type 2 diabetes (T2DM) with and without atherogenic dyslipidaemia (AD) and to explore its association with their NMR lipoprotein profile. The second part of the study compares whether these new inflammatory markers based on NMR have a greater ability to discriminate specific patterns between different study groups than those commonly used in clinical settings such as CRP.

Chapter 7 contain the general discussion and the conclusions of this thesis, respectively.

1.7. References

(1) Gornik, O.; Lauc, G. Glycosylation of Serum Proteins in Inflammatory Diseases. *Dis. Markers* **2008**, *25* (4–5), 267–278.

(2) Fuertes-Martín, R.; Taverner, D.; Vallvé, J.-C.; Paredes, S.; Masana, L.; Correig Blanchar, X.; Amigó Grau, N. Characterization of 1 H NMR Plasma Glycoproteins as a New Strategy To Identify Inflammatory Patterns in Rheumatoid

Arthritis. *J. Proteome Res.* **2018**, *17* (11), 3730–3739.

(3) Gabay, C.; Kushner, I. Acute-Phase Proteins and Other Systemic Responses to Inflammation. *N. Engl. J. Med.* **1999**, *340* (6), 448–454.

(4) Alper, J. Searching for Medicine's Sweet Spot. *Science* **2001**, *291* (5512), 2338–2343.

(5) Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; et al. *Essentials of Glycobiology*; Cold Spring Harbor Laboratory Press, 2015.

(6) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **2006**, *126* (5), 855–867.

(7) Kushner, I. Regulation of the Acute Phase Response by Cytokines. *Perspect. Biol. Med.* **1993**, *36* (4), 611–622.

(8) Bekesi, R. A. L. M. and J. G. *Plasma Glycoproteins in Various Disease States Including Carcinoma*; Waverly Press, 1962; Vol. 22.

(9) Dall'Olio, F., Vanhooren, V., Chen, C.C., Slagboom, P.E., W.; M., F. N-Glycomic Biomarkers of Biological Aging and Longevity: A Link with Inflammaging. *Ageing Res. Rev.* **2013**, *12* (2), 685–698.

(10) Pezer, M.; Rudan, I.; Campbell, H. Mechanisms of Disease: The Human N-Glycome. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (8), 1574–1582.

(11) Anderson, L.; Anderson, N. G. High Resolution Two-Dimensional Electrophoresis of Human Plasma Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74* (12), 5421–5425.

(12) Jeyarajah, E. J.; Cromwell, W. C.; Otvos, J. D. Lipoprotein Particle Analysis by Nuclear Magnetic Resonance Spectroscopy. *Clin. Lab. Med.* **2006**, *26* (4), 847–870.

(13) Mallol, R.; Rodriguez, M. A.; Brezmes, J.; Masana, L.; Correig, X. Human Serum/Plasma Lipoprotein Analysis by NMR: Application to the Study of Diabetic Dyslipidemia. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *70*, 1–24.

(14) Otvos, J. D.; Shalaurova, I.; Wolak-Dinsmore, J.; Connelly, M. A.; Mackey, R. H.; Stein, J. H.; Tracy, R. P. GlycA: A Composite Nuclear Magnetic Resonance Biomarker of Systemic Inflammation. *Clin. Chem.* **2015**.

(15) Otvos, J. D.; Guyton, J. R.; Connelly, M. A.; Akapame, S.; Bittner, V.; Kopecky, S. L.; Lacy, M.; Marcovina, S. M.; Muhlestein, J. B.; Boden, W. E. Relations of GlycA and Lipoprotein Particle Subspecies with Cardiovascular Events and Mortality: A Post Hoc Analysis of the AIM-HIGH Trial. *J. Clin. Lipidol.* **2018**.

(16) Akinkuolie, A. O.; Pradhan, A. D.; Buring, J. E.; Ridker, P. M.; Mora, S. Novel Protein Glycan Side-Chain Biomarker and Risk of Incident Type 2 Diabetes Mellitus. *Arterioscler. Thromb. Vasc. Biol.* **2015**, *35* (6), 1544–1550.

(17) Akinkuolie, A. O.; Buring, J. E.; Ridker, P. M.; Mora, S. A Novel Protein Glycan Biomarker and Future Cardiovascular Disease Events. *J. Am. Heart Assoc.* **2014**, *3* (5), e001221.

(18) Ritchie, S. C.; Würtz, P.; Nath, A. P.; Abraham, G.; Havulinna, A. S.; Fearnley, L. G.; Sarin, A.-P.; Kangas, A. J.; Soininen, P.; Aalto, K.; et al. The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection. *Cell Syst.* **2015**, *1* (4), 293–301.

(19) Colley, K. J.; Varki, A.; Kinoshita, T. *Cellular Organization of Glycosylation*; Cold Spring Harbor Laboratory Press, 2015.

(20) Freeze, H. H.; Hart, G. W.; Schnaar, R. L. *Glycosylation Precursors*; Cold Spring Harbor Laboratory Press, 2015.

(21) Stanley, P.; Taniguchi, N.; Aebi, M. *N-Glycans*; Cold Spring

Harbor Laboratory Press, 2015.

(22) Brockhausen, I.; Stanley, P. O-GalNAc Glycans. *Essentials Glycobiol.* **2017**.

(23) Brailsford, J. A.; Danishefsky, S. J. Probing the Stability of Nonglycosylated Wild-Type Erythropoietin Protein via Reiterative Alanine Ligations. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (19), 7196–7201.

(24) Maverakis, E.; Kim, K.; Shimoda, M.; Gershwin, M. E.; Patel, F.; Wilken, R.; Raychaudhuri, S.; Ruhaak, L. R.; Lebrilla, C. B. Glycans in the Immune System and The Altered Glycan Theory of Autoimmunity: A Critical Review. *J. Autoimmun.* **2015**, *57*, 1–13.

(25) Connelly, M. A.; Gruppen, E. G.; Otvos, J. D.; Dullaart, R. P. F. Inflammatory Glycoproteins in Cardiometabolic Disorders, Autoimmune Diseases and Cancer. *Clin. Chim. Acta* **2016**, *459*, 177–186.

(26) Domann, P. J.; Pardos-Pardos, A. C.; Fernandes, D. L.; Spencer, D. I. R.; Radcliffe, C. M.; Royle, L.; Dwek, R. A.; Rudd, P. M. Separation-Based Glycoprofiling Approaches Using Fluorescent Labels. *Proteomics* **2007**, *7* (S1), 70–76.

(27) Capati, A.; Ijare, O. B.; Bezabeh, T. Diagnostic Applications of Nuclear Magnetic Resonance–Based Urinary Metabolomics. *Magn. Reson. Insights* **2017**, *10*, 1178623X1769434.

(28) Emwas, A.-H. M. The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research. In *Metabonomics*; Humana Press, New York, NY, 2015; pp 161–193.

(29) Ala-Korpela, M.; Korhonen, A.; Keisala, J.; Hörkkö, S.; Korpi, P.; Ingman, L. P.; Jokisaari, J.; Savolainen, M. J.; Kesäniemi, Y. A. 1H NMR-Based Absolute Quantitation of Human Lipoproteins and Their Lipid Contents Directly from Plasma. *J.*

Lipid Res. **1994**, *35* (12), 2292–2304.

(30) Mallol, R.; Rodriguez, M. A.; Brezmes, J.; Masana, L.; Correig, X. Human Serum/Plasma Lipoprotein Analysis by NMR: Application to the Study of Diabetic Dyslipidemia. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *70*, 1–24.

(31) Fan, T. W.-M.; Lane, A. N. Applications of NMR Spectroscopy to Systems Biochemistry. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *92–93*, 18–53.

(32) Nagana Gowda, G. A.; Raftery, D. Recent Advances in NMR-Based Metabolomics. *Anal. Chem.* **2017**, *89* (1), 490–510.

(33) Larive, C. K.; Barding, G. A.; Dinges, M. M. NMR Spectroscopy for Metabolomics and Metabolic Profiling. *Anal. Chem.* **2015**, *87* (1), 133–146.

(34) Lindon, J. C.; Nicholson, J. K.; Holmes, E. *The Handbook of Metabonomics and Metabolomics*; Elsevier, 2007.

(35) Soininen, P.; Kangas, A. J.; Würtz, P.; Tukiainen, T.; Tynkkynen, T.; Laatikainen, R.; Järvelin, M.-R.; Kähönen, M.; Lehtimäki, T.; Viikari, J.; et al. High-Throughput Serum NMR Metabonomics for Cost-Effective Holistic Studies on Systemic Metabolism. *Analyst* **2009**, *134* (9), 1781.

(36) Beckonert, O.; Keun, H. C.; Ebbels, T. M. D.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Metabolic Profiling, Metabolomic and Metabonomic Procedures for NMR Spectroscopy of Urine, Plasma, Serum and Tissue Extracts. *Nat. Protoc.* **2007**, *2* (11), 2692–2703.

(37) Tiainen, M.; Soininen, P.; Laatikainen, R. Quantitative Quantum Mechanical Spectral Analysis (QQMSA) of 1H NMR Spectra of Complex Mixtures and Biofluids. *J. Magn. Reson.* **2014**, *242*, 67–78.

(38) Bell, J. D.; Brown, J. C. C.; Sadler, P. J. NMR Studies of Body Fluids. *NMR Biomed.* **1989**, *2* (5–6), 246–256.

(39) Virginie Silvestre; Stéphane Goupry; Michel Trierweiler; Richard Robins, A.; Akoka*, S. Determination of Substrate and Product Concentrations in Lactic Acid Bacterial Fermentations by Proton NMR Using the ERETIC Method. *Anal. Chem* **2001**, *73* (8), 1862–1868.

(40) Farrant, R. D.; Hollerton, J. C.; Lynn, S. M.; Provera, S.; Sidebottom, P. J.; Upton, R. J. NMR Quantification Using an Artificial Signal. *Magn. Reson. Chem.* **2010**, *48* (10), 753–762.

(41) Ala-Korpela, M. Serum Nuclear Magnetic Resonance Spectroscopy: One More Step toward Clinical Utility. *Clin. Chem.* **2015**, *61* (5), 681–683.

(42) Torri, G. M.; Torri, J.; Gulian, J. M.; Vion-Dury, J.; Viout, P.; Cozzone, P. J. Magnetic Resonance Spectroscopy of Serum and Acute-Phase Proteins Revisited: A Multiparametric Statistical Analysis of Metabolite Variations in Inflammatory, Infectious and Miscellaneous Diseases. *Clin. Chim. Acta.* **1999**, *279* (1–2), 77–96.

(43) Mallol, R.; Amigó, N.; Rodríguez, M. A.; Heras, M.; Vinaixa, M.; Plana, N.; Rock, E.; Ribalta, J.; Yanes, O.; Masana, L.; et al. Liposcale: A Novel Advanced Lipoprotein Test Based on 2D Diffusion-Ordered ^{1}H NMR Spectroscopy. *J. Lipid Res.* **2015**, *56* (3), 737–746.

(44) Joshi, A. A.; Lerman, J. B.; Aberra, T. M.; Afshar, M.; Teague, H. L.; Rodante, J. A.; Krishnamoorthy, P.; Ng, Q.; Aridi, T. Z.; Salahuddin, T.; et al. GlycA Is a Novel Biomarker of Inflammation and Subclinical Cardiovascular Disease in Psoriasis. *Circ. Res.* **2016**, *119* (11).

(45) McGarrah, R.; Craig, D.; Haynes, C.; Dowdy, Z. E.; Shah, S.; Kraus, W. GlycA, a Novel Biomarker of Systemic Inflammation, Improves Cardiovascular Risk Prediction in a High-Risk Coronary Catheterization Cohort. *J. Am. Coll. Cardiol.* **2015**, *65* (10), A1606.

CHAPTER 2

2. Development of a glycoprotein characterization method by ^{1}H-NMR spectroscopy

CHAPTER 3

3. Human serum/plasma glycoprotein analysis by ^{1}H-NMR, an emerging method of inflammatory assessment

3.1. Introduction

Plasma glycoproteins belong to the large family of acute-phase proteins (APPs), which are directly related to inflammatory disorders.[1] Research in this area is expanding as several studies suggest that the change in the concentration of APPs and the pattern of glycation of these proteins can influence cellular changes in a large number of diseases [2–4] so they can be regarded as diagnostic markers.[5]

It should be noted that glycoproteome analysis is much more complex than proteome analysis because, unlike proteins in which amino acid sequences are unique, oligosaccharides and polysaccharides from glycans are normally composed of an enormous diversity of both linear and branched sugar residues, which increase the complexity of the glycoprotein structures.[6] For this reason, the analysis of glycoproteins is often a technical challenge. In recent years, nuclear magnetic resonance (NMR) has played a major role as an analytical tool for metabolomic studies with biological fluids, especially for serum and plasma samples. A clear example of this is its application for the determination of lipoproteins.[7] Unlike other techniques, NMR is capable of quantifying metabolites in a reproducible and effective way, so it is widely used in large epidemiological studies and has started to be introduced in routine clinical practice.[8] In this review, we will focus on NMR as a promising technique for quantifying glycoproteins in serum or plasma.

This review reflects the growing interest in determining proton NMR (^{1}H-NMR) circulating glycoproteins and their application in the

clinical field, which will pave the way for these markers to be used in the prediction and monitoring of several diseases.

3.2. Glycoproteins, a biochemical approach

3.2.1. Proteins, general concepts

Proteins are synthesized in the body through a process called translation, which occurs in the cytoplasm and involves converting genetic codes into proteins. Each type of protein consists of a precise sequence of amino acids that allows it to fold up into a three-dimensional shape, or conformation. Proteins are responsible for nearly every task of cellular life: providing the cell with its structural support in the case of cytoskeletal proteins, favouring chemical reactions in the case of enzymes, controlling intracellular traffic and the flow of substances between the cell and the outside, or regulating gene expression.

However, proteins can undergo numerous chemical modifications in their structure that have important modulating effects on their biological function, alter their cellular location and capacity to interact with other proteins, or even determine their own degradation. These modifications occur in proteins once they have synthesized and are called post-translational modifications.

3.2.2. Post-translational modifications (PTMs) of proteins: Glycosylation

PTMs increase the functional diversity of the proteome. These modifications include phosphorylation, myristoylation, farnesylation, cysteine oxidation, ubiquitination, acetylation, phosphorylation, glycosylation, methylation, nitrosylation, etc, and influence almost all aspects of normal cell biology and pathogenesis.[9] Glycosylation is the addition of one or more chains of carbohydrates (glycans) to a protein. This is the main chemical modification of most plasma-membrane and secretory proteins.[10] Glycoproteins participate in many key biological processes including cell adhesion, molecular trafficking and clearance, receptor activation, signal transduction, and endocytosis. Most of the proteins in blood plasma (except for albumin) are highly glycosylated, and the glycosylation of these and other secreted proteins can provide solubility, hydrophilicity and negative loading, thus reducing unwanted intermolecular interactions and protecting them against proteolysis or simply varying their function. Cell surface membrane proteins, such as receptors, adhesion molecules and channels, are also typically glycosylated, and this modification can also change their function.[11]

Glycosylation is regarded as the most complex PTM because of the large number of enzymatic steps involved.[12] Most of the proteins secreted in eukaryotic cells are translocated to the endoplasmic reticulum (ER) where they are folded, modified and subjected to quality control mechanisms. The protein component of all glycoproteins is synthesized in the rough ER (RER). They then make their way through multiple stacks of the Golgi apparatus (from *cis*

Golgi to *trans* Golgi), finally being distributed to various destinations from the trans-Golgi network. This is where the addition of glycans to the polypeptide chain of the protein occurs by complex dynamic interactions between hundreds of enzymes such as glycosyltransferase reactions, enzymes that transfer activated forms of monosaccharides from nucleotide sugars and lipid-linked sugar intermediates to acceptors including proteins, lipids, and growing glycan chains.[11] Some of these activated forms of monosaccharides are mannose, fucose, galactose, N-Acetylglucosamine (GlcNAc), N-Acetylgalactosamine (GalNAc), N-acetylneuraminic acid (Neu5Ac) or sialic acid, among others. **Figure C3.1** shows an overview of the cellular organization of glycosylation.

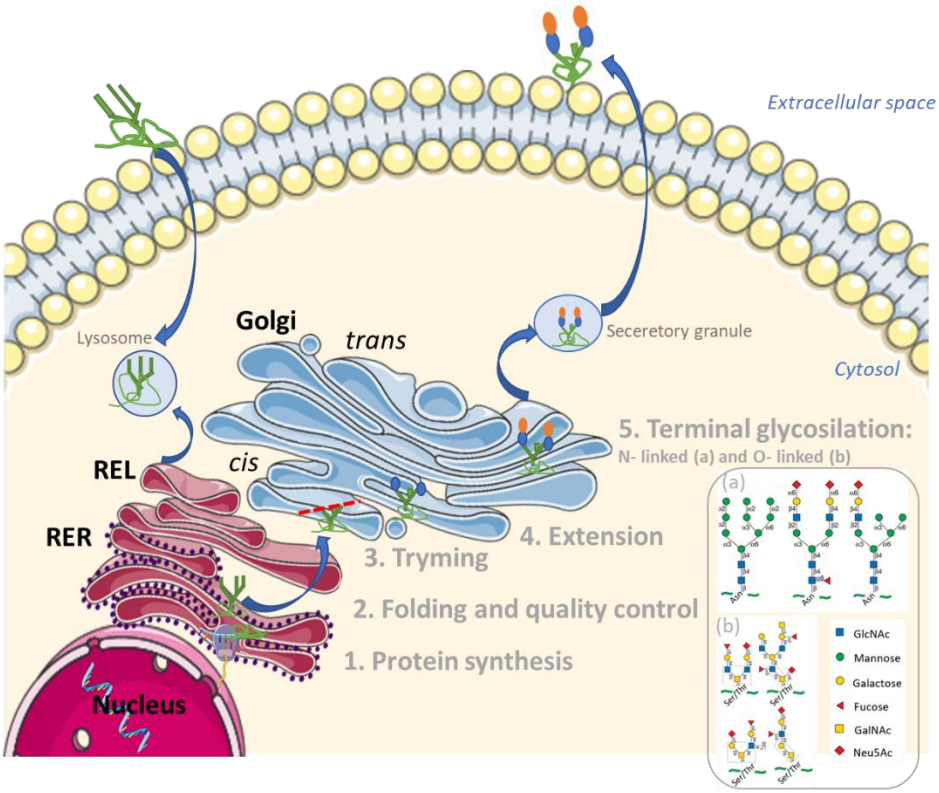


Figure C3.1. Initiation and maturation of glycoproteins in the ER–Golgi–plasma membrane pathway. This illustration outlines an overview of the mechanisms for initiation, trimming, and elongation of the glycoprotein in a human cell. Orange and blue spheres represent the addition of glycans chains to proteins (in green) in the Golgi apparatus. Examples of N-glycans structures (a) and O-Glycans structures (b) are also represented.

Protein glycosylation encompasses N-glycans, O-glycans, and glycosaminoglycans (frequently referred to as proteoglycans).[13] N- and O-glycosylation are the most commonly detected types. The structures of N- and O-linked oligosaccharides are very different, and different sugar residues are usually found in each type.[14] In all N-linked oligosaccharides, GlcNAc is linked to the amide nitrogen of the asparagine (Asn) of a consensus peptide sequence Asn-X-Ser [15] (with X being any amino acid except proline) and they always contain mannose. This glycosylation usually has several branches each terminating with a negatively charged Neu5AC or sialic acid residue.[14,16,17] However, O-linked oligosaccharides are linked to the hydroxyl group of serine (Ser) or threonine (Thr) via GalNAc or (in collagens) to the hydroxyl group of hydroxylysine via galactose and they are generally short, often containing only one to four sugar residues. [14,16,17]

It is important to note that O-glycosylation is more abundant intracellularly and has been associated mostly with protein signalling and intracellular mechanisms, while N-glycosylation is predominant in circulating proteins.[18] In addition, N-glycan synthesis can be easily

altered by pathophysiological conditions such as inflammatory and autoimmune diseases and in the pathophysiological process of aging, which is why N-glycans are emerging as powerful and reliable biomarkers of several diseases[15] as we shall see below.

3.3. Clinical importance of glycoproteins

Both N-glycans and O-glycans play an important role in the functions of the glycoprotein involved in various cell recognition signals and pathological situations,[17,19–24] so they are potentially potent and reliable biomarkers of various diseases. Because of the large number of biological processes in which glycans participate, it is not surprising that defects in the synthesis of glycans can be the direct cause of numerous diseases and, therefore, markers of the disease.[5] It was not until the 1980s that NMR began to arouse great interest in the search for clinically relevant markers, includingAPPs.[25] Since before 1987, it has been reported that the concentration of plasma glycoproteins changes in a number of clinical disorders characterised by inflammation (e.g. different types of cancer, rheumatoid arthritis, some liver diseases, trauma, etc) and pregnancy.[26–30] The altered synthesis of N-glycans is thought to underlie these pathological conditions.[15] It is important to note that most serum glycoproteins have N-linked, and less frequently O-linked, sugars in their structure. Nonetheless, there is still much to learn about the role of glycans in disease mechanisms. However, as more information on protein glycation emerges, it becomes increasingly clear that glycation is

strictly regulated and that the binding of glycan to proteins is of paramount physiological importance.[31]

In general, the most studied glycoproteins of clinical importance are the glycoproteins of the cell membrane, whose glycans, called glycocalyx, play important roles in the immune response. For example, selectins are a widely studied family of membrane proteins that are glycosylated and play a crucial role in the recruitment of leukocytes, the onset of the immune response and the onset of inflammation.[32] It has been shown that the deregulation of selectins or their glycoprotein ligand are associated with atherosclerosis, thrombosis and even the metastasis of tumours.[33]

We want to place special emphasis on the glycoproteins of human plasma, which in recent years have been attracting considerable interest as possible biomarkers of disease.[34] Human plasma glycoproteins belong to the large family of APPs, which are characterized by increasing or decreasing their concentration (positive or negative acute-phase proteins respectively) by up to 25 percent during inflammatory disorders.[1] Most of these APPs are glycosylated proteins secreted from hepatocytes.[35] The alterations in the glycosylation of these proteins indicate cellular changes in a large number of diseases, which is why they can be regarded as diagnostic markers of a disease. Numerous changes in the glycosylation of serum proteins have been reported for inflammatory diseases. **Table C3.1** shows some examples of glycation changes in serum glycoproteins that have been associated with various inflammatory diseases.

Table C3.1: Examples of serum protein glycation changes and their association with disease

Glycoprotein	Glycation change	Related diseases	References
Alpha 1-Acid Glycoprotein (AGP)	Highly branched N-linked glycan	Cirrhosis and HCC, congenital disorders, RA, SLE	176,177
	Increased sialylation	Cancer	178
	Decreased sialylation	Cirrhosis and HCC	176,179
	Increased fucosylation	liver cancer	180
Alpha-Fetoprotein	Elevated bisecting N-acetylglucosamine, decrease in sialylation and increase in fucosylation	Cirrhosis, hepatitis and HCC	181–183
Alpha -1-antitrypsin (ATT)	Increased fucosylation alpha	Hepatitis C, HCC	184,185
	Decrease in branching, predominance of alpha 2-6 linked sialic acid and less alpha 2-3 linked sialic acid	Breast and ovarian cancer	186
	Oligosaccharide branching and increased sialic acid content	Acute general inflammation	187
	Increased glycan branching	RA	188
Transferrin	Increased branching and fucosylation of N-glycans. Increasing peripheral N-acetylglucosamine residues	Ovarian, breast and colon cancer, HCC, Cirrhosis, hepatitis	189,190
	Increased fucosylation and sialic acid-linked to galactose	Liver disease	191
Haptoglobin (Hp)	Increased fucose and N-acetylglucosamine	Alcoholic liver disease	192

	Increased fucosylation	Various types of cancer and RA	[193–196]
	O-glycans ad fucosylation	Prostate cancer	[197]
Inmunoglobulin G (IgG)	Decreased galactose	RA, SLE, IBD, ovarian cancer, prostate cancer	[198–204]
	Increased N-acetylglucosamine residues (controversy)	RA	[205,206]
Inmunoglobulin A (Ig A)	Reduced galactosylation of O-linked glycosylation	Nephropathy	[207,208]

Abbreviations: HCC, Hepatocellular Carcinoma; SLE, Systemic Lupus Erythematosus; RA, Rheumatoid arthritis; IBD, inflammatory bowel disease.

3.4. Measurement techniques for glycoprotein determination

3.4.1. Traditionally used techniques to measure glycated proteins

Various techniques have been used to measure individual glycated proteins. Examples of these are Lectin analysis,[36] DNA sequencing equipment-Fluorophore Assisted Carbohydrate Electrophoresis (DSA-FACE),[15] chromatographic methods such as high performance liquid chromatography (HPLC)[37] or hydrophilic interaction liquid chromatography (HILIC),[36] MALDI-TOF analysis of tryptic glycopeptides,[36] electrophoresis,[27] MS[38] or the recently developed glycoblotting method that combines the BlotGlycoABC bead and MALDI-TOF MS to detect abnormal glycosylation patterns in the whole serum glycoproteins.[39,40] In a recent review, A. Conelly et al. described the current assays of glycoproteins in biological fluids, among which were the enzyme linked immunosorbent assays (ELISAs), electrochemiluminescence immunoassay (ECLIA), luminex based assays, radioimmunoassays (RIA) and nephelometric assays, which quantify the amount of protein in biological samples.[41]

Most of these techniques measure a specific individual glycoprotein. However, measuring the glycan portion of inflammatory proteins is becoming increasingly useful for diagnostic purposes.[41] Some of the newest high-performance techniques such as mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR), which have recently been introduced to the clinical laboratory, have been used for this purpose.

It is important for techniques to be cost-effective so throughput must be high and molecular measurements absolute. For this reason, MS and NMR have been the most used techniques. MS is used for more detailed characterizations and is based on mass difference, but it produces complicated spectroscopic data and is expensive. NMR, however, is currently the only methodology capable of providing reproducible quantifications of high-performance metabolites in a cost-effective manner.[8] It has also been widely used in recent years to quantify lipoproteins in a fully optimized way and has generated considerable medical advances.[7] Although NMR is not as specific as MS, one major advantage is that in a very short time it can provide a complete metabolic profile of a serum or plasma sample.[8]

In this review, we focus on ^{1}H-NMR applications based on methods for quantifying circulating glycoproteins. The great versatility of NMR and the integrated computational methods of systems biology provides robust and reliable tools for biomedical research.

3.4.2. Serum/ plasma NMR glycoprotein analysis

As mentioned in the section above, measuring lipoproteins in plasma and serum has been a key issue in recent years. Similarly, the measurement of glycoproteins is becoming increasingly important, particularly because they can be quantified from the same experiment used for lipoproteins, which means considerable savings and maximum profitability. Another major advantage associated with NMR is the low experimental variability between laboratories. In fact, greater variability is due to interpersonal variability itself and measurements on different days.[42]

3.4.2.1. Sample handling and preparation

Serum and plasma fractions are taken from blood samples that have undergone various biochemical protocols after collection.[43] In the case of serum, coagulation factors (i.e. fibrinogen) and blood cells are removed by centrifugation, while plasma is typically obtained from blood samples with an added anticoagulant agent (i.e. heparin or EDTA). These agents produce high intensity peaks (EDTA) or overlapping signals (heparin) in the NMR spectra,[44] which is why for some NMR applications serum may be preferred to plasma. However, the metabolic profiles detected in plasma and serum by NMR are comparable although signals from EDTA complicate the plasma spectrum profile.[45]

One major advantage of NMR in the study of plasma and/or serum is that measurements can often be made with minimal sample preparation. Normally, in NMR experiments on biofluids samples only require the addition of phosphate buffer in a small volume of deuterated solvent, and an internal standard for chemical shift reference and quantitative normalization.

3.4.2.2. Sample storage

For the analysis of lipoproteins and other plasma/serum metabolites by NMR, some storage issues have to be considered. Samples can be stored in good refrigerated conditions for several days at 2–4°C, and up to 7 days at temperatures below 4°C.[46] They can be successfully stored at -20°C for a moderate period of time (up to 1–2 months), but some enzymes, such as plasma esterase, may still be

active at this temperature. Therefore, for longer storage periods -70°C or -80°C is required.[47] In terms of stability, some studies report a high degree of stability of glycoproteins measured by NMR in frozen samples and stored for more than 10 years.[48]

However, the protocols established for the preservation of serum and plasma samples may be used in order to minimize a possible variability in the results obtained by different analytical platforms.[44,49]

3.4.2.3. Processing of NMR spectra for glycoprotein profiling

The procedure for obtaining the spectrum before quantification of the glycoproteins has been extensively described.[7] Briefly, the spectrum goes through several phases before it is for glycoprotein quantification. First it is obtained by applying one or more pulse sequences. Generally, the most standard pulses are the Nuclear Overhauser Effect Spectroscopy (NOESY)-presaturation sequence that acquires a quantitative serum spectrum by suppressing the water peak, the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence that acquires low molecular weight metabolites, a diffusion-edited pulse sequence with bipolar gradients and, finally, the longitudinal eddy-current delay (LED) with presaturation of the water signal. Most authors use a single pulse for glycoprotein profiling, the most common being CPMG, although LED or NOESY are also used, as can be seen in **table C3.1S** of **supporting information.** The next step is to apply an algorithm to quantify the region in which glycoproteins resonate. Several algorithms have been described in the literature; binning, peak alignment, and combinations of peak alignment and data reduction such as PARS, the curve-fitting algorithm, the peak alignment tools in

HiRes, and targeted profiling.[50] **Figure C3.2** shows a scheme of the analysis of glycoproteins by ^{1}H-NMR.

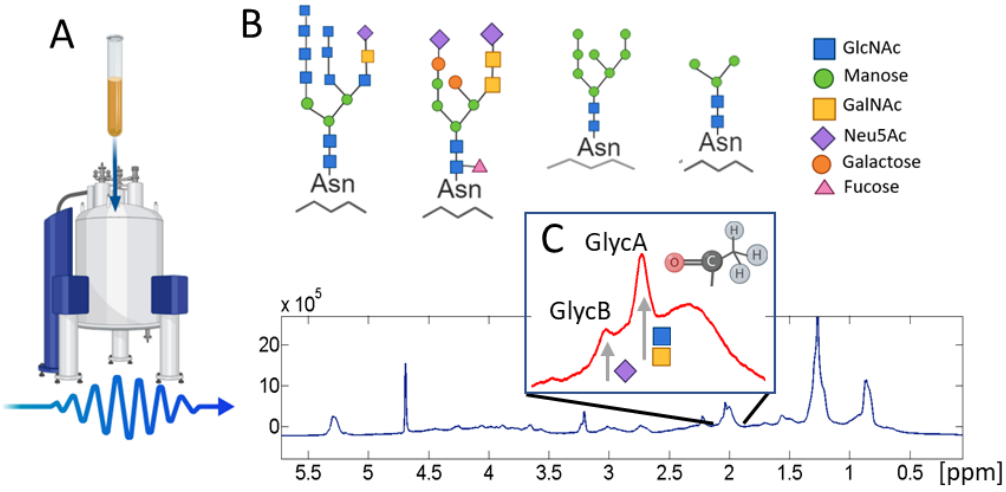


Figure C3.2. ^{1}H-NMR glycoprotein analysis methodology. A, sample tube and spectrometer; B, examples of N-glycans with different residues attached to the protein chain by asparagine (Asn); C, ^{1}H-NMR spectrum produced by the sample in which the region of the glycoproteins is marked. The chemical group producing this signal is indicated.

In the case of glycoproteins, the region is a composite signal with a prominent peak centered at approximately 2.03 ppm of the ^{1}H-NMR spectrum. This signal is produced by the -$COCH_3$ acetyl groups of N-acetylglucosamine and N-acetylgalactosamine and N-acetylneuraminic acid.[51] The resonances of the sugar ring protons of the glycoproteins are not clearly discernible in the plasma spectra

because of extensive overlap with the more intense signals from glucose.[25]

Otvos et al. were the first to call the main peak of this signal "GlycA" (Glycoprotein acetylation) in 2015.[52] Before this, other authors had referred to this same peak as N-acetyl glycoprotein (NAC or NAGs) or N-Acetyl glucosamine (NAG). The low detection sensitivity of NMR means that species present at concentrations less than about 20 μmol/L are undetectable under the conditions of measurement. So, only a small subset of acute phase glycoproteins make meaningful contributions to the GlycA signal.[52] It appears that measured GlycA concentrations are mainly due to contributions from α1-acid glycoprotein, haptoglobin, α1-antitrypsin, α1-antichymotrypsin, and transferrin.[25,52] With the exception of transferrin, the circulating concentrations of the proteins that constitute the GlycA signal increase during the acute-phase response.[48] This heterogeneous composition is a challenge for future research since NMR alone cannot accurately measure the concentration of each of the individual proteins in the signal. However, complementary studies using NMR and other techniques such as immunoassays and even machine learning techniques have shown that α1-acid glycoprotein is the major contributor to the signal, followed by α1-antitrypsin.[25,48,52,53]

During the present decade, several research groups have focused on developing methods to determine ^{1}H-NMR glycoproteins. Some companies that specialize in the analysis of other metabolites, such as lipoproteins, have expanded their services by also offering the analysis of glycoproteins since, as mentioned above, they are obtained

from the same spectrum. A clear example of this is the NMR-algorithm (Lipoprofile®) at LabCorp, Inc. (formerly LipoScience, Inc.), which quantifies GlycA in plasma.[52] Another example is Nightingale Health Ltd (formerly Brainshake, Ltd.), which also quantifies GlycA in serum samples.[54] Also noteworthy is Biosfer Teslab S,L (Liposcale®). This company recently developed a method for analysing glycoproteins by ^{1}H-NMR spectra analysis which obtains parameters other than GlycA such as GlycB (concentration of acetyl groups of N-acetylneuraminic acid) and ratios H/W GlycA and H/W GlycB, which provide information on the function in each case. This function depends on its height, which is related to the concentration, and its width, which is related to the flexibility and the aggregation of the molecules generating the signal. Higher and narrower signal peaks have been related to some inflammatory pathologies.[51] The glycoprotein profiling methods mentioned will be described in the section below.

3.4.2.4. Glycoprotein profiling methods

Considering that in NMR the integrated surface (area) of the absorption signal is proportional to the number of nuclei that pass from the fundamental state to the excited state, integrating isolated signals is the classical approach to estimating the concentration of glycoprotein acetyls. Many authors use this method of integration to quantify the signal.[45,55–63] However, this approach is very sensitive to baseline distortions and it is not recommended for overlapping peaks. The results in many cases are expressed in ‘units of detected protons’ with one unit corresponding to the area of the proton NMR signal detected on 1 mmol/l formate in 1 ml of serum.[45]

Most of the studies reported in the literature, especially since 2015, use the following two methodologies to analyse the ^{1}H-NMR of glycoproteins. This information is reflected in **supporting information table C3.1S.**

3.4.2.4.1. The method by Otvos et al.

The experimental NMR spectroscopy used by Otvos et al. to quantify GlycA is exactly the same as the one used in the numerous clinical studies carried out to assay lipoprotein subclasses and lipids described above. Briefly, they performed a curve fitting method to quantify the particle concentration and mean particle size of various lipoprotein subclasses (*LipoProfile*®)[64] at Liposcience Inc (acquired by Labcorp in 2014). For glycoproteins, they acquired the serum spectra in the same way as NMR LipoProfile® test spectra but in a single block of 8 scans. The GlycA signal was quantified by non-negative linear least-squares deconvolution of the 1.86-2.07 spectral region using proprietary software and the same singular value decomposition computation used for NMR LipoProfile® analysis.[46,65] The deconvolution models include a library of allylic proton reference spectra from 57 isolated lipoprotein subclasses (20 HDL, 9 LDL, 28 VLDL/chylomicron) that were obtained for use in the NMR LipoProfile® deconvolution model, sets of slightly offset narrow Lorentzian signals to model the chemical shift microheterogeneity of glycoprotein GlcNAc methyl signals centered at 2.00 ppm and also at other downfield locations (~2.02-2.05 ppm), a plasma protein reference spectrum background from amino acid residues on albumin and other plasma proteins. Summing the derived amplitudes of the 10

Lorentzian components between 1.99 and 2.01 ppm gives the sample's GlycA signal amplitude (US 2013/0328561 A1 patent). They applied a correction factor of 17.8 µmol/l of glycoprotein N-acetyl methyl group concentration units.[52] They evaluated the contribution of the acute phase proteins α1-antitrypsin, haptoglobin, transferrin, fibrinogen, IgG, α1-acid glycoprotein, α1-antichymotrypsin, and α2-macroglobulin. Their results indicated that from these proteins, IgG, fibrinogen and α2-macroglobulin gave rise to no detectable GlycA NMR signal, while the other five acute-phase glycoproteins appear to have mobile glycan chains that would produce GlycA signals in proportion to their glycan GlcNAc concentrations.[52]

3.4.2.4.2. The method by Ala-Korpela et al.

In 2014, the Finnish company Brainshake (known as Nightingale Health Ltd since 2017) became a serum NMR metabolomics platform that measures more than 200 metabolites, including glycoprotein acetyls, in a highly automated way.[66] The experimental protocols are based on the lipoprotein profile characterization developed by Ala-Korpela, P. Soininen and colleagues in which they set up a curve fitting model by using Lorentzian functions with an in-house algorithm for deconvoluting the signals that mathematically optimized the half-line width, the resonance frequency and the intensity for each Lorentzian function.[67] Their NMR metabolomics platform includes three molecular windows for analysing the NMR spectrum: the LIPO window, the LMWM window, and the LIPID window. The LIPO window uses a Bruker NOESY solvent pre-saturation pulse sequence to analyse the largest

molecules, and lineshape fitting and regression methods to quantify the number of particles in each lipoprotein subclass and their content (mainly cholesterol and triglycerides). The LMWM window uses a Bruker 1D CPMG pulse sequence to analyse low molecular weight metabolites. And the LIPID window uses lipid extraction procedures to provide information about saturated and unsaturated fatty acid families, free and esterified cholesterol, sphingolipids and phosphoglycerides.[8] The peak of the glycoprotein acetyls is seen in the LIPO and LMWM window. For each metabolite a ridge regression model is applied for quantification to overcome the problems of heavily overlapping spectral data.[68] Low-molecular-weight metabolites and lipid extract measures are quantified in mmol/L using regression modelling calibrated against a set of manually fitted metabolite measures. The calibration data is quantified using iterative lineshape fitting analyses and PERCH NMR software (PERCH Solutions Ltd., Kuopio, Finland). Absolute quantification cannot be directly established for the lipid extract measures because of experimental variation in the lipid extraction protocol. [68]

3.4.2.4.3. Other methodologies

In 1987 Bell et al. were the first to assign the broad peaks centred at 2.04 ppm and 2.08 ppm to N-acetyl protons of N-acetylated carbohydrate side-chains associated with 'acute-phase' plasma glycoproteins. They used a Bruker AM500 spectrometer operating at 500 MHz. To estimate the concentration of acute-phase glycoproteins in blood plasma by NMR, standard additions were made of a mixture of α-acid glycoprotein, α-antitrypsin, haptoglobin and transferrin.

Peak intensities were compared by weighing paper traces. They suggested that the concentration of glycoproteins responsible for the two signals (I, 2.04 ppm and II, 2.08 ppm) was 13 mg/ml in the plasma of normal subjects.[25]

De Meyer et al. implemented the adaptive intelligent (AI)-Binning algorithm.[50] Briefly, each point in the NMR spectral range of a specific bin is evaluated as a potential (candidate) new bin edge. This candidate bin edge virtually splits the current bin into two new bins, and the quality of the two new bins is compared with the quality of the current bin. In order to assess this bin quality, a measure known as bin value is defined. The detection of the α-1 acid glycoprotein peaks was more straightforward with AI-Binning than with standard binning. No arbitrary parameters, reference spectra, a priori knowledge or data modifications are required for this protocol.[50] Correia G. et al. used the R package BATMAN (Bayesian AuTomated Metabolite Analyser for NMR spectra) for deconvolution and relative metabolite quantification in the spectra.[69] Lecuyer et al. used another integrative methodology, NMRPipe, to slice each 1D NMR spectrum into buckets of 0.001 ppm, containing NMR signals.[70] The intelligent buckets were scaled to the total summed integrals for each spectrum and then used as the final input for the multivariable models.[71]

Finally, many authors use the entire normalized spectrum to make statistical analyses with SIMCA,[72–74] Orthogonal Projections to Latent Structures (O-PLS),[75–77] Partial Least Squares-discriminant Analysis (PLS-DA),[78,79] or Statistical Total Correlation Spectroscopy[80] (STOQSY).[81]

It should be noted that most of the studies described in this review use the methodology of Otvos (LipoScience-Labcorp) and Ala-Korpela, Soininen et al. (Brainshake-Nightingale). Both methodologies provide the parameter GlycA as the absolute N-acetyl group concentration. As far as we know, no comparative study of the two methods has been published. However, to our knowledge, the first study in which GlycA data analysed with Vantera, a clinical NMR analyzer designed for clinical use, have been reported is the PREVENT study. A high correlation (a coefficient of 0.983) was reported between the Vantera analyser from LiposScience (Labcorp) and Brainshake (Nightingale Health Ltd) GlycA analysis methods.[82] The studies in which other methodologies are used are not comparable with these two methods. Therefore, the N-acetyl NMR information needs to be unified in a single inflammatory marker such as GlycA. It would also be interesting to perform a comparative study of the results of the methodologies mentioned to quantify the N-acetyl groups of glycoproteins so that the studies can be comparable between them.

3.5. ^{1}H-NMR glycoproteins clinical studies

Although we have seen that several changes in the glycation of proteins give rise to disease and that they can be measured by several techniques, in this review we focus on the most recent conclusions in the literature on glycoprotein studies, particularly those drawn from ^{1}H-NMR studies.

CHAPTER 3

A broad search of the literature was performed using Elsevier's scientific database SCOPUS, PubMed, and the Google scholar database. SCOPUS (with cut-off dates between 1999 and 2019) was the main database searched and the following search terms were used: "Glycoproteins OR GlycA OR Glyc A" AND "NMR OR nuclear magnetic resonance OR H-NMR" AND "serum OR plasma" AND "marker OR biomarker". The search was completed with manual reference checks on Google Scholar and the PubMed database with the same keywords. The initial result of the search was 239 documents that were subsequently filtered by species (humans), English language and fields of no interest such as agricultural sciences. Finally, we discarded the articles that spoke of specific glycoproteins not detected by ^{1}H-NMR and articles that included NMR but only referred to lipids. As shown in the trend graph (**Figure C3.3**), the detection of glycoproteins by ^{1}H-NMR has been a topic of great interest in recent years and the number of studies is progressively increasing.

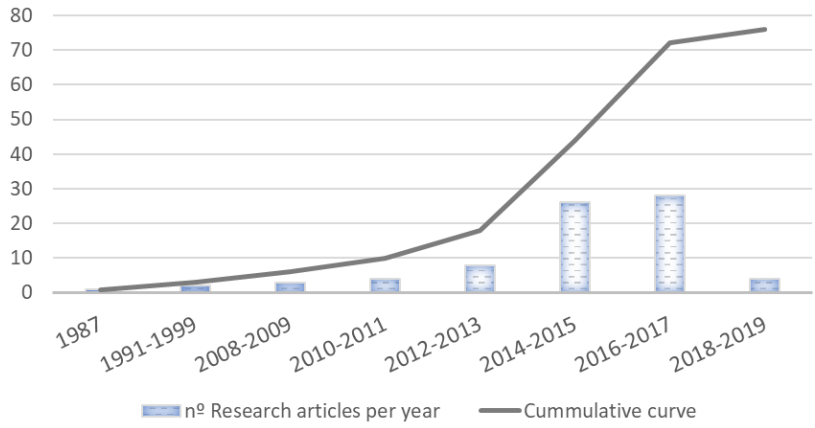


Figure C3.3: Trend graph of the number of research articles per year in recent years.

3.5.1. Former studies on ^{1}H-NMR detection and identification of glycoproteins

Prior to 1990, the literature search offered little clinical data on the detection of glycoproteins or glycans by ^{1}H-NMR. However, some early works in 1983 and 1984 by Nicholson and colleagues reported assignments to peaks in the ^{1}H-NMR spectrum[83,84] and mentioned the assignment of the spectrum peaks for N-acetyls from sugars of glycoproteins.[83] The first studies on the characterization of glycoproteins were carried out by Bell et al. in 1987 and determined the N-acetyl protons of highly mobile N-acetylated carbohydrate side-chains associated with plasma glycoproteins (mainly α1-acid glycoprotein, α1-antitrypsin, haptoglobin, transferrin and immunoglobulins).[25] In 1999, when even the signals of the N-acetyl groups in the ^{1}H-NMR spectrum were not unequivocal, one of the first studies compared the glycoprotein signals with the levels of immunoglobulins detected by other biochemical methods.[45] It was then theorized that both analytical approaches should be used with different strategies because NMR parameters were more suitable for longitudinal studies of chronic situations.[45] These previous studies suggested that glycoproteins of acute-phase reactants, which reflect both acute and chronic inflammation, may be useful for the detection, prognosis, and therapeutic monitoring of tissue damage marked by inflammation in several pathologies.

3.5.2. Clinical applications

^{1}H-NMR has been used to characterize glycoproteins in several pathologies since the first studies mentioned above. A summary can be found in **table C3.2.** A more detailed explanation of each of the studies in this section is reflected in **table C3.1S** of **supporting information.**

Table C3.2: Summary of ^{1}H-NMR glycoprotein's clinical applications

Clinical study topic		Main findings	References
Tumours and cancer	OC, CSCC, BC, LC, CRC	Increased circulating N-acetyl glycoproteins levels and increased GlcNAc-branching of N-glycans.	25,57,60,72,73,79
Metabolic diseases	Obesity	Association between GlycA and the leptin/adiponectin ratio	94
		Correlation between GlycA and TG and lipids	96
		Correlation GlycA and branched chain amino acids	97
		Strong relationship of CRP, GlycA, and GlycB and insulin resistance	99
	Diabetes Mellitus	α1-acid glycoprotein as a predictor of future glycemia	110
		Associations of GlycA with higher IL-6 and CRP	112
		Associations of GlycA with future T2DM	82,111
		GlycA had a more robust correlation with CRP, plasma glucose, and measures of adiposity and insulin resistance than GlycB	99,113
	MetS	Increased levels of glycosylated acute-phase proteins (GlycA) associated with MetS	94,115,209
Cardiovascular risk	Healthy individuals	GlycA/alpha1-acid glycoproteins or baseline circulating glycoprotein N-acetyl methyl groups are associated with CVD and longitudinal risk of all-cause mortality.	88,89,111,123,124,129,163
	High-risk individuals	GlycA and GlycB strongly associated with future major adverse CVE	129
		GlycA and hsCRP was statistically significant for the outcome of death	130,131
		GlycA, and small and medium-size HDL particles proved to be independent predictors of cardiac death.	132,133

	Life expectancy	Higher GlycA levels had lower life expectancy.	134
	All-cause mortality	Positive association between α1-antitrypsin and increased risk of liver diseases, heart failure and COPD, and significant association between α1-acid glycoprotein and heart failure and chronic lower respiratory diseases	53
		GlycA related to increased risk of alcoholic liver disease, chronic renal failure, glomerular diseases, COPD, inflammatory polyarthropathies and hypertension	151
HIV-infection		Higher GlycA levels in HIV-infected patients	137
Chronic inflammatory diseases	RA	GlycA is higher in RA patients than in controls.	51,139,140
	SLE	GlycA levels increased with each unit increase in SELDAI.	142,143
		GlycA has been shown to be a good marker of systemic inflammation in lupus-nephritis.	141
	Psoriasis	GlycA is increased in psoriasis.	146
	IBD	GlycA in populations with ulcerative colitis and Chron's disease better reflects inflammatory status than other classical markers.	149
	CKD	GlycA was independently associated with albuminuria and inversely related to eGFR.	157
	CHC	Increased severity of fibrosis has been associated with higher NAC plasma levels.	76
Cognitive function and psychological health	Global cognitive function	GlycA is inversely related to global cognition, information processing speed and memory domains.	158
	AD	Elevated circulating glycoproteins were associated with the risk for AD and MCI.	159
Rare vascular diseases	Takayasu arteritis	N-Acetyl glycoproteins are significantly up-regulated in TA patients.	78,161
	Kawasaki disease	High levels of GlycA were confirmed in paediatric population with acute KD disease.	163

Pregnancy		Gradual increase in N-acetyl glycoproteins during pregnancy.	81
Pregnancy		Multiple nutrient intake correlates with GlycA including fibre, LC-PUFA and w-3 LC-PUFA and several vitamins and minerals.	54
Primary aldosteronism		GlycA and hsCRP were statistically significantly higher in obese than in overweight pregnant women.	164
		GlycA levels significantly increased in PA population.	167
Sickle cell disease		GlycA levels are decreased in SCL.	169
Human African Trypanosomiasis		Significant increase of N-acetyl glycoprotein in HAT patients.	170
Sodium intake		Lower GlycA and hsCRP concentrations were both associated with higher 24-h sodium excretion.	116
Tobacco smoking		Similar significant associations between different measures of smoking behaviour and higher GlycA and hsCRP levels.	171
Effect of exercise		Regular exercise significantly reduced plasma GlycA.	68,100,172
Effect of treatments	Anti-TNF and monoclonal antibodies	Decrease in GlycA levels.	141,146
Effect of treatments	Antirretroviral treatment	GlycA was the only marker of inflammation, among hsCRP, IL-6 and D-dimer, that decreased.	138
Toxicity	Statins	Do not affect GlycA levels	133

	Metformin	Lower NAC serum levels in T2DM patients treated in metformin than in untreated patients.	[56]
	Probiotics	Greater gut microbiota richness is negatively linked with low-grade inflammation marker GlycA.	[54]
	Sodium valproate	N-acetyl moieties of glycoprotein significantly increased ($p<0.01$) in valproate sodium induced hepatotoxicity	[74]
Toxicity	Oncologic toxicity	The high acute radiation sequelae were associated with increased signals of N-acetyl glycoproteins	[63]

Abbreviations: OC,Ovarian cancer; CSCC, Cervical squamous cell carcinoma; BC, Breast cancer; LC, lung cancer; CRC, colorectal cancer; GlcNAc, N-Acetyl glucosamine; TG, total triglycerides; CRP, C-reactive protein; IL-6, interleukin-6; T2DM, type 2 diabetes mellitus; MetS, metabolic syndrome; CVD, cardiovascular disease; CVE, cardiovascular event; COPD, chronic obstructive pulmonary disease; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; SELDAI, Systemic Lupus Erythematosus Disease Activity Index; IBD, Inflammatory bowel disease; CKD, chronic kidney disease; CHC, chronic hepatitis C; AD, Alzheimer disease; MCI, mild cognitive impairment; TA, Takayasu arteritis; KD, Kawasaki disease; PUFA, polyunsaturated fatty acids; PA, primary aldosterism; SCL, sickle cell disease.

3.5.2.1. Tumours and cancer

Cancer is the second leading cause of death globally and is estimated to have accounted for 9.6 million deaths in 2018. Biomarker research in this field has been becoming increasingly important in recent years for the diagnosis and prognosis of different kinds of tumour. Evidence in the literature shows that chronic inflammation is a potential factor associated with tumour development.[73] Alterations in glycosylation patterns regulate the development and progression of cancer, potentially serve as important biomarkers and provide a set of specific targets for therapeutic interventions.[21] Changes in glycosylation commonly associated with cancer include sialylation, fucosylation, increased GlcNAc-branching of N-glycans, over-expression of truncated mucin type O-glycans and increased circulating N-acetyl glycoprotein levels.[25] These changes increase structural glycan heterogeneity and alter the function of cells.[85]

It was not until 1988 and 1989 that ^{1}H-NMR began to be used as a new tool for studying the structural and metabolic modifications in cancer patients.[55] In recent years, NMR-based body fluid metabolomic studies have been increasingly performed for the diagnosis and prognosis of the disease.[86]

Breast cancer (BC) is the most frequent cancer among women. It affects 2.1 million women every year and causes the highest number of cancer-related deaths (WHO, 2018). In BC research so far, metabolomics has been generally used for the direct characterization of tumour metabolism alterations.[72] ^{1}H-NMR has been used to

demonstrate that serum metabolite profiles derived from metastatic breast cancer (MBC) patients are different from localized early breast cancer patients (EBC). Compared to EBC patients, MBC patients display higher serum concentrations of N-acetyl glycoproteins (NAC1 $p<0.027$ and NAC2 $p<0.007$),.[72] Suman et al. reinforced these results when they found high levels of NAG among other metabolites such as hydroxybutyrate, lysine, glutamate, glucose, and lactate metabolites in BC patients, which were potentially useful for diagnosing BC progression.[79] However, contrary to these results, in an NMR-based untargeted metabolomic study, Lecuyer et al. reported lower plasma levels of glycoproteins, lipoproteins, lipids, acetone, glycerol-derived compounds, unsaturated lipids and a higher risk of developing breast cancer within the following decade.[71]

Two ^{1}H-NMR studies have been carried out in the cystic fluid of ovarian cancer[57,87] but only one of them focuses on the N-acetyl groups of glycoproteins and shows that they are positively associated with the pathology.[57] Studies on cervical squamous cell carcinoma (CSCC) have also shown an increase in the glycoprotein peak of the ^{1}H-NMR plasma spectrum, as well as the peak of other metabolites, compared with patients with cervical intraepithelial neoplasia (CIN).[61]

Like BC, lung cancer is also one of the most common diseases worldwide (2.09 million cases according to WHO data for 2018). One of the negative points of this cancer is the asymptomatology of the early stages. Therefore, the search for predictive markers that make it possible to identify the presence of cancer is important for a good

prognosis. Chronic obstructive pulmonary disease (COPD) is a predisposing factor for this type of cancer. Deja et al. reported an increase in N-acetylated glycoproteins in patients with multi-stage lung cancer compared to COPD patients. They also pointed out that the N-acetylated glycoprotein signal is a useful marker for distinguishing between different stages of lung cancer.[73]

Chandler et al. also showed a positive association between GlycA and incident colorectal cancer (CRC) and mortality by evaluating the baseline measurements of GlycA in two large cohorts – the Women's health study (WHS) and the Multi-Ethnic Study of Atherosclerosis (MESA) – with a median follow-up period of 19 and 11 years, respectively.[88]

Duprez et al. also pointed out that among other inflammatory markers such as hsCRP, IL-6 and D-dimer, GlycA was the only one that was independently predictive of future cancer. It is curious to note the difference in importance in different ethnicities; GlycA did not predict total cancer in whites, but was a strong predictor in blacks, Chinese and Hispanics.[89]

However, other studies have shown a different trend: a relationship between low levels of NAG and the risk of developing hepatocellular carcinoma (HCC)[75] and urothelial carcinoma (UTUC) in patients compared with healthy controls.[62] In addition, another ^{1}H-NMR metabolomic study in glioma, the most common of all primary central nervous system tumours, showed a lower level of glycoproteins, among other metabolites.[90]

3.5.2.2. Metabolic disorders

Obesity and diabetes are some of the most common metabolic disorders in which ^{1}H-NMR glycoproteins have been studied. Both disorders are closely related to each other and to other pathologies such as cardiovascular diseases (CVDs) or metabolic syndrome (MetS). Systemic inflammation is hypothesized as a central mechanism.

3.5.2.2.1. Obesity

A higher body mass index (BMI) is characterized by higher leptin concentrations and decreased anti-inflammatory adiponectin levels. These parameters are determined with the leptin/adiponectin ratio, which is elevated in obesity and a marker of the adipose tissue production of pro-inflammatory cytokines and of insulin resistance in nondiabetic individuals.[91,92] An important result in terms of ^{1}H-NMR glycoproteins and obesity is the considerable association found between GlycA and the leptin/adiponectin ratio,[93,94] which suggests that GlycA is a marker of adipose tissue-associated low-grade inflammation.

Furthermore, GlycA has been shown to correlate with higher concentrations of triglycerides and other lipid levels, such as LDL cholesterol, in obese non-pregnant subjects[95] and in obese and overweight pregnant women.[96] The level of GlycA has also been shown to correlate with the amount of branched chain amino acids,[95] which along with aromatic amino acids are increased in obesity and insulin resistance,[97] and also in type 2 diabetes.[98] Lorenzo et al. also

found a strong relationship between C-reactive protein (CRP), GlycA and GlycB, and measures of insulin resistance and adiposity. Furthermore, GlycB has weaker relationships with CRP and measures of insulin resistance and adiposity than GlycA.[99]

Strategies to reduce obesity such as an exercise-based lifestyle, diet and even bariatric surgery have been the subject of countless studies. One way to establish whether ^{1}H-NMR glycoproteins are directly related to poorer health in obese patients is to study whether they vary when these strategies are carried out. Barlett et al. evaluated how an exercise-based lifestyle or exercise plus diet interventions for 6 months modulate GlycA in sedentary adults with prediabetes. Their results showed a reduction in GlycA levels, which they associated with a decrease in visceral adiposity.[100] GlycA has also been measured in obese patients undergoing bariatric surgery, which is the most effective therapy in cases of severe obesity. The most typical forms of bariatric surgery are Roux-en-Y gastric bypass and sleeve gastrectomy, both of which lead to substantial weight loss. Manmadhan et al. saw that postoperative changes in GlycA were very positively associated with changes in body weight, high sensitive CRP (hsCRP) and glycated haemoglobin (HbA1c), and inversely with changes in the mean particle size of HDL and adiponectin.[101]

Few studies have been conducted in children or adolescents. One of them, a cohort of 1664 US adolescents from the HEALTHY study, showed that high GlycA values were associated with higher BMI and more related to girls than to boys, which can be explained by

the progression of puberty.[102] A second study in an adolescent population showed a significant reduction in GlycA after 12 weeks of lifestyle intervention consisting of weekly nutrition and health classes delivered by promoters (bilingual, bicultural health educators) and 3 days per week of moderate to vigorous physical activity. Decreases in GlycA were associated with decreases in 2-hour glucose ($p<0.008$) and BMI ($p<0.03$).[103]

3.5.2.2.2. Diabetes Mellitus

Diabetes is a chronic disease that occurs either when the pancreas does not produce enough insulin (type 1 diabetes) or when the body cannot effectively use the insulin it produces (type 2 diabetes).[104] Low-grade systemic inflammation has been associated with the risk of diabetes.[105,106]

The association of circulating levels of inflammatory proteins, in particular APP, in type 2 diabetes (T2DM) is well described in prospective epidemiological studies.[107–109] One of the largest studies was conducted in 2012 by Würtz et al, who investigated the associations of circulating metabolites with fasting and post-loading glycemia before disease onset. They pointed out that α1-acid glycoprotein is a predictor of future glycemia, and underlined the importance of prolonged inflammation as a risk marker for attenuating glucose tolerance.[110] In a set of 26,508 women enrolled in the WHS, Akinkuolie et al. showed the potential role of glycans in the risk of T2DM by showing that several APP was associated with the risk of developing T2DM[111] Connelly et al. confirmed the findings of

Akinkuolie et al.[111] by showing that in 4,525 participants of the Prevention of Renal and Vascular End-stage Disease (PREVEND) study, GlycA was an independent predictor of T2DM even after adjusting for traditional diabetes risk factors and hsCRP.[82] Moreover, in PREVEND the associations of GlycA with future T2DM were similar for men and women while the hsCRP associations appeared to be stronger in women than in men. Another study on patients with T2DM hospitalized for diseases such as congestive heart failure (CHF), cardiac non-CHF, infection, and other noncritical diseases showed that there were differences in inflammatory markers across disease states. GlycA was associated with higher IL-6 and CRP, with values being highest in T2DM patients with infectious diseases.[112]

Lorenzo et al. studied the relation of GlycA, GlycB, and CRP with direct measures of insulin sensitivity. They found that CRP levels and GlycA were higher in T2DM than in isolated impaired glucose tolerance, but GlycB was not increased. They also found that GlycA had a more robust correlation with CRP, plasma glucose and measures of adiposity and insulin resistance than GlycB.[99] In line with these results from Lorenzo et al, Filezova et al. also found a positive association between GlycA and impaired insulin secretion in a population of 5,401 non-diabetic men from the prospective Metabolic Syndrome in Men study (METSIM), in addition to the above-mentioned results by other authors with hyperglycemia, incident T2DM and CVD.[113]

In contrast to these results, some authors have found inverse relationships. Rawat et al. found that residual signals of N-acetyl glycoproteins were lower in diabetes patients with inadequate glycaemic control and with diabetic neuropathy, nephropathy, and cardiovascular disease than in healthy controls.[114] They attribute this to possible greater oxidative damage, but more studies are needed to find a solid explanation. Likewise, Gruppen et al. found in one study that GlycA was not positively associated with T2DM.[115] However, in a similar study, they subsequently verified that GlycA was higher in subjects with either T2DM, MetS or both.[116]

3.5.2.2.3. Metabolic Syndrome (MetS)

MetS is a clustering of cardiovascular risk factors with insulin resistance as a major feature. This syndrome has been defined in various ways, but generally consists of three or more of the following components: hyperglycemia, hypertension, hypertriglyceridemia, low HDL, and increased abdominal circumference and/or BMI at >30 kg/m2.[117]

Pro-inflammatory markers such as white blood cell counts and plasma levels of coagulation factors (fibrinogen and plasminogen activator inhibitor 1), APP such as CRP and serum amyloid A (SAA), pro-inflammatory cytokines (tumour necrosis factor (TNF)-a, IL-1b and IL-6), and chemokines are positively correlated with insulin resistance and the features of the metabolic syndrome in most cases.[118] Little information is available on the relationships of MetS with other pro-inflammatory biomarkers but increased levels of glycosylated APP

(GlycA) have been associated with MetS and negatively related to bilirubin levels.[93] This may represent a quantitative measure of a pro-inflammatory state.

In line with these results, Gruppen et al. found a higher concentration of GlycA in MetS subjects and a positive correlation of GlycA with cholesterol acetyl transferase (LCAT), systolic blood pressure, BMI, waist circumference and plasma triglycerides. The correlation with HDL cholesterol was inverse.[116,119]

3.5.2.3. CVD risk and all-cause mortality prediction

In the clinical field, the hsCRP marker has been strongly associated with CVD,[120,121] as have other widely used markers such as IL-6, which has been more strongly related to all-cause death than CRP.[122] Interestingly, in the last few years, several studies have pointed to ^{1}H-NMR glycoproteins, specifically GlycA, as better predictors of the risk of CVD and all-cause mortality in both the general population and the population at high risk of CVD.[82,88,89,111,123,124]

3.5.2.3.1. Apparently healthy population

GlycA has been measured in apparently healthy individuals from large population studies such as WHS, 'Justification for the Use of Statins in Primary Prevention: An Intervention Trial Evaluating Rosuvastatin' (JUPITER) trial, and MESA. In all of them elevated baseline circulating glycoprotein N-acetyl methyl groups were associated with longitudinal risk of CVD incidence and mortality, among other pathologies such as chronic inflammatory-related severe

hospitalization, cancer or death.[89,124–126] Considering the results, the authors defend GlycA's role as an important predictor of 5 to 15-year risk of CVD.[123,124] In these studies, the predictive capacity of GlycA was independent of age, sex, modifiable lifestyle risk factors, medication, and disease prevalence.

A large study conducted by Ritchie et al. in 2015 focused on GlycA levels in a population of 11,825 individuals from three different cohorts; i) the Dietary, Lifestyle, and Genetic Determinants of Obesity and Metabolic syndrome (DILGOM) study; ii) a large Finnish population survey on risk factors on chronic, noncommunicable diseases (FINRISK); iii) and the Cardiovascular Risk in Young Finns Study (YFS). The results suggested that in apparently healthy individuals, GlycA may be chronically elevated for periods of up to a decade; in individuals with elevated GlycA, cytokines also increased slightly, which suggested a prolonged low-grade inflammatory state. GlycA strongly predicted the future risk of hospitalization and death from infection.[48]

Benson et al. wanted to take one step further and, in 6,479 MESA participants, demonstrated that lower levels of GlycA were related to cardiovascular health (CVH) as defined by Life's simple 7 score (LS7), which includes seven individual health metrics (smoking, physical activity, BMI, diet, total cholesterol, blood pressure, and blood glucose). LS7 scores were categorized into CVH groups classified as “optimal” (12-14), “average” (8-11) or “inadequate” (0-

7). GlycA was independently inversely associated with continuous LS7 scores.[126]

3.5.2.3.2 Prediction of CVD in high-risk populations

Recently, glycoprotein acetyls have been strongly related with myocardial infarction (MI), ischemic stroke (IS) and intracerebral haemorrhage (ICH).[127] GlycA and its role in improving the prediction of CVD risk in high-risk populations has also been described.[128]

Some studies carried out on patients from the angiography registry of the Intermountain Heart Study demonstrated that baseline levels of both GlycA and GlycB were strongly associated with future major adverse cardiovascular events.[129] In addition, GlycA proved to be an independent predictor of cardiac death.[130] Muhlestein et al. studied 2,996 patients undergoing angiography for coronary artery disease (CAD) from the same trial, and determined that the interaction between GlycA and hsCRP was statistically significant for the outcome of death.[131]

On another note, Correia et al. evaluated changes in key metabolites in 28 children undergoing surgery for congenital heart. A relationship was found between inflammation and metabolic derangement in the days after surgery for congenital heart disease with an increase in N-acetylated glycoprotein fragments immediately post-surgery.[69]

3.5.2.3.3. Life expectancy prediction and all-cause mortality

In line with these results, McGarrah et al. studied 7,617 cardiac catheterization patients from the CATHGEN biorepository. They also found a strong association of GlycA with mortality, CAD and all-cause mortality, cardiovascular and non-cardiovascular. They noted that the individuals at highest risk of dying were those who had diabetes and a higher concentration of GlycA, followed by those without diabetes but with high concentrations of GlycA, and finally individuals with diabetes but lower concentrations of GlycA. On the other hand, the GlycA and smaller HDL subclasses had independent but opposite effects on mortality risk prediction, with smaller HDL subclasses being protective.[132]

A sub-study analysis of the trial Atherothrombosis Intervention in Metabolic Syndrome with Low HDL/High Triglycerides and Impact on Global Health Outcomes (AIM-HIGH) conducted by Otvos et al. confirmed and extended the previous findings. GlycA predicted CVD events and mortality in high-risk patients with established CVD who had achieved very low LDL-C levels. All-cause mortality was significantly associated with both GlycA and low levels of small HDL particles.[133]

The remaining life expectancy is an increasingly used measure of survival. Gruppen et al. determined that men and women from the PREVEND study with higher GlycA levels had a lower life expectancy, while high levels of hsCRP were only related to a lower life expectancy in men.[134]

3.5.2.4. Human Immunodeficiency Virus (HIV)-infection

As far as we know, there is only one study in the literature that relates GlycA to HIV infection. Tibuakuu et al. investigated the association of GlycA with CVD risk in HIV infection, since in these patients the risk of CVD is higher than in people who are not infected.[135,136] They showed that GlycA levels were higher in men who were HIV-infected than in those who were not, and higher in men with detectable versus undetectable viral load. In HIV men with plaque, GlycA was positively associated with the extent of coronary artery calcium and total plaque.[137] Further research is needed to see if GlycA levels are predictive of incident CVD events in HIV-infected individuals.

Another study measured GlycA in HIV patients, but it focused on observing the effects of antiretroviral treatment[138] and will be discussed below.

3.5.2.5. Chronic inflammatory diseases related to immune system

3.5.2.5.1. Rheumatoid Arthritis (RA)

RA is a chronic inflammatory disease associated with the development of CVD. GlycA may be a useful marker of disease activity and CVD risk in patients with RA. It has been shown that GlycA is higher in RA patients than in controls[51,139,140] and strongly correlated with all components of DAS28 scores, the measure of disease activity in RA.[140] Fuertes-Martin et al. have pointed out that the parameters H/W GlycA and GlycB ratios are significantly higher

in the RA population than in controls and have hypothesized that the high, narrow shape of the peaks could be an additional marker of systemic inflammation.[51]

Moreover, GlycA has been associated with the presence of coronary artery calcium and prevalent coronary artery disease in patients with RA[140] even though GlycA is predominantly associated with typical systemic inflammation and less with adiposity.[139]

Bearing in mind that some traditional markers used to evaluate RA, such as CRP and ESR, are nonspecific because their concentrations are also increased in other chronic inflammatory diseases, GlycA could be an additional measure of inflammation that will lead to greater accuracy than if only the classic inflammatory parameters are considered, as is conventional in clinical practice.[51,139]

3.5.2.5.2. Systemic Lupus erythematosus (SLE)

Some studies have shown a high association between increased GlycA levels and SLE.[141–143] Two in particular have demonstrated that GlycA levels increased with each unit increase in the Systemic Lupus Erythematosus Disease Activity Index (SELDAI).[142,143]

In addition, GlycA has been shown to be a good marker of systemic inflammation in lupus-nephritis, one of the most severe complications of SLE. Unlike GlycA levels, the CRP concentrations of non-lupus nephritic controls were not significantly different from those of patients with active SLE. It is important to stress the role of GlycA

and BMI in predicting proliferative status over classical inflammation markers.[141]

3.5.2.5.3. Psoriasis

Psoriasis is a chronic inflammatory skin condition associated with chronic systemic inflammation, increased vascular inflammation and a greater risk of incident CV events and CV mortality.[144,145] GlycA was seen to be increased in psoriasis patients and remained significant after adjustment for age, sex, BMI and traditional CV risk factors. Interestingly, treatment of psoriasis with anti-TNF therapy led to a decrease in GlycA levels and vascular inflammation.[146]

3.5.2.5.4. Inflammatory Bowel Disease (IBD)

IBD is a global disease that is increasingly prevalent on all the continents.[147] It is characterized by chronic relapsing intestinal inflammation.[148] Dierckx et al. measured GlycA in populations with ulcerative colitis (UC) and Chron's disease (CD). GlycA reflected the inflammatory status of patients versus controls and the decrease in inflammation in response to treatment better than other classical markers such as PCR and fecal calprotectin (fcal).[149]

3.5.2.6. Other chronic inflammatory diseases

Chronic obstructive pulmonary disease (COPD) is a heterogeneous condition with patients displaying varying clinical and pathophysiological features. The mechanisms and mediators underlying COPD and its comorbidities are poorly understood. However, there is compelling evidence to suggest that increased

oxidative stress and lung inflammation play an important role in its pathophysiology.[150]

In an 8-year follow-up study Ritchie et al. found a significant positive association between GlycA's constituent glycoprotein α1-antitrypsin and increased risk of liver diseases, heart failure and COPD and a significant association between GlycA's constituent glycoprotein α1-acid glycoprotein and heart failure and chronic lower respiratory diseases.[53] Kettunen et al. made a systematic evaluation of GlycA as a reproducible biomarker for disease prediction in a population of 11,861 adults. Their results were the same as some of those mentioned so far and also demonstrated new strong and consistent associations between elevated GlycA and increased risk of alcoholic liver disease, chronic renal failure, glomerular diseases, COPD, inflammatory polyarthropathies and hypertension.[151]

As a curiosity the relation is mentioned in two traditional Chinese medicine studies that detected lower plasma glycoprotein concentrations in patients with COPD and abnormal Savda syndrome than in controls,[152] the Savda syndrome being a set of psychological and emotional stressors.[153] More rigorous studies and Western medicine are needed to establish a scientific hypothesis on the role of ^{1}H-NMR glycoproteins in COPD.

Another chronic low-grade inflammation disease is chronic kidney disease (CKD), characterized by a reduced estimated glomerular filtration rate (eGFR) and/or albuminuria. However, some epidemiological studies have reported contradictory data that support

a relationship between CKD and an increase in CRP,[154] and others have not found any such association with CRP but with other inflammatory markers such as TNF-α and IL-6.[155,156]

Titan et al. investigated the association of GlycA to albuminuria and eGFR in 5,050 middle-aged men and women from the ELSA-Brasil Study. They showed that GlycA was independently associated with albuminuria and inversely related to eGFR. They also showed that GlycA was better than hsCRP at diagnosing albuminuria, which suggests that glycation has an important role in the progression of CKD and in risk assessment.[157]

Finally, Chronic hepatitis C (CHC) has been most widely studied in the context of non-invasive biomarkers. Increased severity of fibrosis has been associated with higher NAC plasma levels.[76]

3.5.2.7. Cognitive function and psychological health

Psychological suboptimal health is a prevalent state with a pathophysiological mechanism that is extremely complicated and poorly understood but inflammation is known to be related. CRP and IL-6 have been associated with cognition, but few studies have measured inflammatory markers as predictors of cognitive function in middle age or of the onset of cognitive complications such as dementia.[158] However, a higher level of N-acetyl-glycoproteins in patients with psychological suboptimal health has been reported.[77] Another study showed an inverse relationship between GlycA and global cognition and also between information processing speed and memory domains.[158]

One of the most studied cognitive diseases with the most mysterious aetiology is Alzheimer's disease (AD). The need for early diagnosis is a growing issue today. MetS and elevated circulating glycoproteins were also associated to risk of AD and mild cognitive impairment (MCI).[159] These results suggest that bringing these factors together would be more conducive to developing AD but further research is needed.

The lack of biomarkers of inflammation in this field and the results discussed above make the search for prediction and prevention biomarkers an increasingly attractive field.

3.5.2.8. Rare vascular diseases

3.5.2.8.1. Takayasu arteritis

Takayasu arteritis (TA) is a rare, idiopathic systemic inflammatory disease affecting large arteries, including the aorta, its major branches and the pulmonary arteries. Arterial inflammation is the core feature of the disease, variably associated with a systemic acute-phase response.[160] Novel biomarkers are required to distinguish inflammatory and non-inflammatory remodelling in Takayasu arteritis. It has been shown that N-Acetyl glycoproteins are significantly up-regulated in TA patients.[161]

A larger study, conducted by Jain et al. in another cohort, confirmed these results and confirmed the role of glutamate and NAG as potentially strong biomarkers of TA.[78]

3.5.2.8.2. Kawasaki disease

Kawasaki disease (KD) is a self-limited vasculitis that typically presents in young children as an acute illness with fever and mucocutaneous changes. It can develop coronary artery aneurysms, and predispose to serious long-term cardiovascular complications.[162] In a paediatric population with acute KD disease, high levels of GlycA were confirmed. It was also found that GlycA and lipoproteins, both measured with ^{1}H-NMR, may be useful for distinguishing acute KD from bacterial or viral illnesses[163] but further research is needed.

3.5.2.9. Pregnancy

Strategies increasingly focus on tracking metabolic changes during pregnancy in order to determine metabolic profiles that may be associated with prenatal disorders. ^{1}H-NMR has been used to discover metabolic biomarkers that personalize the monitoring of the pregnancy[81] but only a few studies have included ^{1}H-NMR glycoproteins.

An untargeted ^{1}H-NMR study of maternal blood plasma has shown a gradual increase in N-acetyl glycoproteins and a direct link between them and LDLc+VLDLc.[81] Another study conducted in overweight pregnant women evaluated the association between intake of dietary nutrients and markers of low-grade inflammation. Multiple nutrients correlated with GlycA – including fibre, LC-PUFA and ω-3 LC-PUFA and several vitamins and minerals – but no correlations were detected between any of the nutrients and hsCRP and

lipopolysaccharide (LPS).[54] Similarly, increased richness in intestinal microbiota was negatively related to GlycA, but no similar relationship was observed between hsCRP and microbiota-rich gut,[54] which suggests that GlycA may have inflammatory pathways different from those of CRP. The same cohort was used to study whether intestinal permeability was related to metabolic risk markers. The results showed that serum zonulin, a protein responsible for regulating paracellular transport in the intestine, was associated with GlycA, among other markers.[96]

Houttu et al. studied whether there were differences in the inflammatory profile of overweight pregnant women and obese pregnant women. In line with what has been mentioned so far, low-grade inflammatory markers, GlycA and hsCRP were statistically significantly higher in obese pregnant women than in overweight pregnant women. The correlation coefficients were also higher between GlycA and lipids than between hsCRP and lipids. Houttu et al also found a correlation between GlycA and branched chain aromatic amino acids in pregnant women, coincidental with higher insulin and glucose concentrations during early pregnancy.[164]

3.5.2.10. Primary aldosteronism (PA)

PA is characterized by the autonomous production of aldosterone, which causes sodium retention, plasma renin suppression, endocrine hypertension and cardiovascular damage, among other things.[165]Aldosterone is associated with key functions in the regulation of blood pressure, but has also been associated with causing

inflammation, fibrosis and blood vessel remodeling.[166] Berends et al. found that GlycA levels were significantly higher in a PA population than in normotensive control subjects and subjects with treated and untreated hypertension, which indicated enhanced low grade chronic inflammation.[167]

3.5.2.11. Sickle cell disease (SCD)

SCD is a multisystem disorder with multiple organ damage associated with recurring inflammation caused by tissue ischemia, reperfusion injury and vascular damage.[168] Although inflammatory markers used in clinics such as interleukins, prostaglandin-E2, Tumour Necrosis Factor-α and CRP are increased in SCD, GlycA is not. This result goes against what is reflected in this review and has been attributed to the fact that haemolysis is observed in SCD patients but not in patients with other pathologies. It should also be noted that haptoglobin is one of the major proteins in the GlycA signal and in this case it depletes rapidly during intravascular haemolysis.[169]

3.5.2.12. Human African Trypanosomiasis (HAT)

Very few studies on metabolomic profiling use ^{1}H-NMR to identify a metabolic signature of a specific parasitic infection. Marked differences have been shown in plasma HAT patients who have a significant increase of creatinine, N-acetyl glycoprotein ($p<0.01$), formate and myoinositol compared to controls.[170]

3.5.2.13. Sodium intake

High sodium intake has been linked to major health issues such as CVD and hypertension. In a large cohort of predominantly healthy men and women with age- and sex adjusted analyses that took into account BMI, lower GlycA and hsCRP concentrations were both associated with higher 24-h sodium excretion, and these relations remained present after other potential covariates were taken into account.[116]

3.5.2.14. Tobacco smoking

Tobacco smoking is one of the major preventable causes of death and cardiovascular disease (CVD) in the world. Higher levels of some inflammatory markers are indicators of exposure to smoking. Kianoush et al. studied the association between smoking and systemic inflammation (GlycA) in 11,509 participants from MESA and 'The Brazilian Longitudinal Study of Adult Health' (ELSA-Brasil) cohorts. They found similar significant associations between different measures of smoking behaviour and higher GlycA and hsCRP levels.[171]

3.5.2.15. Effect of exercise

Increased physical activity and weight loss are effective ways to reduce inflammation. Only a few studies have described how these lifestyle changes contribute to reducing GlycA.[68,110] One study mentioned above with a 6-month intervention of resistance exercise alone or combined with diet in overweight and pre-diabetes individuals showed that GlycA levels decreased significantly by 2%.[100] Barber et

al. demonstrated significantly reduced plasma GlycA in 1,568 individuals submitted to 14 exercise interventions even after adjustment for age, sex, race, baseline BMI, and baseline GlycA.[172]

A very interesting study by Kujala et al. shows clear differences between physically active and physically inactive age- and sex-matched pairs and twin pairs. They found better metabolic health including a decrease in isoleucine, α1-acid glycoprotein and glucose in the physically active subjects.[68]

3.5.2.16. Effect of treatments

In clinical terms there is a need to be able to measure the effect of pharmacological treatments with reliable markers. GlycA has been measured in a few drug studies to evaluate its anti-inflammatory effect and in the near future it could be a clinically relevant biomarker for monitoring disease severity.

3.5.2.16.1. Modulators of the inflammatory and immune response

Anti- TNFα therapy has been established as an efficacious therapeutic strategy in inflammatory diseases. Interestingly, the treatment of psoriasis with anti-TNFα therapy led to a decrease in GlycA levels and vascular inflammation in close parallel with reductions in atherosclerotic CVD activity.[146]

The effect of some monoclonal antibodies on GlycA levels has also been reported in the literature. Dierckx et al. showed a consistent decrease in GlycA levels in IBD patients during therapy with

Adalimumab, infliximab, vedolizumab and ustekinumab compared to other biomarkers such as fecal calprotectin or CRP,[149] which have had a moderate response at most in some patients with IBD.[173]

3.5.2.16.2. Antiretroviral therapy

Kelesidis et al. investigated how markers of inflammation change in response to treatment in a longitudinal antiretroviral therapy study. They found that GlycA was the only marker of inflammation, above hsCRP, IL-6 and D-dimer, that decreased across all the treatment groups conducting an initial antiretroviral therapy with atazanavir, raltegravir and darunavir.[138] Even though future studies are needed to determine the role of protein glycans in HIV-1 infection, this finding suggests that GlycA could be a marker for evaluating the success of treatment in these patients.

3.5.2.16.3. Statins

Statins inhibit a key step in the biosynthetic pathway of sterol by reducing cholesterol and contributing to the prevention of cardiovascular disease. Although statins decrease some inflammation markers such as CRP, they do not appear to affect GlycA levels. More studies are required to confirm this.[133]

3.5.2.16.4. Metformin

Metformin is a widely prescribed medication that has been used to treat T2DM. NAC serum levels decrease in metformin-treated T2DM patients compared to untreated patients.[56]

3.5.2.16.5. Probiotics

The use of probiotics is an emerging approach for reducing chronic inflammation, but few studies have evaluated the effect of probiotics on inflammatory markers. In the literature, greater gut microbiota richness is negatively linked with the inflammation marker GlycA[54] in overweight pregnant women. More studies are needed to confirm that probiotics can play a role in reducing GlycA levels.

3.5.2.17. Toxicity

Drug-induced hepatotoxicity is an important healthcare issue in the sense that side effects in patients can be serious. Currently, toxicity is only detected when the tissue has already been badly damaged.[74] For this reason, new markers are required to warn of the toxicity of drugs and to better monitor the treatment. Huo et al. used ^{1}H-NMR to find that the N-acetyl moieties of glycoprotein were significantly increased ($p<0.01$) in sodium valproate-induced hepatotoxicity in epileptic patients, among other metabolites such as glucose, lactate, acetoacetate, VLDLc/LDLc, lysophosphatidylcholines, phosphatidylcholines, choline, creatine, amino acids, pyruvate and uric acid.[74]

It should be noted that predicting the toxicity of radiotherapy treatment is also a new research avenue in oncology because of the temporary toxicity generated that seriously affects the patient's quality of life. In a case of head and neck squamous cell carcinoma (HNSCC), the toxicity of the treatment was studied via ^{1}H-NMR of human blood

serum. The high acute radiation sequelae were associated with increased N-acetyl glycoprotein signals, concordant with a significant increase in CRP levels and suggestive of an elevated inflammatory state.[63]

3.6. Concluding remarks and future perspectives

In this review we have discussed the clinical importance of glycoproteins in the onset of some diseases. Specifically, we have summarized the main results of the clinical studies carried out to date on plasma glycoprotein concentrations detected using ^{1}H-NMR spectroscopy, since it is the main high-performance metabolomic technique capable of quantifying serum or plasma glycoproteins in a sensitive and robust way.

Although the role of glycans in disease mechanisms is still not fully understood, the plasma level of glycosylation has been associated with different diseases, most of which have a marked inflammatory component. As has been mentioned in section 3.3, plasma glycoproteins belong to the family of APP, which increase in concentration when there are inflammatory processes. To a great extent, they could be considered to be disease diagnostic markers. It is important to note that the ^{1}H-NMR technique is a quantitative technique that measures a global state of glycation, but it cannot identify exactly which proteins are involved. As we have mentioned in this review, the most abundant proteins detected are α1-acid

glycoprotein, α1-antitrypsin, haptoglobin, transferrin and immunoglobulins, with the parameters NAG, NAC or GlycA being total measures of systemic inflammation.

Section 3.5 shows how interest in ^{1}H-NMR glycoprotein research in the clinical field has been increasing from year to year. The most widely studied diseases within this field are tumours, metabolic diseases, cardiovascular diseases, chronic inflammatory diseases, and other diseases also discussed in this review. However, more studies are needed, for example, on rare vascular diseases or cognitive diseases. All results have in common an increase in GlycA levels – or circulating levels of N-acetyl glycoproteins – with respect to a control group. In addition, high associations with CRP, IL-6 and even with the leptin/adiponectin index are widely described. The variable GlycB has been shown to be associated to these parameters but less strongly. However, this has been reported in so few studies[51,99,101,129] that firm conclusions cannot be drawn until more studies are performed.

There is general consensus that GlycA is a robust marker of systemic inflammation. Special mention should be made of the association found with CRP. Although CRP is the inflammatory biomarker that has been most studied since ancient times, it has been shown to be prone to fluctuations for a variety of reasons.[174] Joshi et al. found an association of GlycA levels with subclinical coronary atherosclerosis among in psoriatic patients independently of other traditional CVD risk factors such as hsCRP.[146] Another example is the independent association of obesity with subclinical atherosclerosis and

intima-media carotid artery regardless of hsCRP status in the MESA study,[175] which indicates that the usefulness of hsCRP for predicting CVD in patients with chronic systemic inflammation is limited, as it is non-specific and may not fully capture all forms of inflammation.[48] In general, the associations between GlycA and morbidity and mortality have been higher and CRP independent, suggesting that GlycA better captures systemic inflammation than CRP because of its composite nature.[82,88,89,111,123,124]

These results raise some open questions. The first is that the two markers may follow distinct inflammatory pathways because some studies have reported different behaviour between GlycA and hsCRP.[101] The second is that GlycA may integrate more multiple inflammatory pathways by capturing the global signal of several proteins and, therefore, better captures the degree of systemic inflammation.

Some authors suggest that the use of both markers provides additional information about the status of inflammatory disease.[51,115] However, regardless of their possibly distinct inflammatory origin, GlycA has clear advantages over CRP: while hsCRP needs to be measured several times on consecutive days to ensure the risk of CVD, GlycA needs only one measure. On the other hand, Otvos et al. in previous studies[52] ensures a high reliability of GlycA because its measures are similar in both serum and plasma samples, in fasting and non-fasting states and also after short or long-term storage, having GlycA a lower intra-individual variability than hsCRP.

Even though most of the results are consistent, some drawbacks need to be pointed out. The first is that the mechanism of glycation in human disease is not yet fully understood because of the highly complicated structures of glycans and their mechanism of action. Therefore, more research is needed. Secondly, there is a lack of a consensus on the technique for quantifying glycoproteins by NMR spectroscopy. For a rigorous transfer to the clinic in the future, the techniques used by different laboratories should be validated and standardized so that these markers can become part of conventional clinical practice.

Although more research is needed on GlycA and pharmacological treatments, the results of the studies lead us to hypothesize that GlycA's response to treatments may help to improve treatment follow-up and bring us closer to the concept of personalized medicine in the future.

In summary, the characterization of glycoproteins by ^{1}H-NMR has two main advantages. First, the versatility of the ^{1}H-NMR technique, which from a single spectrum gives the inflammatory information provided by the glycoprotein profile as well as other information provided by metabolites such as lipoproteins. Second, the applications reviewed here demonstrate that GlycA is potentially a key biomarker in a wide range of diseases. These aspects are fundamental to the future use of ^{1}H-NMR glycoprotein quantification in clinical practice.

3.7. Supporting information

This section contains the **table C3-1S**, that includes the research articles found in the literature related to the determination of glycoproteins through ^{1}H-NMR. It reflects the most important aspects such as the main objective, number of participants, the employed key methods and the main results related to glycoproteins.

Table C3-1S. Research articles found in the literature related to the determination of glycoproteins through ^{1}H-NMR

Main objective	Participants	Key methods	Main results related to glycoproteins	Clinical topic	Reference
Tumours and cancer					
To characterize N-acetyl protons of highly mobile N-acetylated carbohydrate side-chains associated with 'acute-phase' plasma glycoproteins.	> 10 healthy adults, 6 MG, 1 melanoma and 5 RA and > 10 pairs from bothmother and cord at birth	Bruker AM500 spectrometer Standard additions of glycoproteins (spiking) Plasma samples - 20ºC storage	The intensities of peaks in the spectrum of maternal plasma are greater than those of the cord plasma. Peaks are also more intense in maternal plasma than in plasma from non-pregnant women. The intensities of peaks in the spectra of plasma from subjects with melanoma and RA are greater than those of normal plasma.	Pregnancy, melanoma and RA	Bell 1987[25]
To study the variations in NAG and NANA glycosylated residues in three clinical situations: cancerous pathologies, acute inflammatory processes and autologous bone marrow transplantation (BMT).	225 patients 49 controls	Bruker AM 400 MHz spectrometer Plasma samples - 20ºC	i) The distribution of glycosylated residues varies with the origin of the cancerous tissue; ii) The level of these residues is a function of tumour development; iii) The concentrations in NAG and NANA are well correlated with the standard biological parameters of acute phase and leucocyte activation.	Cancerous pathologies	Kriat et al. 1991[55]

An unassigned and prominent resonance in the region from d 2.0–2.1ppm has frequently been found in the in vivo MR spectra of cancer patients. 1H-NMRS on the aspirated cyst fluid (in vitro) of patients confirmed the observation.	11 Ovarian tumour patients	Bruker DMX-500 spectrometer Cyst fluid stored at -70ºC COSY	N-acetyl groups from glycoproteins and/or glycolipids may contribute to the d 2.0–2.1ppm resonance complex in ovarian cyst fluid.	Ovarian tumour	Kolwijck et al 2009[57]
To examine the association of baseline GlycA concentration with total death, incident CVD, chronic inflammatory-related non-CVD, ChrIRD events and total cancer, and to compare these associations with other commonly used clinical biomarkers of chronic	6523 men and women from MESA (healthy at baseline). Median follow-up 12.1 y.	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Relative risk per SD of GlycA, IL-6, and D-dimer for total death; for total CVD; and for ChrIRD. Only GlycA was predictive for total cancer. Women had 7% higher values of all inflammatory biomarkers than men and a significantly lower GlycA prediction coefficient than men in predicting total cancer.	All-cause mortality (CVD, ChrIRD, cancer)	Duprez et al.2016 [89]

inflammation (hsCRP, IL-6, and D-dimer).					
To reveal the changes in the metabolic profiles during BC progression. Detection of early stage and late stage altered metabolic pattern.	72 early- and late-stage BC patients 50 healthy subjects	Bruker Avance III 800 MHz NMR spectrometer CPMG pulse Plasma samples	The levels of hydroxybutyrate, lysine, glutamate, glucose, NAG and lactate were highly distinguished in BC. ROC curve showed that glutamate, lactate and NAG levels of metabolic alterations separate EBC and LBC with an AUC value of 0.7	Breast cancer: EBC and LBC	Suman et al.2018 [79]
To investigate whether metabolomic profiles, generated from a simple blood draw from healthy women, could help predict the risk of developing breast cancer in the 10–15 subsequent years.	206 breast cancer cases diagnosed during a 13-year follow-up 396 matched controls SU.VI.MAX cohort	Bruker AVANCE III 500 MHz NMR spectrometer NOESY1D and CPMG pulses Plasma samples	Women characterized by higher fasting plasma levels of valine, lysine, arginine, glutamine, creatine, creatinine and glucose, and lower plasma levels of lipoproteins, lipids, glycoproteins, acetone, glycerol-derived compounds and unsaturated lipids had a higher risk of developing breast cancer.	Risk of breast cancer development	Lecuyer et al. 2018[71]

To investigate the metabolic signatures of a glioma in plasma in order to assess the diagnostic potential of this approach and gain novel insights into the metabolism of glioma and its systemic effects.	70 glioma patients 70 controls	Varian Unity Inova 600 spectrometer CPMG pulse Plasma samples	Patients with a glioma were associated with lower concentrations of isoleucine, leucine, valine, lactate, alanine, **glycoprotein**, glutamate, citrate, creatine, myo-inositol, choline, tyrosine, phenylalanine, 1-methylhistidine, α-glucose, β-glucose, and higher concentrations of very low-density lipoprotein, low density lipoprotein (LDL), unsaturated lipids, and pyruvate.	Glioma	Kelimu et al.2016 [90]
To examine the association between GlycA and incident CRC and mortality	Discovery cohort: 27,495 participants from the WHS study (median follow-up 19 y.); Replication cohort: 6,784 participants from the MESA study (median follow-up 11 y.)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	In WHS, adjusted HRs per SD increment of GlycA for CRC incidence and mortality were 1.19 (1.06±1.35; p = **0.004**) and 1.24 (1.00±1.55; p = **0.05**), respectively. Replicated findings in MESA showed that HRs per SD of GlycA for CRC incidence and mortality were 1.32 (1.06±1.65; p = **0.01**) and 1.54 (1.06±2.23; p = **0.02**), respectively, after adjusting for age, sex, and race.	CRC	Chandler et al.2016 [88]

To investigate whether metabolic differences could be detected between HCC cases and matched controls from a prospective cohort study using serum samples collected prior to diagnosis.	114 primary HCC cases 222 controls	Bruker Avance III 800 MHz spectrometer CPMG and NOESY pulses Serum samples	Sixteen metabolites of either endogenous or exogenous origin including lower levels of N-acetyl glycoproteins were found to be significantly associated with HCC risk.	Hepatocellular carcinoma	Fages et al. 2015[75]
1H NMR-based metabolomic analysis of serum samples from patients with UTUC.	39 UTUC patients 34 healthy controls	Bruker AV 500 MHz spectrometer CPMG pulse Serum samples	Serum LDL, VLDL, valine and glycoprotein levels followed a decreasing trend, whereas serum PUFA and 3,7-dimethyluric acid levels showed an increasing trend in UTUC patients compared with healthy controls.	UTUC	Li et al.2015 [62]
To use 1H-NMR to distinguish between the metabolic fingerprints of COPD and lung cancer patients	77 NSCLC 22 COPD	Bruker Avance II 600MHz CPMG Plasma samples stored at -80ºC	Increased N-acetylated glycoproteins were observed in all lung cancer patients compared with the COPD group. These metabolite biomarkers may prove useful in distinguishing lung cancer states: isoleucine, acetoacetate, and creatine as well as the two NMR signals of N-acetylated glycoproteins and glycerol.	Lung cancer and COPD	Deja et al.2014 [73]

To investigate molecular processes that reflect acute radiation sequelae in HNSCC patients using NMR-based metabolomics of blood serum.	45 patients with HNSCC: low ARS (26 patients), high ARS (19 patients)	Bruker 400.13 MHz Avance III spectrometer NOESY, CPMG, DIFF, JRES pulses Serum samples	The high ARS group was characterized by the increased signals arising from **NAG** and acetate as well as the decreased signals of branched amino acids, alanine, creatinine, choline containing compounds and carnitine. Serum glucose is low in the high ARS group. NAG is positively correlated with CRP, platelet count (PLT), ESR and absolute monocyte count and is also negatively correlated with albumin and mean platelet volume	Acute radiation sequelae in HNSCC	Bogusz ewicz et al. 2016[63]
To carry out a 1H-NMR-based metabolic phenotyping study to identify coordinated metabolic serum changes associated with advanced metastatic breast cancer (MBC) in comparison to the localized early disease (EBC).	46 EBC 39 MBC Validation: 61 EBC and 51 MBC	Bruker Avance III spectrometer 800MHz Standard 1H 1D NMR pulse sequences, NOESY and CPMG Serum samples	9 statistically significant differences between EBC and MBC patients: histidine, acetoacetate, glycerol, pyruvate, glycoproteins (NAC1 p<0.027 and NAC2 p<0.007), mannose, glutamate and phenylalanine	Breast cancer: EBC and MBC	Jobard et al. 2014[72]

To find a possible correlation between the t biochemical serum contents of a1-GP, a1-AT, Hp, CRP, IgA, IgM, IgG and Tf) and MRS data (NAG and NANA) in selected pathologies	40 patients (cancer, inflammatory and infectious diseases, diabetes) 10 controls	Bruker AM 400-WB spectrometer Serum samples -80ºC storage	High correlation between MRS data and the most abundant acute-phase glycoproteins a1-GP, a1-AT and Hp. No correlation with Ig levels. Biochemical and MRs variables are independently sensitive to the inflammatory status of the patient. However, the combination of the two sets of variables fails to provide additional sensitivity and specificity.	Cancer, inflammatory and infectious diseases, diabetes	Torri 1999[45]
To use 1H NMR-based metabonomics to investigate esophageal cancer metabolic signatures in plasma and urine	108 EC patients 40 healthy subjects	Varian Unity Inova 600MHz NMR spectrometer CPMG, COSY, TOCSY Plasma samples	Compared to controls, EC plasma had higher levels of dimethylamine, a-glucose, b-glucose, citric acid, and lower levels of leucine, alanine, isoleucine, valine, glycoprotein, lactate, acetone, acetate, choline, isobutyrate, unsaturated lipids, VLDL, LDL, 1-methylhistidine	EC	Hasim et al. 2012[61]

To use 1H-NMR based metabonomics to characterize the metabolic profiles of cervical intraepithelial neoplasia (CIN) and cervical squamous cell carcinoma (CSCC).	38 patients with CIN 38 patients with CSCC 38 healthy women	Varian Inova 600MHz spectrometer CPMG Plasma samples stored at -80˚C	Compared with samples from patients with CIN, the plasma of CSCC patients had higher levels of acetate, formate, lactate, isoleucine, leucine, valine, alanine, glutamine, histidine, tyrosine, acetylcysteine, myo-inositol, glycoprotein, α-glucose and β-glucose, together with lower levels of acetone, unsaturated lipid and carnitine.	Cervical carcinoma: CIN/CSCC	Hasim et al. 2013[60]
Metabolic diseases					
To examine changes in GlycA after lifestyle intervention among young, obese, prediabetic Latinos.	27 obese, prediabetic young Latinos 12-week lifestyle intervention	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA was significantly reduced ($p<0.01$) Additional improvements were observed in multiple cardiovascular risk factors, including BMI, total cholesterol and 2-hour glucose. Decreases in GlycA were associated with decreases in 2-hour glucose ($p<0.008$) and BMI ($p<0.03$).	Obese, prediabetic adolescents	Olson et al. 2018[103]

To explore the effect of bariatric surgery on GlycA in severely obese adults.	23 obese non-diabetic women undergoing Roux-en-Y gastric bypass. 31 obese non-diabetic women with sleeve gastrectomy. 14 non-obese controls. Baseline, 6- and 12-month analysis	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Bariatric surgery significantly reduced GlycA by 6 months (451 ± 47 µmol/L vs. 383 ± 50 µmol/L; P < 0.001) with further reduction at 12 months (348 ± 41 µmol/L; P < 0.001) and no difference between procedures. Increased high density lipoprotein particle size was strongly associated with reduced GlycA.	Bariatric surgery for severe obesity	Manmadhan et al. 2019[101]
To contrast whether metabolic phenotyping can provide a better understanding of the unique set of regulatory perturbations that predispose to diabetes and its associated complications/ comorbidities.	38 diabetes patients with good glycaemic control (DB). 35 patients with diabetes complications with inadequate glycaemic control (DC). 50 healthy controls.	Bruker Biospin Avance-III 800 MHz CPMG pulse Serum samples	Residual signals of N-acetyl glycoproteins (NAG) were found to be decreased in patients with diabetes complications compared to diabetes patients and healthy controls.	Diabetes and diabetes complications	Rawat et al. 2019[114]

To study the relation between GlycA and type 2 diabetes and compare it with high-sensitivity C-reactive protein.	26,508 women with a follow-up of 17.2 y., 2,087 T2DM cases (WHS study)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	The relative risk of GlycA with type 2 diabetes was somewhat higher for individuals with baseline BMI < vs =25 kg. Both GlycA and hsCRP were significantly associated with the risk of incident type 2 diabetes.	T2DM	Akinkuolie et al. 2015[111]
To investigate whether intestinal microbiota composition and serum metabolic and inflammatory profiles differ and are interrelated in overweight and obese women during early pregnancy.	52 overweight and 47 obese pregnant women in early pregnancy	GlycA- Brainshake LTD (now Nightingale Health LTD protocol)[8,66] Serum samples	Low-grade inflammatory markers, GlycA and hsCRP, were statistically significantly higher in obese pregnant women than in overweight pregnant women. The correlation coefficients were also higher between GlycA and lipids than between hsCRP and lipids. GlycA and hsCRP correlated with the following concentrations; isoleucine, leucine and phenylalanine. GlycA also correlated with alanine.	Overweight and obese pregnant women	Houttu et al. 2018[164]

To investigate the associations of GlycA, interleukin-1 receptor antagonist (IL-1RA), and high-sensitivity C-reactive protein (hs-CRP) with insulin secretion, insulin sensitivity, incident type 2 diabetes, hypertension, CVD events, and total mortality in the prospective METSIM study.	5401 men without diabetes at baseline or during the follow-up period (6.8 years)	Bruker AVANCE III 500/600 MHz spectrometer Serum samples	During the follow-up period GlycA was associated with impaired insulin secretion, hyperglycemia, incident type 2 diabetes and CVD.	Insulin sensitivity and secretion	Filezova et al. 2017[113]
To examine the relation of GlycA, GlycB, and CRP with direct measures of insulin sensitivity (insulin sensitivity index [SI]) and insulin secretion (acute insulin response [AIR]).	1,225 participants (278 with T2DM and 947 without diabetes) in the IRAS.	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	1) Adiposity and SI have independent relationships with CRP concentration and GlycA and GlycB NMR signals; 2) Both CRP and GlycA demonstrate a statistically independent relation to insulin SI, suggesting that GlycA may reflect an inflammatory pathway distinct from the pathway related to CRP: 3) All three inflammatory markers are more related to 2-h glucose than to fasting glucose. 4) GlycB has weaker relationships with CRP and measures of	Insulin resistance and insulin secretion	Lorenzo et al. 2017[99]

			insulin resistance and adiposity than GlycA.		
To investigate how obesity, insulin resistance and low-grade inflammation link to circulating metabolites, and whether the connections are due to genetic or environmental factors.	1368 (531 monozygotic (MZ) and 837 dizygotic (DZ)) twins of healthy young adults. FinnTwin16 and FinnTwin12 cohorts	Bruker AVANCE III spectrometer operating at 500 MHz Serum samples	Fat, especially in the abdominal area, together with HOMA-IR and CRP correlated significantly with an atherogenic lipoprotein profile, higher levels of branched-chain and aromatic amino acids, higher levels of **glycoprotein**, and a more saturated fatty acid profile.	Obesity, insulin resistance, low-grade inflammation	Bogl et al. 2016[95]
To evaluate how exercise-based lifestyle or exercise plus diet interventions modulate GlycA in persons at risk of T2DM.	169 sedentary adults with prediabetes. 6-month exercise-based lifestyle interventions, 4 intervention groups: (1) low amount/moderate intensity (2) high-amount/moderate intensity (3) high-amount/vigorous-intensity (4) a Clinical Lifestyle	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	At baseline, women had significantly greater concentrations of GlycA. No significant differences between groups were detected. GlycA was reduced on average by 3% in the High-Vig group and on average by 4% in the Clinical Lifestyle intervention. The Low-Mod group reduced GlycA on average by 1% while the High-Mod group increased GlycA on average by 1%.	Prediabetes /exercise and diet-based lifestyle interventions	Barlett et al. 2017[100]

	(combined diet plus low-amount/moderate-intensity exercise)				
To evaluate the analytical performance of the GlycA test, measured on the Vantera® Clinical Analyzer. To study the relationship of GlycA with the risk of T2DM.	4524 individuals from the PREVEND study (general population) Follow-up 8.5 years	Reference: GlycA-LipoScience protocol Comparator: Vantera® Clinical Analyzer, a 400 MHz NMR spectrometer Plasma samples	Participants with higher levels of GlycA were more likely to be older and tended to have higher BMI, blood pressure, glucose and hsCRP levels. GlycA predicted incident T2DM in models adjusted for age, sex, and for BMI, alcohol intake, smoking status, lipid lowering drugs, anti-hypertensive medication, systolic blood pressure, total cholesterol, HDL-C and TG.	T2DM	Connelly et al. 2016[82]
To compare plasma GlycA and Lp-PLA2 mass between subjects without T2DM or MetS and subjects with T2DM and/or MetS	58 subjects with T2DM and/or MetS (group 2) 40 control subjects (group 1)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA and hsCRP were higher, whereas Lp-PLA2 was lower in group 2 vs group 1. GlycA was positively related to hsCRP in each group.	T2DM and/or MetS	Gruppen et al. 2016[209]

To investigate associations of circulating metabolites from high-throughput profiling separately for fasting and 2-h glucose cross-sectionally and prospectively in middle-aged Finnish men and women	1873 individuals 618 after 6.5 years	Bruker AVANCE III spectrometer operating at 500.36 MHz Serum samples stored at -80ºC (NMR Spectroscopy protocol by Soininen et al. 2009)[66]	A1-acid glycoprotein was prospectively associated with both fasting and 2h-glucose ($p<0.05$)	Glycemia	Würt et al. 2012[110]
(i) To determine whether plasma GlycA is elevated in subjects with MetS and (ii) to assess the relationship of GlycA with plasma LCAT activity	58 MetS (46 subjects with T2DM) and 45 controls	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA was found to be elevated in MetS, but not positively associated with the presence of T2DM. GlycA was correlated positively with systolic blood pressure, BMI and waist circumference as well as with plasma triglycerides, and inversely with HDL cholesterol. GlycA was related to higher plasma LCAT activity.	MetS	Gruppen et al. 2015[115]

To compare GlycA and other markers of inflammation among hospitalized, noncritically ill patients with type 2 diabetes	121 T2DM patients: (71 CHF, 21 cardiac-non-CHF, 18 infectious diseases, and 11 in other categories)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA varied significantly across diagnostic categories and values were highest in patients with infectious disease. GlycA was associated with higher IL-6 and CRP and lower hemoglobin and estimated GFR; GlycA was not associated with HbA1c.	T2DM and other noncritical illnesses	Dungan et al. 2015[112]
To test whether plasma GlycA elevations are associated with lower bilirubin in MetS, and to assess the extent to which the association of GlycA with MetS is attenuated when taking account bilirubin levels.	58 MetS (46 T2DM) 63 without MetS (19 T2DM)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA and hs-CRP were higher, coinciding with lower bilirubin in MetS ($p<0.01$ for each). GlycA was strongly correlatedwith hs-CRP. GlycA and hs-CRP were both associated positively with the presence of MetS. GlycA and hs-CRP were negatively related to bilirubin, regardless of MetS and diabetes status.	MetS	Dullaart et al. 2015[93]
Cardiovascular risk					
A new binning algorithm, Adaptive Intelligent Binning (AI-Binning), is presented to characterize hypertensive spectra	40 hypertensive and 40 matched normotensive subjects	Bruker Avance II spectrometer 700.13 MHz CMPG pulse Serum samples	The binning algorithm enabled the relevant metabolites to be identified and suggested the involvement of a-1 acid glycoproteins and choline biochemistry in hypertension.	HBP	Tim De Meyer - 2008[50]

To evaluate changes in key metabolites following congenital heart surgery and to examine the potential of metabolic profiling for stratifying patients in terms of expected clinical outcomes.	28 children undergoing surgery for congenital heart disease (15 underwent tight glycemic control postoperatively and 13 were treated conventionally)	Bruker AVANCE III spectrometer 600 MHz CPMG pulse Plasma samples	Acetate, acetoacetate, acetone, alanine, citrate, formate, glucose, 3-hydroxybutyrate, isoleucine, leucine, N-acetylated glycoprotein, threonine, and valine had significantly higher concentrations in the postoperative samples.	Children with congenital heart disease	Correia et al. 2015[69]
To investigate the associations of plasma metabolic markers with the risk of incident MI, IS and ICH.	912 MI, 1146 IS, and 1138 ICH cases 1466 control subjects	GlycA- Brainshake LTD (now Nightingale Health LTD protocol)[8,66] Serum samples	Glycoprotein acetyls, ketone bodies, glucose, and docosahexaenoic acid were associated with all 3 diseases	CVD prediction	Holmes et al. 2018[127]
To determine if GlycA adds independent value to hsCRP for CV risk prediction.	2996 patients in the Intermountain Heart Collaborative Study who underwent coronary	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA and HS-CRP were moderately correlated. The interaction between GlycA and HS-CRP was statistically significant for the outcome of death. Baseline levels of both GlycA and HS-CRP were found to be independent and	CVD prediction	Muhlestein et al. 2018[131]

	angiography. Median follow-up 7.9 years		additive markers of risk for future major adverse CV events, especially death and hospitalization.		
a) To define the role of GlycA as a potential biomarker of adverse events in patients undergoing cardiac catheterization; (b) to evaluate the independent and incremental predictive performance of GlycA and HDL subclasses; and (c) to understand a priori defined potential interactions between HDL subclasses and GlycA.	7617 individuals in the CATHGEN cardiac catheterization biorepository	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA was associated with the presence and extent of coronary artery disease and with all-cause mortality, cardiovascular mortality and noncardiovascular mortality in models adjusted for 10 cardiovascular risk factors. GlycA and smaller HDL subclasses had independent but opposite effects on mortality risk prediction, with smaller HDL subclasses being protective. Individuals without diabetes who had the greatest quartile of GlycA concentration actually had a greater risk than patients with diabetes who had a lower concentration of GlycA	CVD prediction	McGarrah et al. 2017[132]
To evaluate the association and interaction between various HDL sub-particles, the inflammatory marker GlycA, and future cardiovascular risk.	2,848 patients from the angiography registry of the Intermountain Heart Study	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA, small HDL-P and medium HDL-P were significantly associated with cardiac death, but large HDL-P was not after adjustment for CV risk factors and medications. Only small HDL-P had a significant interaction with GlycA.	CVD prediction	Muhlestein et al. 2016[130]

To test if GlycA is associated with incident CVE and can improve CV risk prediction with traditional risk factors.	6,939 individuals	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA was strongly associated with incident CVE after adjustment for clinical risk factors. Consideration of GlycA in addition to traditional risk factors improves CV risk prediction in a high-risk population.	CVE	McGarrah et al. 2015[128]
To study the association of GlycA and GlycB with CVD.	2,996 patients who underwent coronary angiography for CAD determination/follow-up	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	No significant association was found between GlycA or GlycB and CAD. The incidence of major adverse CV events was significantly higher in patients with higher levels of GlycA and GlycB.	CVD prediction	Mulhestein et al. 2014[129]
To examine the association of baseline GlycA concentration with incident CVD events. To assess whether GlycA provided additional clinical utility for the risk of future cardiovascular events beyond the information conveyed by hsCRP	27,491 initially healthy women. Follow-up 17.2 years for CVD events (JUPITER trial)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	At baseline, increasing quartiles of GlycA were associated with a higher prevalence of traditional CVD risk factors and higher concentrations of hsCRP. GlycA and hsCRP were moderately correlated. The association of GlycA with CVD was attenuated after adjusting for hsCRP	CVD	Akinkuolie et al. 2014[124]

To identify biomarkers for all-cause mortality and enhance risk prediction. A high-throughput profiling of 106 plasma sample biomarkers are quantified by NMR	17,345 individuals	(NMR spectroscopy protocol by Soininen et al. 2009)[66] Plasma and serum samples	Alpha-1-acid glycoprotein, albumin, VLDL size, and citrate biomarkers were predictive of cardiovascular mortality, death from cancer and other nonvascular diseases. Alpha-1-acid glycoprotein was the strongest multivariate predictor of the risk of death from all causes	All-cause mortality	Fischer et al. 2014[123]
To determine differences in life expectancy in men and women from the PREVEND cohort with higher vs. lower levels of GlycA and hsCRP	5526 subjects from PREVEND study Median *follow up* 8.5 years	Vantera® Clinical Analyzer, a 400 MHz NMR spectrometer Plasma samples	Life expectancy in men and women at the end of follow up was lower in the highest vs the lower three quartiles of GlycA (P < .001). For hsCRP, this was only observed in men (P < .001) but not in women (P=0.67).	Life expectancy	Gruppen et al.2019 [134]
To examine the effects of ERN treatment on lipoprotein particles and GlycA and their relations with incident CVD events including mortality in a substudy analysis of the AIM-HIGH trial	3,414 AIM-HIGH participants	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Compared to placebo, ERN treatment lowered VLDL and LDL and increased HDL particle concentrations, increased LDL and HDL particle sizes, but did not affect GlycA. Baseline and in-trial GlycA levels were associated with increased risk of CVD events. None of the lipoprotein particle classes or subclasses were associated with incident CVD. All-cause mortality was significantly	CVD prediction	Otvos et al. 2018[133]

			associated with both GlycA and low levels of small HDL particles.		
To determine whether GlycA levels were associated with CV health as defined by the LS7 score and with each of its individual seven health metrics	6,479 MESA participants without CVD	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	After multivariable adjustment, GlycA remained independently and inversely associated with CV health categories. For each of the individual LS7 metrics (Blood Pressure, Blood Glucose, Total Cholesterol, Smoking, Diet, Physical Activity, and BMI), there was an inverse significant relation with GlycA levels.	CV Health	Benson et al. 2018[126]
To decompose the spectral GlycA biomarker by developing imputation models for GlycA's constituent glycoproteins, and use these imputed molecular phenotypes to investigate associations with disease risk.	11,861 adults across two population-based cohorts (DILGOM07 and FINRISK) Median 8 years follow-up	GlycA- Brainshake LTD (now Nightingale Health LTD protocol)[8,66] + immunoassays for AAT, AGP, HP, and TF + imputation models (Machine learning) Serum samples	Imputed AAT was significantly associated with risk of hospitalisation or death for substantially more outcomes (including liver diseases, heart failure and COPD). In contrast, imputed AGP was significantly associated with increased risk from only two outcomes: heart failure and chronic lower respiratory diseases. HP was the strongest predictor of chronic lower respiratory diseases, inflammatory polyarthropathies and atherosclerosis. Multiple individual glycoproteins independently and weakly predict each disease, with the GlycA NMR signal capturing this risk in aggregate.	Morbidity and mortality	Ritchie et al. 2018[53]

To examine the longitudinal association between GlycA and mortality among initially-healthy individuals.	27,524 participants in the WHS follow-up Replicate in 12,527 individuals JUPITER trial	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA for all-cause mortality was significantly increased at 5 years. Similar risk for all-cause mortality was observed in the replication cohort. Risk of CVD and cancer mortality was increased at 5 years	All-cause mortality	Lawler et al. 2016[125]
HIV infection					
To examine the associations between GlycA and subclinical coronary plaque among HIV-infected and HIV-uninfected men participating in MACS.	935 men from MACS (63% HIV-infected individuals)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	1) Higher quartiles of plasma GlycA were significantly positively associated with HCV infection, HIV infection, levels of hsCRP, IL-6, fibrinogen, sCD163, sCD14, CCL2 and CAC. 2) GlycA levels were higher in HIV-infected men than in HIV-uninfected men (397±68 vs 380±60 µmol/L, p=0.0001), and higher for men with detectable vs. undetectable viral load (413±79 vs 393±65 µmol/L, p=0.004). 3) Plasma GlycA levels positively correlated with smoking pack-years and HCV infection status and inversely correlated with HDL cholesterol levels and physical activity level among HIV-infected participants. 4) After adjusting for HIV serostatus,	CVD in HIV-infection	Tibuakuu et al. 2018[137]

			demographic and CVD risk factors, GlycA level was associated with a higher prevalence of CAC and coronary stenosis. 5) Among men with plaque, GlycA was positively associated with the extent of CAC and total plaque.		
Chronic inflammatory diseases					
To determine whether an H-NMR spectroscopic metabolic phenotypin approach could be used to identify signatures reflective of the dynamic, pathological metabolic perturbations associated with fibrosis in CHC patients	50 CHC patients 63 CHC patients validation	Bruker Avance (Avance III) 600 MHz NMR spectrometer CPMG pulse Plasma samples	Increased severity of fibrosis was associated with higher tyrosine, phenylalanine, methionine, citrate and, very-low-density lipoprotein (vLDL) and lower creatine, low-density lipoprotein (LDL), phosphatidylcholine, and NAC	CHC	Sands et al. 2015[76]

To investigate GlycA levels in a cohort of healthy individuals, patients with CD and patients with UC prior to and after therapeutic control of inflammation.	37 Crohn's disease patients and 21 ulcerative colitis patients before and after biologic therapy (ADM, IFX, VDM, UST) 10 healthy controls.	GlycA- Brainshake LTD (now Nightingale Health LTD protocol)[8,66] Serum samples	GlycA levels were significantly higher in patients with active IBD (CD and UC) than in healthy controls. GlycA levels from CD and UC patients dropped to control levels after mucosal healing. Only GlycA post-treatment levels consistently showed a significant difference between responder and non-responder levels	IBD (CD and UC)	Dierckx et al. 2018[149]
To characterize the plasma glycoprotein profile of a cohort of patients with RA versus healthy individuals and to model the activity of RA to identify patterns indicating the severity of the disease	210 RA patients 203 healthy individuals	Bruker Avance III 600 spectrometer NOESY, LED Plasma samples	RA patients presented significantly higher GlycA area and H/W GlycA and GlycB ratios than the control population. GlycA and GlycB variables derived from 1H NMR, along with classic inflammatory parameters, help to improve the classification of individuals with high RA disease activity based on DAS28.	RA	Fuertes-Martín et al. 2018[51]
To investigate if GlycA could be associated with lupus nephritis severity	105 active SLE patients 39 quiescent SLE patients 21 non-lupus nephritis controls 29 healthy controls	GlycA- Brainshake LTD (now Nightingale Health LTD protocol)[8,66] Serum samples	GlycA was correlated to C-reactive protein (CRP), neutrophil count, proteinuria and the SLE disease activity index. Patients with active SLE showed significantly higher GlycA concentration than healthy controls ($p=0.009$), non-lupus nephritic controls ($p=0.04$) and quiescent SLE patients ($p<10-6$). In	SLE and Lupus nephritis	Dierckx et al. 2018[141]

			patients with biopsy-proven active LN, GlycA was higher in proliferative than non-proliferative lupus nephitis		
To conduct a plasma metabolomic analysis to determine the characteristics of patients with COPD with abnormal Savda syndrome using NMR spectroscopy technology	103 COPD patients mild (n=15), moderate (n=38) and severe (n=50)	Inova 600, Varian Medical Systems Spectrometer CPMG pulse Plasma samples	The concentration of metabolites such as glycoprotein in the plasma of patients with COPD with abnormal Savda syndrome was lower than in the plasma of patients with COPD with non-abnormal Savda syndrome and the plasma of healthy subjects	Savda syndrome in COPD	Xu et al. 2015[152]
To investigate the association of GlycA with albuminuria and eGFR in a Brazilian cohort of middle-aged men and women.	5050 participants from ELSA-Brasil Study	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA was higher in older women, smokers, teetotallers, the obese and those with diabetes, hypertension or dyslipidemia. GlycA was independently associated with log albuminuria and inversely related to eGFR.In the ROC curve, GlycA had a higher AUC than hsCRP (AUC 0.67 vs. 0.62, p = 0.06) for the association with albuminuria A2 or A3.	CKD	Titan et al. 2017[157]

To examine whether GlycA levels increased with active disease and to establish whether this could be a more useful biomarker for predicting cardiovascular events in lupus.	52 patients 229 follow-up visits	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Mean GlycA levels in this cohort were higher than those reported for a normal population. GlycA increased significantly with each point increase in SLEDAI- The African American population had lower VLDL, triglycerides and higher levels of GlycA than the other SLE groups.	SLE	Durcan et al. 2016[142]
To investigate the relationships between GlycA and psoriasis, and between GlycA and subclinical CVD	122 psoriasis patients 109 controls	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Psoriasis patients had higher levels of hsCRP and GlycA which remained significant after adjustment for age, sex, BMI and traditional CV risk factors. GlycA is associated with vascular inflammation and coronary artery disease. Anti-TNF therapy decreases GlycA levels.	Subclinical CVD in psoriasis	Joshi et al.2016 [146]
To explore the relationships of GlycA with inflammation and cardiometabolic risk in RA, and explore whether these relationships were similar to those of people without RA.	50 mild-moderate RA patients 39 controls	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA concentrations were significantly elevated in RA versus controls (P = 0.036). In RA, greater GlycA associated with disease activity and inflammation. In BMI-matched controls, these inflammatory associations were absent or weaker. In RA, greater GlycA was associated with more total abdominal	RA and CVD	Barlett et al. 2016[139]

			adiposity and less muscle density. In BMI-matched controls, GlycA was associated with more cardio-metabolic markers: BMI, waist circumference, adiposity measures and insulin resistance		
To test whether GlycA is a biomarker of disease activity and is associated with coronary artery atherosclerosis in patients with RA.	166 RA patients 90 control subjects	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA concentrations were higher in patients with RA than in control subjects. In RA, GlycA was strongly correlated with DAS28 based on erythrocyte sedimentation rate (DAS28-ESR) and DAS28 based on C-reactive protein (DAS28-CRP) and their components, including tender and swollen joint counts, global health score, ESR and CRP. For each quartile increase in GlycA, the odds of having coronary artery calcium increased by 48 %.	RA	Ormseth et al. 2015[140]

To test the hypothesis that GlycA concentrations are elevated in patients with SLE and associated with other markers of inflammation and coronary atherosclerosis.	116 SLE patients 84 control subjects	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Patients with SLE had higher concentrations of GlycA than control subjects. In patients with SLE, concentrations of GlycA were significantly associated with ESR, CRP, e-selectin, intracellular adhesion molecule-1, and triglycerides.	SLE	Chung et al.2016 [143]
To characterize biological processes associated with GlycA by leveraging population-based omics data and health records from >10,000 individuals.	11,825 individuals from 3 cohorts: DILGOM07 (300 female), +FINRISK (7599 individuals) and + YFS (3596 individuals follow-up)	Bruker AVANCE III, 500 MHz spectrometer (the 3 cohorts) Serum samples	In apparently healthy individuals, GlycA can be chronically elevated for periods of up to a decade. In individuals with high GlycA, modest elevation of numerous cytokines is suggestive of a prolonged low-grade inflammatory state. High GlycA levels correlated with an increased risk of hospitalization and death from various infections (septicemia and pneumonia). GlycA levels persists for over a decade.	Chronic inflammation and long-term risks	Ritchie et al. 2015[48]
1H NMR spectroscopy-based metabolic phenotyping was used to identify biomarkers	47 RA patients 51 healthy subjects	Bruker Avance 600 MHz spectrometer Plasma samples	Cholesterol, lactate, acetylated glycoprotein, and lipid signatures were found to be candidate biomarkers for disease severity.	RA	Laurids en et al. 2010[58]

in the plasma of patients with RA.		stored at or below -25 °C for a period of up to 19 months			
Cognitive function and psychological health					
To present a new three-molecular-window approach that gives specific molecular data on macromolecular lipid–protein aggregates such as lipoprotein particles, on various low-molecular-weight metabolites, and on individual lipid molecules together with their degree of (poly)(un)saturation.	180 elderly people (54 related to MCI, with severely increased risk of AD).	AVANCE 500 DRX spectrometer 1D CPMG pulse sequence Serum samples	Positive association between MCI and the MetS. Low relative amount of n3 fatty acids appears more indicative of MCI than low serum n3 or polyunsaturated fatty acid concentration as such. Elevated circulating glycoproteins in the risk of AD.	MCI and AD	Tukiainen et al 2008[159]

To develop a biomarker panel to provide support for objective diagnostic laboratory tests for psychological suboptimal health.	22 psychological suboptimal health patients 23 healthy controls	Bruker 600 MHz AVANCE III NMR spectrometer CPMG Plasma samples	A biomarker panel containing phenylalanine, glutamine, tyrosine, citrate, N-acetyl-glycoproteins and TMAO was identified and there was a high correlation with the state of psychological suboptimal health.	Psychological suboptimal health	Tian et al.2016 [77]
To examine the association of the inflammatory markers CRP, fibrinogen, WBC and GlycA, measured in young adulthood and of GlycA change over 13 years follow-up with cognitive function in midlife	507 participants (13 years follow up)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	The highest quintile of GlycA change, but not the baseline inflammation measures, was inversely related to global cognition as well as to information processing speed and memory domains.	Cognitive function	Cohen-Manheim et al. 2015[158]
Rare vascular diseases					
To confirm previous TA findings in a larger group of patients and to study their correlation with disease activity	Total: 45 active TA patients and 53 inactive TA patents 57 TA patients (active and inactive) follow-up	Bruker Avance III 800 MHz NMR spectrometer CPMG pulse Serum samples	The sera of TA patients were characterized by elevated levels of LDL, NAG, glucose, glutamate, phosphoglyceride, glycerol, glycerophophocholine, and decreased levels of glucogenic amino acids, lactate	TA	Jain et al. 2018[78]

	3.1 months 43 healthy controls		and creatine. The key metabolites with highest discriminatory potential in active TA were glutamate and N-acetyl glycoprotein (NAG), both elevated.		
To determine the ability of GlycA concentrations and NMR lipoprotein particle measures to distinguish pediatric patients with acute KD from those with other febrile illnesses.	75 acute KD 36 post-treatment subacute KD 63 convalescent KD 48 febrile controls 48 and age-similar healthy controls	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	GlycA was higher in acute KD subjects than other groups. GlycA and NMR-measured lipoprotein particle parameters may be useful for distinguishing acute KD from bacterial or viral illnesses (ROC AUCs of 0.910 and 0.909 for GlycA combined with either the LDL-P/HDL-P ratio or LDL-P, respectively).	Kawasaki disease	Connelly et al. 2016[163]
To investigate the metabolic profiles of sera derived from TA patients using NMR with an aim to assess (a) whether NMR-based serum metabolomics would allow early identification of TA patients and (b) whether metabolic differences in TA	29 TA patients 30 controls	Bruker Avance III 800 MHz NMR spectrometer CPMG pulse Serum samples	Compared to healthy controls, TA patients had (a) increased serum levels of choline metabolites, LDL cholesterol, NAGs, and glucose and (b) decreased serum levels of lactate, lipids, HDL cholesterol, and glucogenic amino acids.	TA	Guleria et al. 2015[161]

patients are related to the risk of TA progression.					
Pregnancy					
To explore the effects of habitual diet and adherence to the recommended diet on gut microbiota, serum lipidomics and low-grade inflammation, and the relationship of gut microbiota composition to serum lipidomics and inflammatory markers in overweight and obese pregnant women.	100 overweight women in early pregnancy	GlycA- Brainshake LTD (now Nightingale Health LTD protocol) [8,66] Serum samples	Multiple nutrients correlated with GlycA including fibre, LC-PUFA and n-3 LC-PUFA and several vitamins and minerals but no correlations were detected between any of the nutrients and hs-CRP and LPS. Higher gut microbiota richness is negatively linked with low-grade inflammation marker GlycA, which was further related to intake of several nutrients including fibre and LC-PUFA. No similar relationship between hs-CRP and gut microbiota richness, suggesting that the inflammatory pathway of GlycA is different from that of CRP.	Diet in overweight pregnant women	Röytiö et al. 2017[54]

To investigate the extent to which intestinal permeability, measured by serum zonulin concentration, is related to metabolic endotoxemia and metabolic risk markers in overweight pregnant women.	100 overweight women in early pregnancy	GlycA- Brainshake LTD (now Nightingale Health LTD protocol) [8,66] Serum samples	Both LPS and GlycA showed a positive relationship with insulin resistance, serum insulin, triglycerides, and total and LDL-cholesterol, and a negative relationship with insulin sensitivity. Serum zonulin was found to associate positively with LPS, hs-CRP, GlycA, insulin, HOMA2-IR, triglycerides, and total cholesterol.	Intestinal permeability in overweight pregnant women	Mokkala et al. 2017[96]
To evaluate the metabolic adaptations reflected in plasma throughout healthy pregnancies by carrying out an untargeted 1H NMR study	20 non-pregnant women 25 1st T 30 2nd T 12 3rd T 7 post-delivery	Bruker Avance DRX 500 spectrometer Plasma samples Noesy, cpmg and led pulse sequences	Gradually increase of N-acetyl glicoproteins. Direct link between LDL+VLDL and N-acetyl-glycoproteins	Pregnancy	Pinto et al. 2015[81]
Primary aldosteronism					
To determine the extent to which (apo)lipoproteins, lipoprotein particle concentrations, GlycA and BCAA, as determined by NMR	20 primary aldosteronism patients 2,819* control subjects without hypertension 501* subjects with	Vantera® Clinical Analyzer, a 400 MHz NMR spectrometer Plasma samples	GlycA was increased in PA vs the three groups ($P < 0.016$).	Primary aldosteronism	Berends et al. 2019[167]

spectroscopy, were altered in individuals with PA, compared to non-hypertensive control subjects, subjects with untreated hypertension and subjects with medically treated hypertension.	untreated hypertension 878* subjects with treated hypertension *From the PREVEND study				
Sickle cell disease					
To evaluate plasma GlycA levels in a cross-sectional sample of patients with SCD and specifically test levels in patients experiencing an acute painful vaso-occlusive crisis.	488 patients with SCD in "steady state" including 52 healthy controls and 12 patients (from the same group) during an acute pain crisis.	Vantera® Clinical Analyzer, a 400 MHz NMR spectrometer Plasma samples	The mean plasma GlycA level was lower in SCD than in healthy controls. Within the same patient, mean plasma GlycA during acute pain crisis was lower than in steady state, although the difference was not significant.	SCD	Weisman et al. 2018[169]
Human African Trypanosomiasis					
To characterize the metabolic effects of T. brucei	46 HAT patients 21 controls	Bruker Avance 600 MHz CPMG Plasma samples	Among other metabolites, NAG is significantly higher in disease ($p<0.01$)	HAT	Lamour et al. 2015[170]

rhodesiense infection in humans					
Sodium intake					
To investigate the association of 24-h sodium excretion with the 2 inflammatory markers GlycA and hsCRP in a large population-based cohort of men and women. To assess the role of adiposity in the association between sodium intake and inflammatory markers	3,935 subjects from a general population (PREVEND study)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	The proinflammatory biomarkers GlycA and hsCRP are inversely related to higher 24-h sodium excretion when taking into account the variation in adiposity.	Sodium intake	Gruppen et al. 2016[116]
Tobacco smoking					
To determine whether smoking is associated with systemic inflammation as measured by GlycA levels. We also sought to compare the strength of the association of smoking	11,509 participants (6,774 from the MESA and 4,735 from ELSA-Brasil)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Compared with people who had never smoked, former and current smokers had significantly higher adjusted means of GlycA levels. Each 5-unit increase in pack-years of smoking was associated with higher GlycA levels among former and current smokers. Among former smokers, each 5-year increase in time since quitting smoking was associated	Smoke	Kianoush et al. 2017[171]

and GlycA versus the association of smoking and hsCRP.			with lower GlycA levels and each 10-unit increase in number of cigarettes/days was associated with higher GlycA among current smokers. Results were significant for the association between all measures of smoking behaviour and GlycA and hsCRP.		
Effect of excercise					
To study whether persistent physical activity compared with inactivity has a global effect on serum metabolome and leads to reduced cardiometabolic disease risk	16 same-sex twin pairs >30-year discordance for physical activity and 1,037 age-and sex-matched pairs. Median follow-up 5 years	Bruker AVANCE III 500 MHz spectrometer NOESY and CPMG pulses Serum samples	Isoleucine, α1-acid glycoprotein, and glucose were lower in the physically active than in the inactive individuals (P<0.001 in meta-analysis) (findings persisted after adjustment for BMI).	Exercise interventions	Kujala et al. 2013[68]
To examine the effects of regular exercise on the inflammatory marker GlycA across seven studies and 14 exercise interventions	1,568 individuals	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Regular exercise significantly reduced plasma GlycA even after adjustment for age, sex, race, baseline BMI, and baseline GlycA. Changes in GlycA were correlated with changes in traditional inflammatory markers, C-reactive protein, interleukin-6, and fibrinogen. However, these	Exercise interventions	Barber et al. 2018[172]

			correlations were relatively weak (range r: 0.21e0.38, p < 0.0001).		
Effect of treatments					
To use 1H NMR and UPLC/MS to study type 2-DM in patients non-treated and treated with metformin hydrochloride	20 non treated type 2-DM patients 15 treated type 2-DM	Bruker AV 600 spectrometer CPMG pulse Serum samples	NAC was lower in serum from metformin treated patients than in serum from untreated patients	T2DM treatment	Huo et al 2009[56]
To understand how markers of inflammation and immune activation change in response to successful ART	328 HIV-1 infected (week 24 to week 92 follow-up)	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	On average, a 10% decline but not significant in levels of GlycA was apparent over 48 weeks with all the studied treatment combinations.	ART for HIV-1	Kelesidis et al. 2015[138]
Toxicity					
To evaluate the hepatotoxicity of valproate sodium (antiepileptic drug) through new 1H-NMR markers	34 epileptic patients	Bruker AV 600 CPMG (64 scans) Serum samples	N-acetyl moieties of glycoprotein significantly increased (p<0.01) in valproate sodium induced hepatotoxicity, among other metabolites such as glucose, lactate, acetoacetate, VLDL/LDL, lysophosphatidylcholines,	Drug toxicity	Huo et al. 2014[74]

			phosphatidylcholines, choline, creatine, amino acids, pyruvate and uric acid.		
Others					
To determine potential relationships between GlycA and adiposity, insulin resistance, hs-CRP, leptin, adiponectin, and the leptin/adiponectin ratio, and to test whether GlycA is elevated in subjects with impaired fasting glucose and T2DM.	103 fasting subjects (30 with normal fasting glucose, 25 with IFG and 48 with T2DM).	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	Plasma GlycA was correlated positively with BMI, HOMAir, hs-CRP, leptin and the leptin/adiponectin ratio, and inversely with adiponectin. GlycA did not significantly vary with the glucose tolerance category. GlycA was related positively to the leptin/adiponectin ratio, regardless of BMI and HOMAir.	Potential relationships between GlycA and other biomarkers	Dullaart et al. 2015[94]
a) To report on levels of GlycA and the change in GlycA as children move from 6th to 8th grade; b) to examine whether BMI group is associated with GlycA in these children; c) to determine if fitness was associated with	1,664 US adolescents from the HEALTHY study	NMR-algorithm at LabCorp, Inc. (formerly LipoScience)[52] Plasma samples	In the 8th grade the median GlycA values were 27 µmol/L higher for girls than boys. GlycA levels are 19% higher in obese girls than healthy weight girls. Strong evidence ($p<0.001$) that in all sub-groups (6th grade boys and girls, and 8th grade boys and girls) GlycA is higher in higher BMI groups. The lowest levels of GlycA are in the low BMI/high fitness group	Systemic inflammation in adolescents	Jago et al. 2016[102]

GlycA, LP-IR and traditional lipid panel variables; and d) to examine whether fitness and BMI are independently related to GlycA and/or LP-IR.			with the highest levels in the high BMI/low fitness group.		
To use 1H NMR to examine the metabolic profiles of plasma from FD patients before and after treatment by acupuncture.	6 female FD patients 6 female healthy subjects	Varian INOVA 600MHz NMR spectrometer CPMG/BBP-LED Plasma samples	Lower levels of lactate, leucine/isoleucine, NAC, and LDL/VLDL in FD patients than in healthy controls	FD	Wu et al. 2010[59]

3.8. References

(1) Gabay, C.; Kushner, I. Acute-Phase Proteins and Other Systemic Responses to Inflammation. *N. Engl. J. Med.* **1999**, *340* (6), 448–454.

(2) Bergin, D. A.; Reeves, E. P.; Meleady, P.; Henry, M.; McElvaney, O. J.; Carroll, T. P.; Condron, C.; Chotirmall, S. H.; Clynes, M.; O'Neill, S. J.; et al. α-1 Antitrypsin Regulates Human Neutrophil Chemotaxis Induced by Soluble Immune Complexes and IL-8. *J. Clin. Invest.* **2010**, *120* (12), 4236–4250.

(3) McCarthy, C.; Saldova, R.; Wormald, M. R.; Rudd, P. M.; McElvaney, N. G.; Reeves, E. P. The Role and Importance of Glycosylation of Acute Phase Proteins with Focus on Alpha-1 Antitrypsin in Acute and Chronic Inflammatory Conditions. *J. Proteome Res.* **2014**, *13* (7), 3131–3143.

(4) Zhang, S.; Shang, S.; Li, W.; Qin, X.; Liu, Y. Insights on N-Glycosylation of Human Haptoglobin and Its Association with Cancers. *Glycobiology* **2016**, *26* (7), 684–692.

(5) Pezer, M.; Rudan, I.; Campbell, H. Mechanisms of Disease: The Human N-Glycome. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (8), 1574–1582.

(6) Silva, M. L. S. Cancer Serum Biomarkers Based on Aberrant Post-Translational Modifications of Glycoproteins: Clinical Value and Discovery Strategies. *Biochim. Biophys. Acta - Rev. Cancer* **2015**, *1856* (2), 165–177.

(7) Mallol, R.; Rodriguez, M. A.; Brezmes, J.; Masana, L.; Correig, X. Human Serum/Plasma Lipoprotein Analysis by NMR: Application to the Study of Diabetic Dyslipidemia. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *70*, 1–24.

(8) Soininen, P.; Kangas, A. J.; Würtz, P.; Suna, T.; Ala-Korpela, M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. *Circ. Cardiovasc. Genet.* **2015**, *8* (1), 192–206.

(9) Seo, J.; Lee, K.-J. *Post-Translational Modifications and Their Biological Functions: Proteomic Analysis and Systematic Approaches*; 2004; Vol. 37.

(10) Lodish, H.; Berk, A.; Zipursky, S.; Al, E. Glycosylation in the ER and Golgi Complex. *Mol. Cell Biol.* **2000**, No. 4, section 17.7.

(11) Colley, K. J.; Varki, A.; Kinoshita, T. *Cellular Organization of Glycosylation*; Cold Spring Harbor Laboratory Press, 2015.

(12) Unverzagt, C.; Kajihara, Y. Recent Advances in the Chemical Synthesis of N-Linked Glycoproteins. *Curr. Opin. Chem. Biol.* **2018**, *46*, 130–137.

(13) Kazuaki Ohtsubo; and Jamey D. Marth. Glycosylation in Cellular Mechanisms of Health and Disease. *Elsevier Inc.* **2006**, *Cell 126* (5), 855–867.

(14) Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. *Protein Glycosylation in the ER and Golgi Complex*, 4th edition.; W. H. Freeman, 2000.

(15) Vanhooren, V.; Laroy, W.; Libert, C.; Chen, C. N-Glycan Profiling in the Study of Human Aging. *Biogerontology* **2008**, *9* (5), 351–356.

(16) Lakshminarayanan, A.; Richard, M.; Davis, B. G. *Studying Glycobiology at the Single-Molecule Level*; 2018.

(17) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **2006**, *126* (5), 855–867.

(18) Lyons, J. J.; Milner, J. D.; Rosenzweig, S. D. Glycans Instructing Immunity: The Emerging Role of Altered Glycosylation in Clinical Immunology. *Front. Pediatr.* **2015**, *3*, 54.

(19) Fournier, T.; Medjoubi-N, N.; Porquet, D. Alpha-1-Acid Glycoprotein. *Biochim. Biophys. Acta* **2000**, *1482* (1–2), 157–171.

(20) Dall'olio, F.; Vanhooren, V.; Chen, C. C.; Slagboom, P. E.; Wuhrer, M.; Franceschi, C. N-Glycomic Biomarkers of

Biological Aging and Longevity: A Link with Inflammaging. *Ageing Res. Rev.* **2013**, *12*, 685–698.

(21) Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nat. Rev. Cancer* **2015**, *15* (9), 540–555.

(22) Taniguchi, N.; Korekane, H. Branched N-Glycans and Their Implications for Cell Adhesion, Signaling and Clinical Applications for Cancer Biomarkers and in Therapeutics. *BMB Rep.* **2011**, *44* (12), 772–781.

(23) Lefebvre, T.; Dehennaut, V.; Guinez, C.; Olivier, S.; Drougat, L.; Mir, A.-M.; Mortuaire, M.; Vercoutter-Edouart, A.-S.; Michalski, J.-C. Dysregulation of the Nutrient/Stress Sensor O-GlcNAcylation Is Involved in the Etiology of Cardiovascular Disorders, Type-2 Diabetes and Alzheimer's Disease. *Biochim. Biophys. Acta - Gen. Subj.* **2010**, *1800* (2), 67–79.

(24) Myslicki, J. P.; Shearer, J.; Hittel, D. S.; Hughey, C. C.; Belke, D. D. O-GlcNAc Modification Is Associated with Insulin Sensitivity in the Whole Blood of Healthy Young Adult Males. *Diabetol. Metab. Syndr.* **2014**, *6* (1), 96.

(25) Bell, J. D.; Brown, J. C. C.; Nicholson, J. K.; Sadler, P. J. Assignment of Resonances for 'Acute-Phase' Glycoproteins in High Resolution Proton NMR Spectra of Human Blood Plasma. *FEBS Lett.* **1987**, *215* (2), 311–315.

(26) Lipton, A.; Harvey, H. A.; Delong, S.; Allegra, J.; White, D.; Allegra, M.; Davidson, E. A. Glycoproteins and Human Cancer:1. Circulating Levels in Cancer Serum. *Cancer* **1979**, *43* (5), 1766–1771.

(27) Silverman, L. M.; Dermer, G. B.; Tökes, Z. A. Electrophoretic Patterns for Serum Glycoproteins Reflect the Presence of Human Breast Cancer. *Clin. Chem.* **1977**, *23* (11).

(28) Hakomori, S.-I. Aberrant Glycosylation in Cancer Cell Membranes as Focused on Glycolipids: Overview and Perspectives1 Introduction: Glycosylation as a Modulatory

Mechanism of Cellular Function and Cell-Social Interaction. *CANCER Res.* **1985**, *45*, 2405–2414.

(29) Patel, P. S.; Adhvaryu, S. G.; Balar, D. B. Serum Glycoconjugates in Patients with Anemia and Myeloid Leukemia. *Tumori J.* **1988**, *74* (6), 639–644.

(30) Rye, P. D.; Walker, R. A. Analysis of Glycoproteins Released from Benign and Malignant Human Breast: Changes in Size and Fucosylation with Malignancy. *Eur. J. Cancer Clin. Oncol.* **1989**, *25* (1), 65–72.

(31) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. Vertebrate Protein Glycosylation: Diversity, Synthesis and Function. *Nat. Rev. Mol. Cell Biol.* **2012**, *13* (7), 448–462.

(32) Austrup, F.; Vestweber, D.; Borges, E.; Löhning, M.; Bräuer, R.; Herz, U.; Renz, H.; Hallmann, R.; Scheffold, A.; Radbruch, A.; et al. P- and E-Selectin Mediate Recruitment of T-Helper-1 but Not T-Helper-2 Cells into Inflamed Tissues. *Nature* **1997**, *385* (6611), 81–83.

(33) Bedard, P. W.; Kaila, N. Selectin Inhibitors: A Patent Review. *Expert Opin. Ther. Pat.* **2010**, *20* (6), 781–793.

(34) Chatterjee, B. P.; Mondal, G.; Chatterjee, U. Glycosylation of Acute Phase Proteins: A Promising Disease Biomarker. *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.* **2014**, *84* (4), 865–874.

(35) Baumann, H.; Gauldie, J. The Acute Phase Response. *Immunol. Today* **1994**, *15* (2), 74–80.

(36) Pilobello, K. T.; Mahal, L. K. Lectin Microarrays for Glycoprotein Analysis. *Methods Mol. Biol.* **2007**, *385*, 193–203.

(37) Gornik, O.; Lauc, G. Glycosylation of Serum Proteins in Inflammatory Diseases. *Dis. Markers* **2008**, *25* (4–5), 267–278.

(38) Qi, Y.-J.; Ward, D. G.; Pang, C.; Wang, Q.-M.; Wei, W.; Ma, J.; Zhang, J.; Lou, Q.; Shimwell, N. J.; Martin, A.; et al. Proteomic Profiling of N-Linked Glycoproteins Identifies ConA-Binding Procathepsin D as a Novel Serum Biomarker for Hepatocellular

Carcinoma. *Proteomics* **2014**, *14* (2–3), 186–195.

(39) Miura, Y.; Hato, M.; Shinohara, Y.; Kuramoto, H.; Furukawa, J.; Kurogochi, M.; Shimaoka, H.; Tada, M.; Nakanishi, K.; Ozaki, M.; et al. BlotGlycoABCTM, an Integrated Glycoblotting Technique for Rapid and Large Scale Clinical Glycomics. *Mol. Cell. Proteomics* **2008**, *7* (2), 370–377.

(40) Gizaw, S. T.; Ohashi, T.; Tanaka, M.; Hinou, H.; Nishimura, S.-I. Glycoblotting Method Allows for Rapid and Efficient Glycome Profiling of Human Alzheimer's Disease Brain, Serum and Cerebrospinal Fluid towards Potential Biomarker Discovery. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (8), 1716–1727.

(41) Connelly, M. A.; Gruppen, E. G.; Otvos, J. D.; Dullaart, R. P. F. Inflammatory Glycoproteins in Cardiometabolic Disorders, Autoimmune Diseases and Cancer. *Clin. Chim. Acta* **2016**, *459*, 177–186.

(42) Lenz, E. M.; Bright, J.; Wilson, I. D.; Morgan, S. R.; Nash, A. F. P. A 1H NMR-Based Metabonomic Study of Urine and Plasma Samples Obtained from Healthy Human Subjects. *J. Pharm. Biomed. Anal.* **2003**, *33* (5), 1103–1115.

(43) Beckonert, O.; Keun, H. C.; Ebbels, T. M. D.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Metabolic Profiling, Metabolomic and Metabonomic Procedures for NMR Spectroscopy of Urine, Plasma, Serum and Tissue Extracts. *Nat. Protoc.* **2007**, *2* (11), 2692–2703.

(44) Aru, V.; Lam, C.; Khakimov, B.; Hoefsloot, H. C. J.; Zwanenburg, G.; Lind, M. V.; Schäfer, H.; van Duynhoven, J.; Jacobs, D. M.; Smilde, A. K.; et al. Quantification of Lipoprotein Profiles by Nuclear Magnetic Resonance Spectroscopy and Multivariate Data Analysis. *TrAC Trends Anal. Chem.* **2017**, *94*, 210–219.

(45) Torri, G. M.; Torri, J.; Gulian, J.-M.; Vion-Dury, J.; Viout, P.; J. Cozzone, P. Magnetic Resonance Spectroscopy of Serum and Acute-Phase Proteins Revisited: A Multiparametric Statistical Analysis of Metabolite Variations in Inflammatory, Infectious

and Miscellaneous Diseases. *Clin. Chim. Acta* **1999**, *279* (1–2), 77–96.

(46) Jeyarajah, E. J.; Cromwell, W. C.; Otvos, J. D. Lipoprotein Particle Analysis by Nuclear Magnetic Resonance Spectroscopy. *Clin. Lab. Med.* **2006**, *26* (4), 847–870.

(47) Sears, B.; Deckelbaum, R. J.; Janiak, M. J.; Shipley, G. G.; Small, D. M. Temperature-Dependent 13C Nuclear Magnetic Resonance Studies of Human Serum Low Density Lipoproteins. *Biochemistry* **1976**, *15* (19), 4151–4157.

(48) Ritchie, S. C.; Würtz, P.; Nath, A. P.; Abraham, G.; Havulinna, A. S.; Fearnley, L. G.; Sarin, A.-P.; Kangas, A. J.; Soininen, P.; Aalto, K.; et al. The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection. *Cell Syst.* **2015**, *1* (4), 293–301.

(49) Tuck, M. K.; Chan, D. W.; Chia, D.; Godwin, A. K.; Grizzle, W. E.; Krueger, K. E.; Rom, W.; Sanda, M.; Sorbara, L.; Stass, S.; et al. Standard Operating Procedures for Serum and Plasma Collection: Early Detection Research Network Consensus Statement Standard Operating Procedure Integration Working Group. *J. Proteome Res.* **2009**, *8* (1), 113–117.

(50) De Meyer, T.; Sinnaeve, D.; Van Gasse, B.; Tsiporkova, E.; Rietzschel, E. R.; De Buyzere, M. L.; Gillebert, T. C.; Bekaert, S.; Martins, J. C.; Van Criekinge, W. NMR-Based Characterization of Metabolic Alterations in Hypertension Using an Adaptive, Intelligent Binning Algorithm. *Anal. Chem.* **2008**, *80* (10), 3783–3790.

(51) Fuertes-Martín, R.; Taverner, D.; Vallvé, J.-C.; Paredes, S.; Masana, L.; Correig Blanchar, X.; Amigó Grau, N. Characterization of ^{1}H NMR Plasma Glycoproteins as a New Strategy To Identify Inflammatory Patterns in Rheumatoid Arthritis. *J. Proteome Res.* **2018**, *17* (11), 3730–3739.

(52) Otvos, J. D.; Shalaurova, I.; Wolak-Dinsmore, J.; Connelly, M. A.; Mackey, R. H.; Stein, J. H.; Tracy, R. P. GlycA: A Composite Nuclear Magnetic Resonance Biomarker of Systemic

Inflammation. *Clin. Chem.* **2015**, *61* (5), 714–723.

(53) Ritchie, S. C.; Kettunen, J.; Brozynska, M.; Nath, A. P.; Havulinna, A. S.; Männistö, S.; Perola, M.; Salomaa, V.; Ala-Korpela, M.; Abraham, G.; et al. Elevated Alpha-1 Antitrypsin Is a Major Component of GlycA-Associated Risk for Future Morbidity and Mortality. *bioRxiv* **2018**, 309138.

(54) Röytiö, H.; Mokkala, K.; Vahlberg, T.; Laitinen, K. Dietary Intake of Fat and Fibre According to Reference Values Relates to Higher Gut Microbiota Richness in Overweight Pregnant Women. *Br. J. Nutr.* **2017**, *118* (05), 343–352.

(55) Kriat, M.; Vion-Dury, J.; Fayre, R.; Maraninchi, D.; Harlé, J. R.; Confort-Gouny, S.; Sciaky, M.; Fontanarava, E.; Viout, P.; Cozzone, P. J. Variations of Plasma Sialic Acid and N-Acetylglucosamine Levels in Cancer, Inflammatory Diseases and Bone Marrow Transplantation: A Proton NMR Spectroscopy Study. *Biochimie* **1991**, *73* (1), 99–104.

(56) Huo, T.; Cai, S.; Lu, X.; Sha, Y.; Yu, M.; Li, F. Metabonomic Study of Biochemical Changes in the Serum of Type 2 Diabetes Mellitus Patients after the Treatment of Metformin Hydrochloride. *J. Pharm. Biomed. Anal.* **2009**, *49* (4), 976–982.

(57) Kolwijck, E.; Engelke, U. F.; van der Graaf, M.; Heerschap, A.; Blom, H. J.; Hadfoune, M.; Buurman, W. A.; Massuger, L. F.; Wevers, R. A. *N* -Acetyl Resonances in *in Vivo* and *in Vitro* NMR Spectroscopy of Cystic Ovarian Tumors. *NMR Biomed.* **2009**, n/a-n/a.

(58) Lauridsen, M. B.; Bliddal, H.; Christensen, R.; Danneskiold-Samsøe, B.; Bennett, R.; Keun, H.; Lindon, J. C.; Nicholson, J. K.; Dorff, M. H.; Jaroszewski, J. W.; et al. ^{1} H NMR Spectroscopy-Based Interventional Metabolic Phenotyping: A Cohort Study of Rheumatoid Arthritis Patients. *J. Proteome Res.* **2010**, *9* (9), 4545–4553.

(59) Wu, Q.; Zhang, Q.; Sun, B.; Yan, X.; Tang, Y.; Qiao, X.; Chen, Q.; Yu, S.; Liang, F. 1H NMR-Based Metabonomic Study on the Metabolic Changes in the Plasma of Patients with Functional

Dyspepsia and the Effect of Acupuncture. *J. Pharm. Biomed. Anal.* **2010**, *51* (3), 698–704.

(60) Hasim, A.; Ma, H.; Mamtimin, B.; Abudula, A.; Niyaz, M.; Zhang, L.; Anwer, J.; Sheyhidin, I. Revealing the Metabonomic Variation of EC Using 1H-NMR Spectroscopy and Its Association with the Clinicopathological Characteristics. *Mol. Biol. Rep.* **2012**, *39* (9), 8955–8964.

(61) Hasim, A.; Ali, M.; Mamtimin, B.; Ma, J.; Li, G.; Abudula, A. Metabonomic Signature Analysis of Cervical Carcinoma and Precancerous Lesions in Women by 1H NMR Spectroscopy. *Exp. Ther. Med.* **2012**, *3* (6), 945–951.

(62) Li, P.; Tao, J.; Wei, D.; Yang, X.; Lu, Z.; Deng, X.; Cheng, Y.; Gu, J.; Yang, X.; Wang, Z.; et al. Serum Metabolomic Analysis of Human Upper Urinary Tract Urothelial Carcinoma. *Tumor Biol.* **2015**, *36* (10), 7531–7537.

(63) Boguszewicz, Ł.; Hajduk, A.; Mrochem-Kwarciak, J.; Skorupa, A.; Ciszek, M.; Heyda, A.; Składowski, K.; Sokół, M. 1H NMR Based Metabolomic Approach to Monitoring of the Head and Neck Cancer Treatment Toxicity. *Metabolomics* **2016**, *12* (6), 102.

(64) Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W. Quantification of Plasma Lipoproteins by Proton Nuclear Magnetic Resonance Spectroscopy. *Clin. Chem.* **1991**, *37* (3), 377–386.

(65) Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W.; Krauss, R. M. Development of a Proton Nuclear Magnetic Resonance Spectroscopic Method for Determining Plasma Lipoprotein Concentrations and Subspecies Distributions from a Single, Rapid Measurement. *Clin. Chem.* **1992**, *38* (9), 1632–1638.

(66) Soininen, P.; Kangas, A. J.; Würtz, P.; Tukiainen, T.; Tynkkynen, T.; Laatikainen, R.; Järvelin, M.-R.; Kähönen, M.; Lehtimäki, T.; Viikari, J.; et al. High-Throughput Serum NMR Metabonomics for Cost-Effective Holistic Studies on Systemic Metabolism †. *Analyst* **2009**, *134* (9), 1781–1785.

(67) Ala-Korpela, M. 1H NMR Spectroscopy of Human Blood Plasma. *Prog. Nucl. Magn. Reson. Spectrosc.* **1995**, *27* (5–6), 475–554.

(68) Kujala, U. M.; Mäkinen, V.-P.; Heinonen, I.; Soininen, P.; Kangas, A. J.; Leskinen, T. H.; Rahkila, P.; Würtz, P.; Kovanen, V.; Cheng, S.; et al. Long-Term Leisure-Time Physical Activity and Serum Metabolome. *Circulation* **2013**, *127* (3), 340–348.

(69) Correia, G. D. S.; Wooi Ng, K.; Wijeyesekera, A.; Gala-Peralta, S.; Williams, R.; MacCarthy-Morrogh, S.; Jiménez, B.; Inwald, D.; Macrae, D.; Frost, G.; et al. Metabolic Profiling of Children Undergoing Surgery for Congenital Heart Disease. *Crit. Care Med.* **2015**, *43* (7), 1467–1476.

(70) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6* (3), 277–293.

(71) Lécuyer, L.; Victor Bala, A.; Deschasaux, M.; Bouchemal, N.; Nawfal Triba, M.; Vasson, M.-P.; Rossary, A.; Demidem, A.; Galan, P.; Hercberg, S.; et al. NMR Metabolomic Signatures Reveal Predictive Plasma Metabolites Associated with Long-Term Risk of Developing Breast Cancer. *Int. J. Epidemiol.* **2018**, *47* (2), 484–494.

(72) Jobard, E.; Pontoizeau, C.; Blaise, B. J.; Bachelot, T.; Elena-Herrmann, B.; Trédan, O. A Serum Nuclear Magnetic Resonance-Based Metabolomic Signature of Advanced Metastatic Human Breast Cancer. *Cancer Lett.* **2014**, *343* (1), 33–41.

(73) Deja, S.; Porebska, I.; Kowal, A.; Zabek, A.; Barg, W.; Pawelczyk, K.; Stanimirova, I.; Daszykowski, M.; Korzeniewska, A.; Jankowska, R.; et al. Metabolomics Provide New Insights on Lung Cancer Staging and Discrimination from Chronic Obstructive Pulmonary Disease. *J. Pharm. Biomed. Anal.* **2014**, *100*, 369–380.

(74) Huo, T.; Chen, X.; Lu, X.; Qu, L.; Liu, Y.; Cai, S. An Effective Assessment of Valproate Sodium-Induced Hepatotoxicity with

UPLC–MS and 1HNMR-Based Metabonomics Approach. *J. Chromatogr. B* **2014**, *969*, 109–116.

(75) Fages, A.; Duarte-Salles, T.; Stepien, M.; Ferrari, P.; Fedirko, V.; Pontoizeau, C.; Trichopoulou, A.; Aleksandrova, K.; Tjønneland, A.; Olsen, A.; et al. Metabolomic Profiles of Hepatocellular Carcinoma in a European Prospective Cohort. *BMC Med.* **2015**, *13* (1), 242.

(76) Sands, C. J.; Guha, I. N.; Kyriakides, M.; Wright, M.; Beckonert, O.; Holmes, E.; Rosenberg, W. M.; Coen, M. Metabolic Phenotyping for Enhanced Mechanistic Stratification of Chronic Hepatitis C-Induced Liver Fibrosis. *Am. J. Gastroenterol.* **2015**, *110* (1), 159–169.

(77) Tian, J.; Xia, X.; Wu, Y.; Zhao, L.; Xiang, H.; Du, G.; Zhang, X.; Qin, X. Discovery, Screening and Evaluation of a Plasma Biomarker Panel for Subjects with Psychological Suboptimal Health State Using 1H-NMR-Based Metabolomics Profiles. *Sci. Rep.* **2016**, *6* (1), 33820.

(78) Jain, A.; Kumar, D.; Guleria, A.; Misra, D. P.; Zanwar, A.; Chaurasia, S.; Kumar, S.; Kumar, U.; Mishra, S. K.; Goel, R.; et al. NMR-Based Serum Metabolomics of Patients with Takayasu Arteritis: Relationship with Disease Activity. *J. Proteome Res.* **2018**, *17* (9), 3317–3324.

(79) Suman, S.; Sharma, R. K.; Kumar, V.; Sinha, N.; Shukla, Y. Metabolic Fingerprinting in Breast Cancer Stages through 1H NMR Spectroscopy-Based Metabolomic Analysis of Plasma. *J. Pharm. Biomed. Anal.* **2018**, *160*, 38–45.

(80) Olivier Cloarec, †; Marc-Emmanuel Dumas, †; Andrew Craig, †; Richard H. Barton, †; Johan Trygg, ‡; Jane Hudson, §; Christine Blancher, §; Dominique Gauguier, §; John C. Lindon, †; Elaine Holmes, † and; et al. Statistical Total Correlation Spectroscopy: An Exploratory Approach for Latent Biomarker Identification from Metabolic 1H NMR Data Sets. *Anal. Chem* **2005**, *77* (5), 1282–1289.

(81) Pinto, J.; Barros, A. S.; Domingues, M. R. M.; Goodfellow, B.

J.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Following Healthy Pregnancy by NMR Metabolomics of Plasma and Correlation to Urine. *J. Proteome Res.* **2015**, *14* (2), 1263–1274.

(82) Connelly, M. A.; Gruppen, E. G.; Wolak-Dinsmore, J.; Matyus, S. P.; Riphagen, I. J.; Shalaurova, I.; Bakker, S. J. L.; Otvos, J. D.; Dullaart, R. P. F. GlycA, a Marker of Acute Phase Glycoproteins, and the Risk of Incident Type 2 Diabetes Mellitus: PREVEND Study. *Clin. Chim. Acta* **2016**, *452*, 10–17.

(83) Nicholson, J. K.; Buckingham, M. J.; Sadler, P. J. High Resolution 1H NMR Studies of Vertebrate Blood and Plasma. *Biochem. J.* **1983**, *211* (3), 605–615.

(84) Nicholson, J. K.; O'Flynn, M. P.; Sadler, P. J.; Macleod, A. F.; Juul, S. M.; Sönksen, P. H. Proton-Nuclear-Magnetic-Resonance Studies of Serum, Plasma and Urine from Fasting Normal and Diabetic Subjects. *Biochem. J.* **1984**, *217* (2), 365–375.

(85) Stowell, S. R.; Ju, T.; Cummings, R. D. Protein Glycosylation in Cancer. *Annu. Rev. Pathol. Mech. Dis.* **2015**, *10* (1), 473–510.

(86) Tiziani, S.; Lopes, V.; Günther, U. L. Early Stage Diagnosis of Oral Cancer Using 1 H NMR-Based Metabolomics 1,2. *Neoplasia* **2009**, *11*, 269–276.

(87) Boss, E. A.; Moolenaar, S. H.; Massuger, L. F. A. G.; Boonstra, H.; Engelke, U. F. H.; De Jong, J. G. N.; Wevers, R. A. *High-Resolution Proton Nuclear Magnetic Resonance Spectroscopy of Ovarian Cyst ¯uid*; 2000.

(88) Chandler, P. D.; Akinkuolie, A. O.; Tobias, D. K.; Lawler, P. R.; Li, C.; Moorthy, M. V.; Wang, L.; Duprez, D. A.; Jacobs, D. R.; Glynn, R. J.; et al. Association of N-Linked Glycoprotein Acetyls and Colorectal Cancer Incidence and Mortality. *PLoS One* **2016**, *11* (11), e0165615.

(89) Duprez, D. A.; Otvos, J.; Sanchez, O. A.; Mackey, R. H.; Tracy, R.; Jacobs, D. R. Comparison of the Predictive Value of GlycA and Other Biomarkers of Inflammation for Total Death, Incident

Cardiovascular Events, Noncardiovascular and Noncancer Inflammatory-Related Events, and Total Cancer Events. *Clin. Chem.* **2016**, *62* (7), 1020–1031.

(90) Kelimu, A.; Xie, R.; Zhang, K.; Zhuang, Z.; Mamtimin, B.; Sheyhidin, I. Metabonomic Signature Analysis in Plasma Samples of Glioma Patients Based on ^{1}H-Nuclear Magnetic Resonance Spectroscopy. *Neurol. India* **2016**, *64* (2), 246.

(91) López-Jaramillo, P.; Gómez-Arbeláez, D.; López-López, J.; López-López, C.; Martínez-Ortega, J.; Gómez-Rodríguez, A.; Triana-Cubillos, S. The Role of Leptin/Adiponectin Ratio in Metabolic Syndrome and Diabetes. *Horm. Mol. Biol. Clin. Investig.* **2014**, *18* (1), 37–45.

(92) Finucane, F. M.; Luan, J.; Wareham, N. J.; Sharp, S. J.; O'Rahilly, S.; Balkau, B.; Flyvbjerg, A.; Walker, M.; Højlund, K.; Nolan, J. J.; et al. Correlation of the Leptin:Adiponectin Ratio with Measures of Insulin Resistance in Non-Diabetic Individuals. *Diabetologia* **2009**, *52* (11), 2345–2349.

(93) Dullaart, R. P. F.; Gruppen, E. G.; Connelly, M. A.; Lefrandt, J. D. A Pro-Inflammatory Glycoprotein Biomarker Is Associated with Lower Bilirubin in Metabolic Syndrome. *Clin. Biochem.* **2015**, *48* (16–17), 1045–1047.

(94) Dullaart, R. P. F.; Gruppen, E. G.; Connelly, M. A.; Otvos, J. D.; Lefrandt, J. D. GlycA, a Biomarker of Inflammatory Glycoproteins, Is More Closely Related to the Leptin/Adiponectin Ratio than to Glucose Tolerance Status. *Clin. Biochem.* **2015**, *48* (12), 811–814.

(95) Bogl, L. H.; Kaye, S. M.; Rämö, J. T.; Kangas, A. J.; Soininen, P.; Hakkarainen, A.; Lundbom, J.; Lundbom, N.; Ortega-Alonso, A.; Rissanen, A.; et al. Abdominal Obesity and Circulating Metabolites: A Twin Study Approach. *Metabolism.* **2016**, *65* (3), 111–121.

(96) Mokkala, K.; Pellonperä, O.; Röytiö, H.; Pussinen, P.; Rönnemaa, T.; Laitinen, K. Increased Intestinal Permeability, Measured by Serum Zonulin, Is Associated with Metabolic Risk

Markers in Overweight Pregnant Women. *Metabolism* **2017**, *69*, 43–50.

(97) Newgard, C. B.; An, J.; Bain, J. R.; Muehlbauer, M. J.; Stevens, R. D.; Lien, L. F.; Haqq, A. M.; Shah, S. H.; Arlotto, M.; Slentz, C. A.; et al. A Branched-Chain Amino Acid-Related Metabolic Signature That Differentiates Obese and Lean Humans and Contributes to Insulin Resistance. *Cell Metab.* **2009**, *9* (4), 311–326.

(98) Wang, T. J.; Larson, M. G.; Vasan, R. S.; Cheng, S.; Rhee, E. P.; McCabe, E.; Lewis, G. D.; Fox, C. S.; Jacques, P. F.; Fernandez, C.; et al. Metabolite Profiles and the Risk of Developing Diabetes. *Nat. Med.* **2011**, *17* (4), 448–453.

(99) Lorenzo, C.; Festa, A.; Hanley, A. J.; Rewers, M. J.; Escalante, A.; Haffner, S. M. Novel Protein Glycan–Derived Markers of Systemic Inflammation and C-Reactive Protein in Relation to Glycemia, Insulin Resistance, and Insulin Secretion. *Diabetes Care* **2017**, *40* (3), 375–382.

(100) Bartlett, D. B.; Slentz, C. A.; Connelly, M. A.; Piner, L. W.; Willis, L. H.; Bateman, L. A.; Granville, E. O.; Bales, C. W.; Huffman, K. M.; Kraus, W. E. Association of the Composite Inflammatory Biomarker GlycA, with Exercise-Induced Changes in Body Habitus in Men and Women with Prediabetes. *Oxid. Med. Cell. Longev.* **2017**, *2017*, 1–12.

(101) Manmadhan, A.; Lin, B.-X.; Zhong, J.; Parikh, M.; Berger, J. S.; Fisher, E. A.; Heffron, S. P. Elevated GlycA in Severe Obesity Is Normalized by Bariatric Surgery. *Diabetes, Obes. Metab.* **2019**, *21* (1), 178–182.

(102) Jago, R.; Drews, K. L.; Otvos, J. D.; Willi, S. M.; Buse, J. B. Novel Measures of Inflammation and Insulin Resistance Are Related to Obesity and Fitness in a Diverse Sample of 11–14 Year Olds: The HEALTHY Study. *Int. J. Obes.* **2016**, *40* (7), 1157–1163.

(103) Olson, M. L.; Rentería-Mexía, A.; Connelly, M. A.; Vega-López, S.; Soltero, E. G.; Konopken, Y. P.; Williams, A. N.;

Castro, F. G.; Keller, C. S.; Yang, H. P.; et al. Decreased GlycA after Lifestyle Intervention among Obese, Prediabetic Adolescent Latinos. *J. Clin. Lipidol.* **2018**.

(104) Tabish, S. A. Is Diabetes Becoming the Biggest Epidemic of the Twenty-First Century? *Int. J. Health Sci. (Qassim).* **2007**, *1* (2), V–VIII.

(105) Duncan, B. B.; Schmidt, M. I.; Pankow, J. S.; Ballantyne, C. M.; Couper, D.; Vigo, A.; Hoogeveen, R.; Folsom, A. R.; Heiss, G.; Atherosclerosis Risk in Communities Study. Low-Grade Systemic Inflammation and the Development of Type 2 Diabetes: The Atherosclerosis Risk in Communities Study. *Diabetes* **2003**, *52* (7), 1799–1805.

(106) Hotamisligil, G. S. Inflammation and Metabolic Disorders. *Nature* **2006**, *444* (7121), 860–867.

(107) Karstoft, K.; Pedersen, B. K. Exercise and Type 2 Diabetes: Focus on Metabolism and Inflammation. *Immunol. Cell Biol.* **2016**, *94* (2), 146–150.

(108) Pradhan, A. D.; Manson, J. E.; Rifai, N.; Buring, J. E.; Ridker, P. M. C-Reactive Protein, Interleukin 6, and Risk of Developing Type 2 Diabetes Mellitus. *JAMA* **2001**, *286* (3), 327–334.

(109) Badawi, A.; Klip, A.; Haddad, P.; Cole, D. E.; Bailo, B. G.; El-Sohemy, A.; Karmali, M. Type 2 Diabetes Mellitus and Inflammation: Prospects for Biomarkers of Risk and Nutritional Intervention. *Diabetes. Metab. Syndr. Obes.* **2010**, *3*, 173–186.

(110) Wurtz, P.; Tiainen, M.; Makinen, V.-P.; Kangas, A. J.; Soininen, P.; Saltevo, J.; Keinanen-Kiukaanniemi, S.; Mantyselka, P.; Lehtimaki, T.; Laakso, M.; et al. Circulating Metabolite Predictors of Glycemia in Middle-Aged Men and Women. *Diabetes Care* **2012**, *35* (8), 1749–1756.

(111) Akinkuolie, A. O.; Pradhan, A. D.; Buring, J. E.; Ridker, P. M.; Mora, S. Novel Protein Glycan Side-Chain Biomarker and Risk of Incident Type 2 Diabetes Mellitus. *Arterioscler. Thromb. Vasc. Biol.* **2015**, *35* (6), 1544–1550.

(112) Dungan, K.; Binkley, P.; Osei, K. GlycA Is a Novel Marker of Inflammation Among Non-Critically Ill Hospitalized Patients with Type 2 Diabetes. *Inflammation* **2015**, *38* (3), 1357–1363.

(113) Fizelova, M.; Jauhiainen, R.; Kangas, A. J.; Soininen, P.; Ala-Korpela, M.; Kuusisto, J.; Laakso, M.; Stančáková, A. Differential Associations of Inflammatory Markers With Insulin Sensitivity and Secretion: The Prospective METSIM Study. *J. Clin. Endocrinol. Metab.* **2017**, *102* (9), 3600–3609.

(114) Rawat, A.; Misra, G.; Saxena, M.; Tripathi, S.; Dubey, D.; Saxena, S.; Aggarwal, A.; Gupta, V.; Khan, M. Y.; Prakash, A. 1H NMR Based Serum Metabolic Profiling Reveals Differentiating Biomarkers in Patients with Diabetes and Diabetes-Related Complication. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2019**, *13* (1), 290–298.

(115) Gruppen, E. G.; Connelly, M. A.; Otvos, J. D.; Bakker, S. J. L.; Dullaart, R. P. F. A Novel Protein Glycan Biomarker and LCAT Activity in Metabolic Syndrome. *Eur. J. Clin. Invest.* **2015**, *45* (8), 850–859.

(116) Gruppen, E. G.; Connelly, M. A.; Dullaart, R. P. F. Higher Circulating GlycA, a pro-Inflammatory Glycoprotein Biomarker, Relates to Lipoprotein-Associated Phospholipase A2 Mass in Nondiabetic Subjects but Not in Diabetic or Metabolic Syndrome Subjects. *J. Clin. Lipidol.* **2016**, *10* (3), 512–518.

(117) Sutherland, J. P.; McKinley, B.; Eckel, R. H. The Metabolic Syndrome and Inflammation. *Metab. Syndr. Relat. Disord.* **2004**, *2* (2), 82–104.

(118) Esser, N.; Legrand-Poels, S.; Piette, J.; Scheen, A. J.; Paquot, N. Inflammation as a Link between Obesity, Metabolic Syndrome and Type 2 Diabetes. *Diabetes Res. Clin. Pract.* **2014**, *105*, 141–150.

(119) Gruppen, E. G.; Riphagen, I. J.; Connelly, M. A.; Otvos, J. D.; Bakker, S. J. L.; Dullaart, R. P. F. GlycA, a Pro-Inflammatory Glycoprotein Biomarker, and Incident Cardiovascular Disease: Relationship with C-Reactive Protein and Renal Function. *PLoS*

One **2015**, *10* (9), e0139057.

(120) Ridker, P. M.; Hennekens, C. H.; Buring, J. E.; Rifai, N. C-Reactive Protein and Other Markers of Inflammation in the Prediction of Cardiovascular Disease in Women. *N. Engl. J. Med.* **2000**, *342* (12), 836–843.

(121) Bassuk, S. S.; Rifai, N.; Ridker, P. M. High-Sensitivity C-Reactive Protein: Clinical Importance. *Curr. Probl. Cardiol.* **2004**, *29* (8), 439–493.

(122) Harris, T. B.; Ferrucci, L.; Tracy, R. P.; Corti, M. C.; Wacholder, S.; Ettinger, W. H.; Heimovitz, H.; Cohen, H. J.; Wallace, R. Associations of Elevated Interleukin-6 and C-Reactive Protein Levels with Mortality in the Elderly. *Am. J. Med.* **1999**, *106* (5), 506–512.

(123) Fischer, K.; Kettunen, J.; Würtz, P.; Haller, T.; Havulinna, A. S.; Kangas, A. J.; Soininen, P.; Esko, T.; Tammesoo, M.-L.; Mägi, R.; et al. Biomarker Profiling by Nuclear Magnetic Resonance Spectroscopy for the Prediction of All-Cause Mortality: An Observational Study of 17,345 Persons. *PLoS Med.* **2014**, *11* (2), e1001606.

(124) Akinkuolie, A. O.; Buring, J. E.; Ridker, P. M.; Mora, S. A Novel Protein Glycan Biomarker and Future Cardiovascular Disease Events. *J. Am. Heart Assoc.* **2014**, *3* (5), e001221.

(125) Lawler, P. R.; Akinkuolie, A. O.; Chandler, P. D.; Moorthy, M. V.; Vandenburgh, M. J.; Schaumberg, D. A.; Lee, I.-M.; Glynn, R. J.; Ridker, P. M.; Buring, J. E.; et al. Circulating N-Linked Glycoprotein Acetyls and Longitudinal Mortality Risk. *Circ. Res.* **2016**, *118* (7), 1106–1115.

(126) Benson, E.-M. A.; Tibuakuu, M.; Zhao, D.; Akinkuolie, A. O.; Otvos, J. D.; Duprez, D. A.; Jacobs, D. R.; Mora, S.; Michos, E. D. Associations of Ideal Cardiovascular Health with GlycA, a Novel Inflammatory Marker: The Multi-Ethnic Study of Atherosclerosis. *Clin. Cardiol.* **2018**, *41* (11), 1439–1445.

(127) Holmes, M. V.; Millwood, I. Y.; Kartsonaki, C.; Hill, M. R.;

Bennett, D. A.; Boxall, R.; Guo, Y.; Xu, X.; Bian, Z.; Hu, R.; et al. Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *J. Am. Coll. Cardiol.* **2018**, *71* (6), 620–632.

(128) McGarrah, R.; Craig, D.; Haynes, C.; Dowdy, Z. E.; Shah, S.; Kraus, W. GlycA, a Novel Biomarker of Systemic Inflammation, Improves Cardiovascular Risk Prediction in a High-Risk Coronary Catheterization Cohort. *J. Am. Coll. Cardiol.* **2015**, *65* (10), A1606.

(129) Muhlestein, J. B.; May, H.; Winegar, D.; Rollo, J.; Connelly, M.; Otvos, J.; Anderson, J. GlycA and GlycB, Novel NMR Biomarkers of Inflammation, Strongly Predict Future Cardiovascular Events, but Not the Presence of Coronary Artery Disease (CAD), among Patients Undergoing Coronary Angiography: The Intermountain Heart Collaborative Study. *J. Am. Coll. Cardiol.* **2014**, *63* (12), A1389.

(130) Muhlestein, J. B.; May, H.; Winegar, D.; Rollo, J.; Connelly, M.; Otvos, J.; Anderson, J. Differential Association of High-Density Lipoprotein Particle Subclasses and GlycA, a Novel Inflammatory Marker, in Predicting Cardiac Death among Patients Undergoing Angiography: The Intermountain Heart Collaborative Study. *J. Am. Coll. Cardiol.* **2016**, *67* (13), 162.

(131) Muhlestein, J. B.; May, H. T.; Galenko, O.; Knowlton, K. U.; Otvos, J. D.; Connelly, M. A.; Lappe, D. L.; Anderson, J. L. GlycA and HsCRP Are Independent and Additive Predictors of Future Cardiovascular Events among Patients Undergoing Angiography: The Intermountain Heart Collaborative Study. *Am. Heart J.* **2018**, *202*, 27–32.

(132) McGarrah, R. W.; Kelly, J. P.; Craig, D. M.; Haynes, C.; Jessee, R. C.; Huffman, K. M.; Kraus, W. E.; Shah, S. H. A Novel Protein Glycan–Derived Inflammation Biomarker Independently Predicts Cardiovascular Disease and Modifies the Association of HDL Subclasses with Mortality. *Clin. Chem.* **2017**, *63* (1), 288–296.

(133) Otvos, J. D.; Guyton, J. R.; Connelly, M. A.; Akapame, S.;

Bittner, V.; Kopecky, S. L.; Lacy, M.; Marcovina, S. M.; Muhlestein, J. B.; Boden, W. E. Relations of GlycA and Lipoprotein Particle Subspecies with Cardiovascular Events and Mortality: A Post Hoc Analysis of the AIM-HIGH Trial. *J. Clin. Lipidol.* **2018**.

(134) Gruppen, E. G.; Connelly, M. A.; Sluiter, W. J.; Bakker, S. J. L.; Dullaart, R. P. F. Higher Plasma GlycA, a Novel pro-Inflammatory Glycoprotein Biomarker, Is Associated with Reduced Life Expectancy: The PREVEND Study. *Clin. Chim. Acta* **2019**, *488*, 7–12.

(135) Freiberg, M. S.; Chang, C.-C. H.; Kuller, L. H.; Skanderson, M.; Lowy, E.; Kraemer, K. L.; Butt, A. A.; Bidwell Goetz, M.; Leaf, D.; Oursler, K. A.; et al. HIV Infection and the Risk of Acute Myocardial Infarction. *JAMA Intern. Med.* **2013**, *173* (8), 614.

(136) Triant, V. A.; Lee, H.; Hadigan, C.; Grinspoon, S. K. Increased Acute Myocardial Infarction Rates and Cardiovascular Risk Factors among Patients with Human Immunodeficiency Virus Disease. *J. Clin. Endocrinol. Metab.* **2007**, *92* (7), 2506–2512.

(137) Tibuakuu, M.; Fashanu, O. E.; Bs, M. B.; Zhao, D.; Otvos, J. D.; Brown, T. T.; Haberlen, S. A.; Guallar, E.; Budoff, M. J.; Palella, F. J.; et al. GlycA, a Novel Inflammatory Marker, Is Associated with Subclinical Coronary Disease in the Multicenter AIDS Cohort Study Short Title: GlycA and Coronary Plaque in HIV. *UCLA Previously Publ. Work.* **2018**.

(138) Kelesidis, T.; Tran, T. T. T.; Stein, J. H.; Brown, T. T.; Moser, C.; Ribaudo, H. J.; Dube, M. P.; Murphy, R.; Yang, O. O.; Currier, J. S.; et al. Changes in Inflammation and Immune Activation With Atazanavir-, Raltegravir-, Darunavir-Based Initial Antiviral Therapy: ACTG 5260s. *Clin. Infect. Dis.* **2015**, *61* (4), 651–660.

(139) Bartlett, D. B.; Connelly, M. A.; AbouAssi, H.; Bateman, L. A.; Tune, K. N.; Huebner, J. L.; Kraus, V. B.; Winegar, D. A.; Otvos, J. D.; Kraus, W. E.; et al. A Novel Inflammatory Biomarker, GlycA, Associates with Disease Activity in Rheumatoid Arthritis

and Cardio-Metabolic Risk in BMI-Matched Controls. *Arthritis Res. Ther.* **2016**, *18* (1), 86.

(140) Ormseth, M. J.; Chung, C. P.; Oeser, A. M.; Connelly, M. A.; Sokka, T.; Raggi, P.; Solus, J. F.; Otvos, J. D.; Stein, C. M. Utility of a Novel Inflammatory Marker, GlycA, for Assessment of Rheumatoid Arthritis Disease Activity and Coronary Atherosclerosis. *Arthritis Res. Ther.* **2015**, *17* (1), 117.

(141) Dierckx, T.; Goletti, S.; Chiche, L.; Daniel, L.; Lauwerys, B.; Jourde-Chiche, N.; Weyenbergh, J. Van. Serum GlycA Level Is a Candidate Biomarker for Disease Activity in Systemic Lupus Erythematosus and for Proliferative Status of Lupus Nephritis, Independent of Renal Function Impairment. *bioRxiv* **2018**, 493809.

(142) Durcan, L.; Winegar, D. A.; Connelly, M. A.; Otvos, J. D.; Magder, L. S.; Petri, M. Longitudinal Evaluation of Lipoprotein Variables in Systemic Lupus Erythematosus Reveals Adverse Changes with Disease Activity and Prednisone and More Favorable Profiles with Hydroxychloroquine Therapy. *J. Rheumatol.* **2016**, *43* (4), 745–750.

(143) Chung, C. P.; Ormseth, M. J.; Connelly, M. A.; Oeser, A.; Solus, J. F.; Otvos, J. D.; Raggi, P.; Stein, C. M. GlycA, a Novel Marker of Inflammation, Is Elevated in Systemic Lupus Erythematosus. *Lupus* **2016**, *25* (3), 296–300.

(144) Mehta, N. N.; Yu, Y.; Saboury, B.; Foroughi, N.; Krishnamoorthy, P.; Raper, A.; Baer, A.; Antigua, J.; Van Voorhees, A. S.; Torigian, D. A.; et al. Systemic and Vascular Inflammation in Patients With Moderate to Severe Psoriasis as Measured by [18F]-Fluorodeoxyglucose Positron Emission Tomography –Computed Tomography (FDG-PET/CT). *Arch. Dermatol.* **2011**, *147* (9), 1031.

(145) Mehta, N. N.; Azfar, R. S.; Shin, D. B.; Neimann, A. L.; Troxel, A. B.; Gelfand, J. M. Patients with Severe Psoriasis Are at Increased Risk of Cardiovascular Mortality: Cohort Study Using the General Practice Research Database. *Eur. Heart J.* **2010**, *31* (8), 1000–1006.

(146) Joshi, A. A.; Lerman, J. B.; Aberra, T. M.; Afshar, M.; Teague, H. L.; Rodante, J. A.; Krishnamoorthy, P.; Ng, Q.; Aridi, T. Z.; Salahuddin, T.; et al. GlycA Is a Novel Biomarker of Inflammation and Subclinical Cardiovascular Disease in Psoriasis. *Circ. Res.* **2016**, *119* (11).

(147) Kaplan, G. G. The Global Burden of IBD: From 2015 to 2025. *Nat. Rev. Gastroenterol. Hepatol.* **2015**, *12* (12), 720–727.

(148) Zhang, Y.-Z.; Li, Y.-Y. Inflammatory Bowel Disease: Pathogenesis. *World J. Gastroenterol.* **2014**, *20* (1), 91.

(149) Dierckx, T.; Verstockt, B.; Vermeire, S.; van Weyenbergh, J. GlycA, a Nuclear Magnetic Resonance Spectroscopy Measure for Protein Glycosylation, Is a Viable Biomarker for Disease Activity in IBD. *J. Crohn's Colitis* **2018**, 1–6.

(150) Austin, V.; Crack, P. J.; Bozinovski, S.; Miller, A. A.; Vlahos, R. COPD and Stroke: Are Systemic Inflammation and Oxidative Stress the Missing Links? *Clin. Sci. (Lond).* **2016**, *130* (13), 1039–1050.

(151) Kettunen, J.; Ritchie, S.; Anufrieva, O.; Lyytikainen, L.-P.; Hernesniemi, J.; Karhunen, P. J.; Kuukasjarvi, P.; Laurikka, J.; Kahonen, M.; Lehtimaki, T.; et al. The Landscape of Incident Disease Risk for the Biomarker GlycA and Its Mortality Stratification in Angiography Patients. *bioRxiv* **2018**, 280677.

(152) Xu, W.; Upur, H.; Wu, Y.; Mamtimin, B.; Yang, J.; Ga, Y.; You, L. Metabolomic Changes in Patients with Chronic Obstructive Pulmonary Disease with Abnormal Savda Syndrome. *Exp. Ther. Med.* **2015**, *9* (2), 425–431.

(153) Ablimit, A.; Kühnel, H.; Strasser, A.; Upur, H. Abnormal Savda Syndrome: Long-Term Consequences of Emotional and Physical Stress on Endocrine and Immune Activities in an Animal Model. *Chin. J. Integr. Med.* **2013**, *19* (8), 603–609.

(154) Gupta, J.; Mitra, N.; Kanetsky, P. A.; Devaney, J.; Wing, M. R.; Reilly, M.; Shah, V. O.; Balakrishnan, V. S.; Guzman, N. J.; Girndt, M.; et al. Association between Albuminuria, Kidney

Function, and Inflammatory Biomarker Profile in CKD in CRIC. *Clin. J. Am. Soc. Nephrol.* **2012**, *7* (12), 1938–1946.

(155) Upadhyay, A.; Larson, M. G.; Guo, C.-Y.; Vasan, R. S.; Lipinska, I.; O'Donnell, C. J.; Kathiresan, S.; Meigs, J. B.; Keaney, J. F.; Rong, J.; et al. Inflammation, Kidney Function and Albuminuria in the Framingham Offspring Cohort. *Nephrol. Dial. Transplant.* **2011**, *26* (3), 920–926.

(156) Lee, B. T.; Ahmed, F. A.; Hamm, L. L.; Teran, F. J.; Chen, C.-S.; Liu, Y.; Shah, K.; Rifai, N.; Batuman, V.; Simon, E. E.; et al. Association of C-Reactive Protein, Tumor Necrosis Factor-Alpha, and Interleukin-6 with Chronic Kidney Disease. *BMC Nephrol.* **2015**, *16* (1), 77.

(157) Titan, S. M.; Pecoits-Filho, R.; Barreto, S. M.; Lopes, A. A.; Bensenor, I. J.; Lotufo, P. A. GlycA, a Marker of Protein Glycosylation, Is Related to Albuminuria and Estimated Glomerular Filtration Rate: The ELSA-Brasil Study. *BMC Nephrol.* **2017**, *18* (1), 367.

(158) Cohen-Manheim, I.; Doniger, G. M.; Sinnreich, R.; Simon, E. S.; Pinchas-Mizrachi, R.; Otvos, J. D.; Kark, J. D. Increase in the Inflammatory Marker GlycA over 13 Years in Young Adults Is Associated with Poorer Cognitive Function in Midlife. *PLoS One* **2015**, *10* (9), e0138036.

(159) Tukiainen, T.; Tynkkynen, T.; Mäkinen, V.-P.; Jylänki, P.; Kangas, A.; Hokkanen, J.; Vehtari, A.; Gröhn, O.; Hallikainen, M.; Soininen, H.; et al. A Multi-Metabolite Analysis of Serum by 1H NMR Spectroscopy: Early Systemic Signs of Alzheimer's Disease. *Biochem. Biophys. Res. Commun.* **2008**, *375* (3), 356–361.

(160) Tombetti, E.; Mason, J. C. Takayasu Arteritis: Advanced Understanding Is Leading to New Horizons. *Rheumatology* **2019**, *58* (2), 206–219.

(161) Guleria, A.; Misra, D. P.; Rawat, A.; Dubey, D.; Khetrapal, C. L.; Bacon, P.; Misra, R.; Kumar, D. NMR-Based Serum Metabolomics Discriminates Takayasu Arteritis from Healthy

Individuals: A Proof-of-Principle Study. *J. Proteome Res.* **2015**, *14* (8), 3372–3381.

(162) Gordon, J. B.; Kahn, A. M.; Burns, J. C. When Children With Kawasaki Disease Grow Up. *J. Am. Coll. Cardiol.* **2009**, *54* (21), 1911–1920.

(163) Connelly, M. A.; Shimizu, C.; Winegar, D. A.; Shalaurova, I.; Pourfarzib, R.; Otvos, J. D.; Kanegaye, J. T.; Tremoulet, A. H.; Burns, J. C. Differences in GlycA and Lipoprotein Particle Parameters May Help Distinguish Acute Kawasaki Disease from Other Febrile Illnesses in Children. *BMC Pediatr.* **2016**, *16* (1).

(164) Houttu, N.; Mokkala, K.; Laitinen, K. Overweight and Obesity Status in Pregnant Women Are Related to Intestinal Microbiota and Serum Metabolic and Inflammatory Profiles. *Clin. Nutr.* **2018**, *37* (6), 1955–1966.

(165) Vilela, L. A. P.; Almeida, M. Q. Diagnosis and Management of Primary Aldosteronism. *Arch. Endocrinol. Metab.* **2017**, *61* (3), 305–312.

(166) Brown, N. J. Contribution of Aldosterone to Cardiovascular and Renal Inflammation and Fibrosis. *Nat. Rev. Nephrol.* **2013**, *9* (8), 459–469.

(167) Berends, A. M. A.; Buitenwerf, E.; Gruppen, E. G.; Sluiter, W. J.; Bakker, S. J. L.; Connelly, M. A.; Kerstens, M. N.; Dullaart, R. P. F. Primary Aldosteronism Is Associated with Decreased Low-Density and High-Density Lipoprotein Particle Concentrations and Increased GlycA, a pro-Inflammatory Glycoprotein Biomarker. *Clin. Endocrinol. (Oxf).* **2019**, *90* (1), 79–87.

(168) Platt, O. S. Sickle Cell Anemia as an Inflammatory Disease. *J. Clin. Invest.* **2000**, *106* (3), 337–338.

(169) Weisman, J. K.; Meeks, D.; Mendelsohn, L.; Remaley, A. T.; Sampson, M.; Allen, D. T.; Nichols, J.; Shet, A. S.; Thein, S. L. GlycA Is Not a Useful Biomarker of Inflammation in Sickle Cell Disease. *Int. J. Lab. Hematol.* **2018**, *40* (6), 704–709.

(170) Lamour, S. D.; Gomez-Romero, M.; Vorkas, P. A.; Alibu, V. P.; Saric, J.; Holmes, E.; Sternberg, J. M. Discovery of Infection Associated Metabolic Markers in Human African Trypanosomiasis. *PLoS Negl. Trop. Dis.* **2015**, *9* (10), e0004200.

(171) Kianoush, S.; Bittencourt, M. S.; Lotufo, P. A.; Bensenor, I. M.; Jones, S. R.; DeFilippis, A. P.; Toth, P. P.; Otvos, J. D.; Tibuakuu, M.; Hall, M. E.; et al. Association Between Smoking and Serum GlycA and High-Sensitivity C-Reactive Protein Levels: The Multi-Ethnic Study of Atherosclerosis (MESA) and Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *J. Am. Heart Assoc.* **2017**, *6* (8).

(172) Barber, J. L.; Kraus, W. E.; Church, T. S.; Hagberg, J. M.; Thompson, P. D.; Bartlett, D. B.; Beets, M. W.; Earnest, C. P.; Huffman, K. M.; Landers-Ramos, R. Q.; et al. Effects of Regular Endurance Exercise on GlycA: Combined Analysis of 14 Exercise Interventions. *Atherosclerosis* **2018**, *277*, 1–6.

(173) Vermeire, S.; Van Assche, G.; Rutgeerts, P. Laboratory Markers in IBD: Useful, Magic, or Unnecessary Toys? *Gut* **2006**, *55* (3), 426–431.

(174) Lloyd-Jones, D. M.; Liu, K.; Tian, L.; Greenland, P. Narrative Review: Assessment of C-Reactive Protein in Risk Prediction for Cardiovascular Disease. *Ann. Intern. Med.* **2006**, *145* (1), 35–42.

(175) Blaha, M. J.; Rivera, J. J.; Budoff, M. J.; Blankstein, R.; Agatston, A.; O'Leary, D. H.; Cushman, M.; Lakoski, S.; Criqui, M. H.; Szklo, M.; et al. Association Between Obesity, High-Sensitivity C-Reactive Protein ≥2 Mg/L, and Subclinical Atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **2011**, *31* (6), 1430–1438.

(176) Biou, D.; Konan, D.; Féger, J.; Agneray, J.; Leroy, Y.; Cardon, P.; Fournet, B.; Durand, G. Alterations in the Carbohydrate Moiety of Alpha-1-Acid Glycoprotein Purified from Human Cirrhotic Ascitic Fluid. *Biochim. Biophys. Acta* **1987**, *913* (3), 308–312.

(177) Van Dijk, W.; Koeleman, C.; Van het Hof, B.; Poland, D.;

Jakobs, C.; Jaeken, J. Increased Alpha3-Fucosylation of Alpha(1)-Acid Glycoprotein in Patients with Congenital Disorder of Glycosylation Type IA (CDG-Ia). *FEBS Lett.* **2001**, *494* (3), 232–235.

(178) Mackiewicz, A.; Mackiewicz, K. Glycoforms of Serum Alpha 1-Acid Glycoprotein as Markers of Inflammation and Cancer. *Glycoconj. J.* **1995**, *12* (3), 241–247.

(179) Serbource-Goguel, N.; Corbic, M.; Erlinger, S.; Durand, G.; Agneray, J.; Feger, J. Measurement of Serum Alpha 1-Acid Glycoprotein and Alpha 1-Antitrypsin Desialylation in Liver Disease. *Hepatology* **1982**, *3* (3), 356–359.

(180) Mondal, G.; Chatterjee, U.; Das, H. R.; Chatterjee, B. P. Enhanced Expression of A1-Acid Glycoprotein and Fucosylation in Hepatitis B Patients Provides an Insight into Pathogenesis. *Glycoconj. J.* **2009**, *26* (9), 1225–1234.

(181) Ishibashi, K.; Nishikawa, A.; Hayashi, N.; Kasahara, A.; Sato, N.; Fujii, S.; Kamada, T.; Taniguchi, N. N-Acetylglucosaminyltransferase III in Human Serum, and Liver and Hepatoma Tissues: Increased Activity in Liver Cirrhosis and Hepatoma Patients. *Clin. Chim. Acta.* **1989**, *185* (3), 325–332.

(182) Miyoshi, E.; Noda, K.; Yamaguchi, Y.; Inoue, S.; Ikeda, Y.; Wang, W.; Ko, J. H.; Uozumi, N.; Li, W.; Taniguchi, N. The Alpha1-6-Fucosyltransferase Gene and Its Biological Significance. *Biochim. Biophys. Acta* **1999**, *1473* (1), 9–20.

(183) Mondal, G.; Chatterjee, U.; Chawla, Y. K.; Chatterjee, B. P. Alterations of Glycan Branching and Differential Expression of Sialic Acid on Alpha Fetoprotein among Hepatitis Patients. *Glycoconj. J.* **2011**, *28* (1), 1–9.

(184) Comunale, M. A.; Rodemich-Betesh, L.; Hafner, J.; Wang, M.; Norton, P.; Di Bisceglie, A. M.; Block, T.; Mehta, A. Linkage Specific Fucosylation of Alpha-1-Antitrypsin in Liver Cirrhosis and Cancer Patients: Implications for a Biomarker of Hepatocellular Carcinoma. *PLoS One* **2010**, *5* (8), e12419.

(185) Lee, H. B.; Yoo, O. J.; Ham, J. S.; Lee, M. H. Serum Alpha 1-Antitrypsin in Patients with Hepatocellular Carcinoma. *Clin. Chim. Acta.* **1992**, *206* (3), 225–230.

(186) Goodarzi, M. T.; Turner, G. A. Decreased Branching, Increased Fucosylation and Changed Sialylation of Alpha-1-Proteinase Inhibitor in Breast and Ovarian Cancer. *Clin. Chim. Acta.* **1995**, *236* (2), 161–171.

(187) Vaughan, L.; Lorier, M. A.; Carrell, R. W. Alpha 1-Antitrypsin Microheterogeneity. Isolation and Physiological Significance of Isoforms. *Biochim. Biophys. Acta* **1982**, *701* (3), 339–345.

(188) Hrycaj, P.; Sobieska, M.; Mackiewicz, S.; Müller, W. Microheterogeneity of Alpha 1 Acid Glycoprotein in Rheumatoid Arthritis: Dependent on Disease Duration? *Ann. Rheum. Dis.* **1993**, *52* (2), 138–141.

(189) Yamashita, K.; Koide, N.; Endo, T.; Iwaki, Y.; Kobata, A. Altered Glycosylation of Serum Transferrin of Patients with Hepatocellular Carcinoma. *J. Biol. Chem.* **1989**, *264* (5), 2415–2423.

(190) Matsumoto, K.; Maeda, Y.; Kato, S.; Yuki, H. Alteration of Asparagine-Linked Glycosylation in Serum Transferrin of Patients with Hepatocellular Carcinoma. *Clin. Chim. Acta.* **1994**, *224* (1), 1–8.

(191) Vanarsa, K.; Ye, Y.; Han, J.; Xie, C.; Mohan, C.; Wu, T. Inflammation Associated Anemia and Ferritin as Disease Markers in Systemic Lupus Erythematosus. *Arthritis Res. Ther.* **2012**, *14* (4), R182.

(192) Mann, A. C.; Record, C. O.; Self, C. H.; Turner, G. A. Monosaccharide Composition of Haptoglobin in Liver Diseases and Alcohol Abuse: Large Changes in Glycosylation Associated with Alcoholic Liver Disease. *Clin. Chim. Acta.* **1994**, *227* (1–2), 69–78.

(193) Okuyama, N.; Ide, Y.; Nakano, M.; Nakagawa, T.; Yamanaka, K.; Moriwaki, K.; Murata, K.; Ohigashi, H.; Yokoyama, S.;

Eguchi, H.; et al. Fucosylated Haptoglobin Is a Novel Marker for Pancreatic Cancer: A Detailed Analysis of the Oligosaccharide Structure and a Possible Mechanism for Fucosylation. *Int. J. cancer* **2006**, *118* (11), 2803–2808.

(194) Park, S.-Y.; Yoon, S.-J.; Jeong, Y.-T.; Kim, J.-M.; Kim, J.-Y.; Bernert, B.; Ullman, T.; Itzkowitz, S. H.; Kim, J.-H.; Hakomori, S. N-Glycosylation Status of β-Haptoglobin in Sera of Patients with Colon Cancer, Chronic Inflammatory Diseases and Normal Subjects. *Int. J. Cancer* **2010**, *126* (1), 142–155.

(195) Yoon, S.-J.; Park, S.-Y.; Pang, P.-C.; Gallagher, J.; Gottesman, J. E.; Dell, A.; Kim, J.-H.; Hakomori, S.-I. N-Glycosylation Status of Beta-Haptoglobin in Sera of Patients with Prostate Cancer vs. Benign Prostate Diseases. *Int. J. Oncol.* **2010**, *36* (1), 193–203.

(196) Dargan, E.; Thompson, S.; Cantwell, B. M. J.; Wilson, R. G.; Turner, G. A. Changes in the Fucose Content of Haptoglobin in Breast and Ovarian Cancer: Association with Disease Progression. *Glycosylation Dis.* **1994**, *1* (1), 37–43.

(197) Nakano, M.; Nakagawa, T.; Ito, T.; Kitada, T.; Hijioka, T.; Kasahara, A.; Tajiri, M.; Wada, Y.; Taniguchi, N.; Miyoshi, E. Site-Specific Analysis of N-Glycans on Haptoglobin in Sera of Patients with Pancreatic Cancer: A Novel Approach for the Development of Tumor Markers. *Int. J. Cancer* **2008**, *122* (10), 2301–2309.

(198) Tomana, M.; Schrohenloher, R. E.; Koopman, W. J.; Alarcón, G. S.; Paul, W. A. Abnormal Glycosylation of Serum IgG from Patients with Chronic Inflammatory Diseases. *Arthritis Rheum.* **1988**, *31* (3), 333–338.

(199) Parekh, R. B.; Roitt, I. M.; Isenberg, D. A.; Dwek, R. A.; Ansell, B. M.; Rademacher, T. W. Galactosylation of IgG Associated Oligosaccharides: Reduction in Patients with Adult and Juvenile Onset Rheumatoid Arthritis and Relation to Disease Activity. *Lancet (London, England)* **1988**, *1* (8592), 966–969.

(200) Mehta, A. S.; Long, R. E.; Comunale, M. A.; Wang, M.;

Rodemich, L.; Krakover, J.; Philip, R.; Marrero, J. A.; Dwek, R. A.; Block, T. M. Increased Levels of Galactose-Deficient Anti-Gal Immunoglobulin G in the Sera of Hepatitis C Virus-Infected Individuals with Fibrosis and Cirrhosis. *J. Virol.* **2008**, *82* (3), 1259–1270.

(201) Kanoh, Y.; Mashiko, T.; Danbara, M.; Takayama, Y.; Ohtani, S.; Egawa, S.; Baba, S.; Akahoshi, T. Changes in Serum IgG Oligosaccharide Chains with Prostate Cancer Progression. *Anticancer Res.* **2004**, *24* (5B), 3135–3139.

(202) Ercan, A.; Cui, J.; Chatterton, D. E. W.; Deane, K. D.; Hazen, M. M.; Brintnell, W.; O'Donnell, C. I.; Derber, L. A.; Weinblatt, M. E.; Shadick, N. A.; et al. Aberrant IgG Galactosylation Precedes Disease Onset, Correlates with Disease Activity, and Is Prevalent in Autoantibodies in Rheumatoid Arthritis. *Arthritis Rheum.* **2010**, *62* (8), 2239–2248.

(203) Vučković, F.; Krištić, J.; Gudelj, I.; Teruel, M.; Keser, T.; Pezer, M.; Pučić-Baković, M.; Štambuk, J.; Trbojević-Akmačić, I.; Barrios, C.; et al. Association of Systemic Lupus Erythematosus With Decreased Immunosuppressive Potential of the IgG Glycome. *Arthritis Rheumatol.* **2015**, *67* (11), 2978–2989.

(204) Gudelj, I.; Lauc, G.; Pezer, M. Immunoglobulin G Glycosylation in Aging and Diseases. *Cell. Immunol.* **2018**, *333*, 65–79.

(205) Axford, J. S. Glycosylation and Rheumatic Disease. *Adv. Exp. Med. Biol.* **1998**, *435*, 163–173.

(206) Goodarzi, M. T.; Axford, J. S.; Varanasi, S. S.; Alavi, A.; Cunnane, G.; Fitzgerald, O.; Turner, G. A. Sialyl Lewis(x) Expression on IgG in Rheumatoid Arthritis and Other Arthritic Conditions: A Preliminary Study. *Glycoconj. J.* **1998**, *15* (12), 1149–1154.

(207) Barratt, J.; Smith, A. C.; Feehally, J. The Pathogenic Role of IgA1 O-Linked Glycosylation in the Pathogenesis of IgA Nephropathy (Review Article). *Nephrology* **2007**, *12* (3), 275–

284.

(208) Lai, K. N. Pathogenesis of IgA Nephropathy. *Nat. Rev. Nephrol.* **2012**, *8* (5), 275–283.

(209) Gruppen, E. G.; Connelly, M. A.; Vart, P.; Otvos, J. D.; Bakker, S. J.; Dullaart, R. P. GlycA, a Novel Proinflammatory Glycoprotein Biomarker, and High-Sensitivity C-Reactive Protein Are Inversely Associated with Sodium Intake after Controlling for Adiposity: The Prevention of Renal and Vascular End-Stage Disease Study. *Am. J. Clin. Nutr.* **2016**, *104* (2), 415–422.

CHAPTER 4

4. Characterization of ^{1}H-NMR Plasma Glycoproteins as a new strategy to identify inflammatory patterns in Rheumatoid Arthritis

4.1. Abstract

Rheumatoid Arthritis (RA) is a chronic autoimmune inflammatory disease associated with a high index of morbidity and mortality from cardiovascular diseases (CVDs). In this study we used ^{1}H-NMR to characterize the plasma glycoprotein and lipoprotein profiles of a cohort of patients with RA (n=210) versus healthy individuals (n=203) to associate them with the RA disease and its severity. Using ^{1}H-NMR, we developed a lineshape method to characterize the two peaks associated with glycoproteins (GlycA and GlycB) and its derived variables: areas of GlycB (Area GlycB) and GlycA (Area GlycA), shape factors of these two peaks (H/W=Height/Width) and the distance between them (Distance GlycB-GlycA). It was also used the advanced lipoprotein test Liposcale® (CE) to characterize the lipoprotein subclasses. The standard lipid panel and traditional inflammatory markers such as C-reactive protein (CRP), the erythrocyte sedimentation rate (ESR), fibrinogen, the rheumatoid factor (RF), anti-citrullinated peptide antibodies (ACPA) and the DAS28 index has been also determined. RA patients presented a significant 10.65 % increase in the GlycA associated area compared to the control group ($p=2.21x10^{-10}$). They also presented significantly higher H/W GlycA and GlycB ratios than the control population (H/W GlycB $p=7.88x10^{-08}$; H/W GlycA $p=5.61x10^{--08}$). The prediction model that uses the traditional inflammatory variables and the ^{1}H-NMR-derived parameters presented an AUC that was almost 10% higher than the model that only uses the traditional inflammatory variables (from 0.7 to 0.79 AUC). In this study, we have demonstrated that GlycA and GlycB variables derived from ^{1}H-NMR, along with

classic inflammatory parameters helps to improve the classification of individuals with high RA disease activity.

4.2. Introduction

Rheumatoid arthritis (RA) is a systemic disease characterized by chronic inflammatory involvement of the synovial membrane of the joints and autoantibody production (rheumatoid factor [RF] and anti–citrullinated protein antibody [ACPA]).[1,2] This autoimmune process destroys cartilage and joint bones as well as ligaments and tendons. The exact cause of RA is not known but it is considered a multifactorial disease, resulting from the interaction of risk factors such as genetic susceptibility, sex, age, smoking, and infectious, hormonal, dietary, socioeconomic and ethnic agents. Patients with RA may have any of the traditional risk factors for cardiovascular disease (elevated levels of low-density lipoprotein [LDL] cholesterol or high-density lipoprotein [HDL] dysregulations). However, evidence suggests that the prevalence of these risk factors is no higher than in the general population.[3] Studies on the incidence and prevalence of RA in recent decades estimate a prevalence between 0.5 and 1% depending on the population studied, with an annual incidence of 0.02-0.05%.[4] Given that the course of the disease is chronic and progressive in most RA patients, there is a high socioeconomic cost as well as a significant impact on the physical function, productivity, quality and life expectancy of the patients.

Comorbidities may precede or accompany RA and there is evidence to suggest that the systemic inflammation and immune dysfunction that characterize RA play a major role in its development and acceleration.[5] Of these comorbidities, CVD is the leading cause of death in patients with RA. It is well established in the literature that the rate of CVD is higher in RA patients than in the general population.[6,7] The chronic inflammatory state of RA contributes to the onset and development of accelerated atherosclerosis, since the inflammatory process in synovial and atherosclerotic plaques is very similar.[8] In addition, there is an increase in the synthesis of proinflammatory cytokines such as TNF-alpha and IL-6, which facilitates endothelial dysfunction as an initial step in atherosclerosis, leading to the formation and rupture of atherosclerotic plaques.[9]

The classical way of measuring the severity of inflammation through acute-phase reactants or serum or plasma proteins, mainly glycoproteins, was first described many years ago.[10] Serial measurements of the erythrocyte sedimentation rate (ESR), which is largely a measure of fibrinogen, have been used to monitor the progress of inflammatory disorders. Moreover, C reactive protein (CRP) reflects the clinical course of the disease over periods up to three years and is objective, simple and easy to record. However, the acute-phase protein response is not specific to rheumatoid arthritis. To measure the disease activity of patients with RA the DAS28 index is used. It is a composite score of four measures: the number of tender joints (out of 28), the number of swollen joints (out of 28), ESR, and the general health of the patients.[11]

CHAPTER 4

Several studies have shown that glycoproteins play a key role in inflammatory and pathological processes.[12–15] Protein glycosylation is a post-translational process responsible for the attachment of glycan chains to the nitrogen of an asparagine residue (N-linkage) or to the oxygen of a serine or threonine residue (O-linkage) by a covalent bond.[16,17] During inflammation, there are changes in the number of antennary branches, increased sialylation and fucosylation, and decreased galactosylation.[18] The branches are rich in N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), N-acetylneuraminic acid (or sialic acid) and fucose residues.[16,18–20] The plasma levels of circulating glycoproteins rise (positive acute phase proteins) or fall (negative acute phase proteins) during the acute phase response to the presence of some environmental stimuli in the organism.[21] The role of glycosylation in the structural changes of proteins means that they can be used as potential early biomarkers of disease.

Currently, the serum concentrations of glycoproteins such as antibodies, antiproteases, and binding or transport proteins are determined using a variety of methods: for example, enzyme-linked immunosorbent assays (ELISAs), electrochemiluminescence immunoassay (ECLIA), luminex-based assays, radioimmunoassays (RIA) and nephelometric assays. These methods quantify the amount of protein in biological samples.[22] Despite this, to date there is no high-performance, fast, or sufficiently sensitive technology that can measure the glycan portion of inflammatory proteins and quantify a measure of general glycosylation that can be useful for diagnosis, and the assessment of disease severity and treatment efficacy.

Given the imperfections of conventional biomarkers for the diagnosis, prognosis, risk prediction and disease prevention at the individual patient level, there is an ongoing effort using novel high-precision laboratory techniques to discover new biomarkers that could increase the sensitivity and specificity of the current clinical results. The increasing role of proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy in recent years has enabled compounds such as metabolites, lipoproteins and glycoproteins to be detected.[23] As many of the ^{1}H-NMR peaks are overlapped, the total-line-shape (CTLS) fitting strategy has been used to quantify the individual signals by modelling theoretical functions.[24]

It has been reported elsewhere that ^{1}H-NMR spectroscopy can quickly and accurately detect circulating levels of glycoproteins.[21,25] Thanks to the initial work by Bell et al.,[26] the quantification of glycosylated proteins is now technologically viable using the signals associated with the side-chain protons of the N-acetyl-carbohydrate groups, which are covalently bound to plasma glycoproteins. [25]

Recently, a pro-inflammatory glycoprotein biomarker termed *GlycA* has been reported in the literature as a marker of inflammation that can be measured by proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy and is associated with cardiometabolic disease and mortality and other inflammatory diseases.[21,25,27–29] It can also be used to quantify plasmatic glycosylated proteins .[30,31] In particular, this signal is associated with the concentration of specific residues of N-acetylglucosamine and N-acetylgalactosamine in the branched side chains of glycosylated plasma proteins (mainly α1-antitrypsin, α1-

antichymotrypsin, haptoglobin and transferrin). It has been shown that the quantification of these signals is a good indicator of the detection, prognosis and therapeutic monitoring of tissue injury marked by systemic inflammatory processes, as well as cardiovascular risk and type 2 diabetes.[25,32] Additionally, it is demonstrated that higher serum levels N-acetyl signal of glycoproteins has been revealed also as a biomarker of metastatic colorectal cancer.[33] Some associations with classical inflammation markers have been described previously as is the relationship of GlycA and CRP.[25]

Considering that there is a need for more reliable inflammatory risk assessment tools, in this study we aim first to use ^{1}H-NMR to characterize the plasma glycoprotein profile of a cohort of patients with RA versus healthy individuals. Our second aim is to model the activity of RA to identify patterns indicating the severity of the disease.

4.3. Experimental section

4.3.1. Patients

The study population included 210 individuals (n=76 men and n=134 women) with a mean age of 58.00 (±12) and a body mass index of 27.79 (±6) who had been diagnosed with RA and were attending the rheumatology unit of the HUSJ (Hospital Universitari Sant Joan de Reus, Spain). We excluded patients older than 80 and younger than 20, and patients with acute intercurrent diseases. Another population of 203 apparently healthy individuals (CT) without AR and CVDs matched by sex, age, and body mass index were used as the control

group. The study was approved by the ethical clinical research committee of the HUSJ and informed consent was obtained from each patient. Patients were visited between September 2011 and November 2014 and on the same day of the medical visit, blood was collected.

4.3.2 Plasma sample handling and analytical methods

Blood samples were withdrawn from the antecubital vein of each participant at the time of recruitment after a 12-hour overnight fast. EDTA plasma was prepared from venous blood collected into sterile, evacuated tubes (BD, Vacutainer). Plasma was immediately separated by low-speed centrifugation at 4 °C and frozen at -80 °C until biochemical and NMR analysis

4.3.2. Biochemical analysis

For the RA patients, we used traditional biochemical methods to determine the classic inflammatory markers: CRP, ESR, fibrinogen, rheumatoid factor (RF) and anti-citrullinated peptide antibodies (ACPA). The index of clinical activity was assessed by the DAS28 index. Apolipoprotein A1 and B:100 concentrations and the standard lipid profile including total cholesterol, HDL cholesterol and triglycerides were also determined. LDL cholesterol was calculated using the Fridewald formula.[34]

4.3.3. Glycoprotein profiling

Plasma samples (250 µL) were previously diluted with 31 µl deuterated water and 244 µl of 50 mM phosphate buffer solution (PBS)

at pH 7.4 consisting of 30.70 Na_2HPO_4 mM and 19.30 NaH_2PO_4 mM before NMR analysis.

^{1}H-NMR spectra were recorded at 310 K on a Bruker Avance III 600 spectrometer operating at a proton frequency of 600.20 MHz (14.1 T). One-dimensional ^{1}H-NMR pulse experiments were carried out using the nuclear Overhauser effect spectroscopy (NOESY)-presaturation sequence (RD-90° -s_1-90° -s_m-90° ACQ) to suppress the residual water peak. The time s_1 was set to 4 µs, and the time s_m (mixing time) was 100 ms. The 90° pulse length was calibrated for each sample[35] and varied from 12.3 to 20.2 µs. The spectral width was 30 ppm (18,000 Hz) and a total of 64,000 data points and four scans (NS) and a receiver gain value of 14.2 were acquired for each sample in NOESY pulse. The gradient pulse strength was 95% of the maximum strength of 53 G cm^{-1} (0.535 T m^{-1}). The glycoprotein profiling method was performed in the region between 2.15 and 1.90 ppm of chemical shift.[36]

4.3.4. Lipoprotein characterization

The diluted plasma samples used for the glycoprotein profiling were analyzed by using the NMR based Liposcale® test (CE) previously reported.[37] This protocol evaluates lipid concentrations (i.e., triglycerides and cholesterol), size and particle number of three different classes of lipoproteins (VLDL, LDL and HDL), as well as the particle number of nine subclasses (large, medium and small VLDL, LDL and HDL). Briefly, particle concentration and diffusion coefficients were obtained from the measured amplitudes and the attenuation of their spectroscopically distinct lipid methyl group NMR

signals. To determine the lipoprotein size, the methyl signal was surface fitted using a previously reported procedure [37] with the numbers of functions so the nine lipoprotein subclasses could be determined. The NMR functions were associated with a given lipoprotein subclass (large, medium and small VLDL, LDL or HDL) according to their associated NMR size. The mean particle diameter for the subclasses (and the estimated ranges) were as follows: large VLDL particles (VLDL-P), 75.2 nm (>60 nm); medium VLDL-P, 52.1 nm (45-60 nm); small VLDL-P, 37.1 nm (35-45 nm); large LDL particles (LDL-P), 22.8 nm (22.5-27 nm); medium LDL-P, 20.5 nm (20-22.5 nm); small LDL-P 18.9 nm (18-20 nm); large HDL particles (HDL-P), 10.1 nm (9-13 nm); medium HDL-P, 8.7 nm (8.2-9 nm); small HDL-P, 7.8 nm (<8.2 nm) in agreement with previous literature.[38–40]

The particle numbers of each lipoprotein subclass were calculated by dividing the lipid volume by the particle volume of a given class. The lipid volumes were determined using common conversion factors to convert the concentration units of the esterified cholesterol (CE) and triglycerides (TG) contained in the lipoprotein core into volume units by using the averaged molecular volumes 1.058 g/ml and 1.021 g/ml for CE and TG and molar mass respectively.[41] Finally, weighted average VLDL, LDL and HDL particle sizes (in nm diameter units) can be calculated from the various subclass concentrations by summing the known diameter of each subclass multiplied by its relative percentage of subclass particle number as estimated from the intensity of its methyl NMR signal. [42]

4.3.5. Statistical analysis

All statistical analyses were computed in MATLAB, Ver. 7.10.0 using PLS-Toolbox, Ver. 5.2.2 (Eigenvector Research). Firstly, univariate statistical analysis of the glycoprotein variables was conducted to identify differences between the RA patients and the healthy individuals. The Lilliefors test was used to study the normality of each variable. Not all the variables followed a normal distribution, therefore in this study the non-parametric Wilcoxon-Mann-Whitney test was applied. The false discovery rate (FDR) adjustment was applied in all the tests of the present study.[43] A Partial Least Squares-Discriminant Analysis (PLS-DA) with the glycoprotein variables measured by ^{1}H-NMR was conducted to identify the characteristic glycoprotein profiles associated with the AR disease and the healthy group. Additionally, to study the relationship between glycoprotein variables and lipids in the RA population we focused on the associations between the glycoprotein and lipoprotein variables determined by biochemical methods and the Liposcale® test. The associations between the glycoprotein variables and the inflammatory markers CRP, ESR, RF, ACPA and DAS28 were also studied. Spearman correlation coefficients were calculated for variable distributions.

Finally, various auto-scaled and cross-validated PLS-DA models by Venetian blinds cross validation were evaluated to build the best predictor model and identify individuals in the highest 20th percentile of disease activity according to the DAS28 index.[44] The input variables used were the traditional inflammatory markers and ^{1}H-

NMR parameters (glycoprotein and lipoprotein). Because of the high number of variables, the variable selection method Genetic Algorithms (GA)[45] was used to construct more reliable models. Receiver Operating Characteristic (ROC) curves were evaluated to assess the diagnostic ability of these models. Permutation test was applied on all PLS-DA models.

4.4. Results

The RA group consisted of 210 individuals (n=76 men and n=134 women) with a mean age of 58 (±12.2) and a body mass index of 27.79 (±6) distributed in 4 groups according to their DAS28 index. The demographic, clinical and biochemical characteristics of patients are shown in **table C4.1**.

We deconvoluted mathematically the glycoprotein region as follows: the raw spectra was fitted with five analytical Voigt functions corresponding to the five signals: low molecular weight (LMW), GlycB, GlycA, G-lipid, and baseline based on their chemical shift. For each of these functions the total area (proportional to the concentration), the height, the position (characteristic of the magnetic environment) and the width (related to the flexibility and the aggregation state of the molecules generating the signal) were determined. The baseline function was used to fit the background within the region 1.80 – 1.92 ppm. The function LMW was used to fit the low molecular weight metabolites that resonate in the region 2.12 – 2.14 ppm. The functions GlycB and GlycA were used to fit the signal

produced by the -$COCH_3$ acetyl groups of N-acetylglucosamine and N-acetylgalactosamine (GlycA) and N-acetylneuraminic acid (GlycB), which resonates in the region 2.02 – 2.07 [36]; and the G-lipid function was used to fit the signal produced by the CH_2=C protons of lipids in the sample, which resonates in the region 1.95 – 2.01 ppm. **Figure C4.1** focuses on the glycoprotein profiling region of the spectrum.

Table C4.1. Demographic, clinical and biochemical characteristics of patients and controls.

		RA group				Control group			
		Number	**%**	**Mean**	**SD**	**Number**	**%**	**Mean**	**SD**
Gender	Women	134	63.8			131	64.5		
	Men	76	36.2			72	35.5		
Age (years)				58	12			45	16
BMI (Kg/m2)				27.79	5.89			25.30	2.99
Waist circumference (cm)				92	15				
Systolic blood pressure				137	21				
Diastolic blood pressure				81	12				
HDL cholesterol (mg/dL)				65	19				
LDL cholesterol (mg/dL)				118	31				
Triglycerides (mg/dL)				105	56				
Glucose (mg/dL)				95	23				
Current smoker (n, %)	No	154	73.3						
	Yes	56	26.7						
Diabetes mellitus (n, %)	No	185	88.1						
	Yes	25	11.9						
Hypertension (n, %)	No	85	40.5						
	Yes	125	59.5						
Dyslipidemia (n, %)	No	124	59.0						
	Yes	86	41.0						
Disease duration (years)				9.3	9.2				
DAS28				3.45	1.27				

	Disease remission (<2.6)	57	27.1		
	Low disease activity (2.6-3.2)	39	18.6		
	Moderate disease activity (3.2-5.6)	95	45.2		
	High disease activity (>5.1)	19	9.0		
HAQ (0-2.5)				0.44	0.52
Rheumatoid factor + (n, %)	Negative	58	27.6		
	Positive	152	72.4		
ACPA+	Negative	40	19.0		
	Positive	170	81.0		
ESR (mm/h)				37	26
CPR (mg/dL)				0.72	0.84
Fibrinogen (mg/dL)				442	98
DMARs	No	51	24.3		
	Yes	159	75.7		
Biological agent	No	169	80.5		
	Yes	41	19.5		
Corticoids	No	105	50.0		
	Yes	105	50.0		
NSAIDs	No	89	42.4		
	yes	121	57.6		

HAQ=health assessment questionnaire index, ACPA=citrullinated anti-cyclic peptide antibodies, ESR=erythrocyte sedimentation rate, CRP=C-reactive protein, DAS28=disease activity score, DMARDs=disease-modifying antirheumatic drugs, NSAIDs=non-steroidal antiinflammatory drugs.

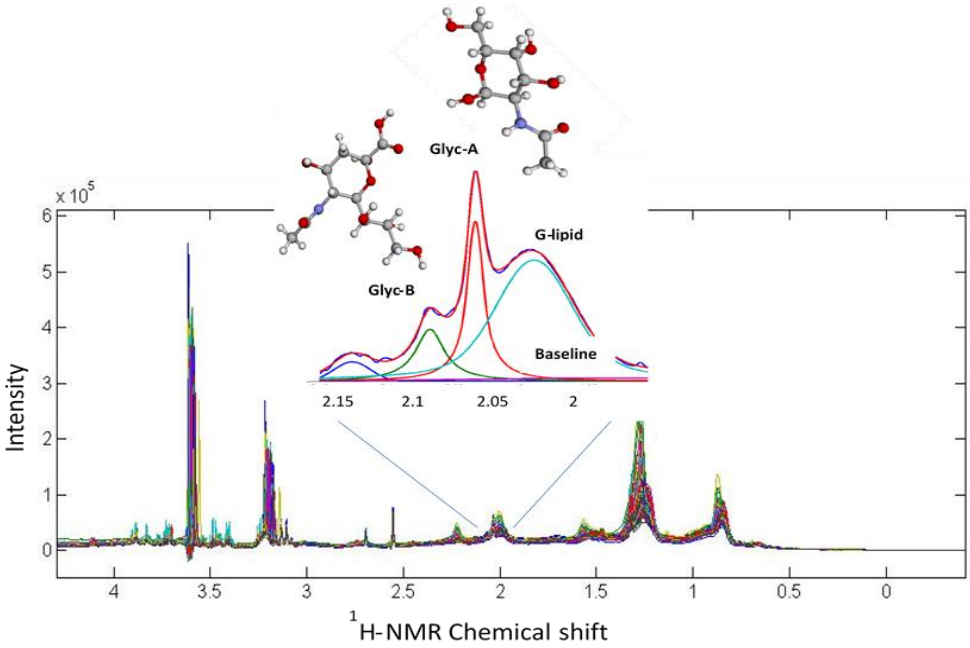


Figure C4.1. Mathematically treated proton nuclear magnetic resonance spectrum (phase corrected, referenced and scaled) in which the region where the signal produced by glycosylated proteins is indicated. The raw spectra were fitted with five analytic Voigt functions corresponding to five signals LMW (blue), GlycB (green), GlycA (red), G-lipid (cyan), and baseline (purple).

Additionally, for GlycB and GlycA functions we calculated the derived parameters H/W=Height/Width to capture the shape of the peaks and the parameter distance GlycB-GlycA to measure the distance between these two signals.

In this study, due to the high degree of correlation between the variables obtained by the peak deconvolution, we shall only consider the following five variables: areas of GlycB (Area GlycB) and GlycA

(Area GlycA), shape factors of these two peaks (H/W=Height/Width) and the distance between them (Distance GlycB-GlycA). The area of these functions gives us information about the concentration of acetyl groups of N-acetylglucosamine and N-acetylgalactosamine (GlycA) and N-acetylneuraminic acid (GlycB) in plasma. The shape factor ratio H/W provides information on what the function is like in each case depending on its height, which is related to the concentration, and its width, which is related to the flexibility and the aggregation of the molecules generating the signal. The variable Distance GlycB-GlycA indicates the relative distance between the chemical shift of both peaks.

The parameters mentioned above were extracted by a lineshape fitting process from the NOESY spectra using the normalized root mean squared error (NRMSE) in the interval 1.90 – 2.15 ppm. The error in all samples was smaller than 2.93 %.

The results of the univariate analysis showed that RA patients presented a significant 10.65 % increase in the GlycA associated area compared to the CT group ($p=2.21x10^{-10}$). They also presented significantly increase in the H/W GlycA ratio (12.70%) and GlycB ratio (16.76%) than the control population (H/W GlycB $p=7.88x10^{-08}$; H/W GlycA $p=5.61x10^{-08}$). The distance between the two functions in the RA group was significantly greater than in the control group (see Table 2). Area GlycB, Area GlycA and the distance parameters did not follow a normal distribution and, as has been explained in statistical analysis section, the non-parametric Wilcoxon-Mann-Whitney test was used to calculate the significance for all parameters.

Table C4.2. Results of univariate analysis between RA and healthy individuals (CT)

	AR		CT		
Variable	**Median**	**Iqr**	**Median**	**Iqr**	***p*-value**
Area GlycB (a.u.)	$2.14x10^{-01}$	$8.56x10^{-02}$	$2.22x1^{0-01}$	$7.39x10^{-02}$	$1.89x10^{-01}$
Area GlycA (a.u.)	$5.26x10^{-01}$	$1.12x10^{-01}$	$4.70x10^{-01}$	$1.09x10^{-01}$	$2.21x10^{-10}$*
H/W GlycB	$3.94x10^{-03}$	$1.81x10^{-03}$	$3.27x10^{-03}$	$1.03x10^{-03}$	$7.88x10^{-08}$*
H/W GlycA	$1.83x10^{-02}$	$4.27x10^{-03}$	$1.65x10^{-02}$	$5.46x10^{-03}$	$5.61x10^{-08}$*
Distance GlycB-GlycA (ppm)	$3.30x10^{-02}$	$9.16x10^{-04}$	$3.21x10^{-02}$	$9.17x10^{-04}$	$3.68x10^{-22}$*

Significant values ($p<0.05$) are marked (*).Median and interquartile range (Iqr) are reported. The Wilcoxon-Mann-Whitney test has been used to calculate significance. *P*-values adjusted by FDR.

Figure C4.2a shows the score plot of a PLS-DA model ($p=5.18x10^{-15}$, Q2=0.128) of three latent variables (LV) built to discriminate RA patients from CT individuals by using the 5 ^{1}H-NMR-derived glycoprotein parameters as the input dataset. Figure 2b shows the loading plot of the same PLS-DA model to illustrate the contribution of each variable to each LV.

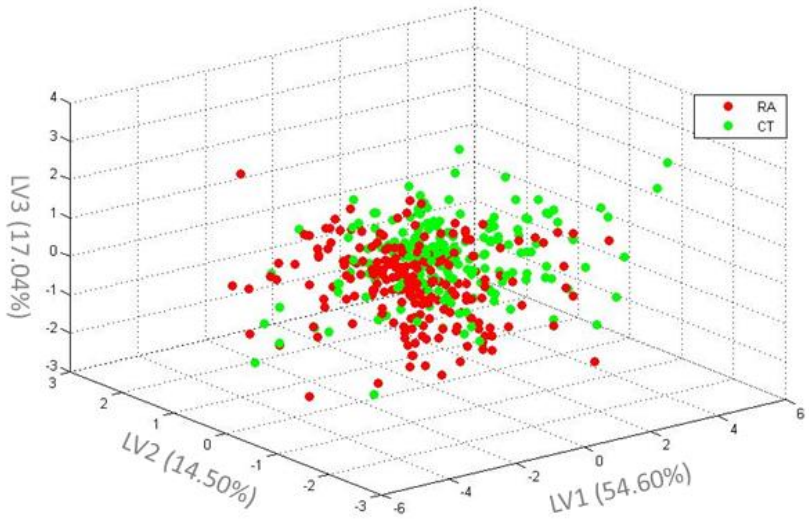


Figure C4.2a. 3D Score plot of the PLS-DA model (3 latent variables) for CT individuals and RA patients. RA patients are shown in red and control patients in green.

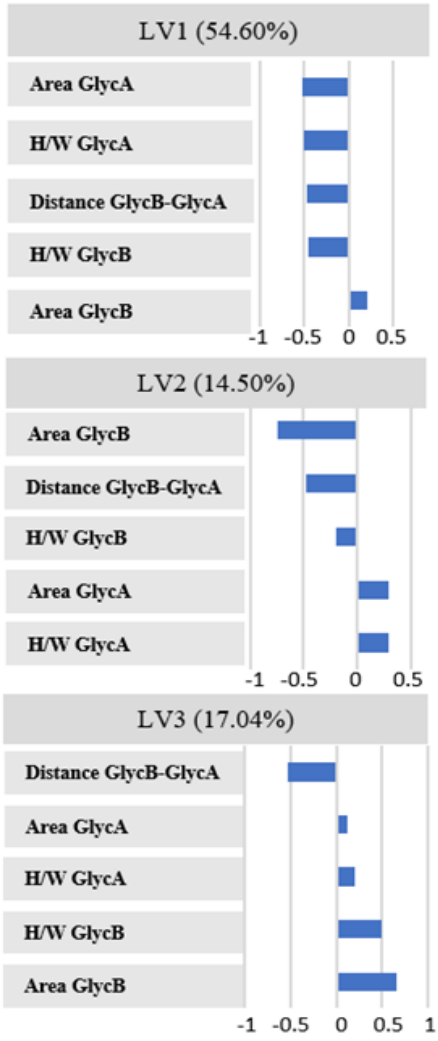


Figure C4.2b. PLS-DA loading plot showing the NMR glycoprotein-derived parameters contributing to the separation of the PLS-DA scores in RA patients and CT individuals.

The relationship between glycoprotein variables and clinical inflammatory markers was also studied. **Table C4.3** shows the Spearman correlation coefficients for Fibrinogen, DAS28, ESR, CRP, RF, ACPA and Albumin. There was a significant, positive but modest

correlation between the five glycoprotein variables and Fibrinogen, ESR, CRP and DAS28. There was also a significant and positive correlation between H/W GlycB and RF and ACPA. Correlations with Albumin are negative but significant. **Figure C4.2S** of the Supporting Information shows some dispersion plots of these results.

Table C4.3. Associations between ^{1}H-NMR glycoproteins variables and clinical inflammatory markers.

		Area GlycB	Area GlycA	H/W GlycB	H/W GlycA	Distance GlycB-GlycA
Fibrinogen (mg/dl)	r	0.247*	0.452*	0.477*	0.414*	0.397*
	p	1.08×10^{-03}	1.69×10^{-10}	2.14×10^{-11}	4.65×10^{-09}	2.29×10^{-08}
ESR (mm/h)	r	0.341*	0.424*	0.436*	0.381*	0.302*
	p	2.49×10^{-06}	1.89×10^{-09}	5.61×10^{-10}	9.07×10^{-08}	4.33×10^{-05}
CRP (mg/dl)	r	0.278*	0.456*	0.468*	0.457*	0.356*
	p	1.87×10^{-04}	1.34×10^{-10}	4.49×10^{-11}	1.34×10^{-10}	7.70×10^{-07}
RF	r	0.127	0.109	0.196*	0.068	0.047
	p	1.03×10^{-01}	1.65×10^{-01}	1.19×10^{-02}	4.02×10^{-01}	5.74×10^{-01}
ACPA	r	-0.019	0.090	0.143	0.042	0.019
	p	8.20×10^{-01}	2.56×10^{-01}	6.84×10^{-02}	6.09×10^{-01}	8.20×10^{-01}
DAS28	r	0.214*	0.285*	0.259*	0.229*	0.127
	p	5.35×10^{-03}	1.25×10^{-04}	5.59×10^{-04}	2.83×10^{-03}	1.01×10^{-01}
Albumin (g/dl)	r	-0.144	-0.168*	-0.132	-0.183*	-0.257*
	p	6.79×10^{-02}	3.18×10^{-02}	8.87×10^{-02}	1.86×10^{-02}	6.33×10^{-04}

Spearman correlation coefficients (r) and p-value (*p*) for each glycoprotein variable and the inflammatory markers obtained by traditional biochemistry. *P*-values adjusted by FDR. Significant values ($p<0.05$) are marked (*). ESR: Erythrocyte sedimentation rate; CRP: C-reactive protein; RF: Rheumatoid Factor; ACPA: anti-citrullinated peptide antibodies; DAS28: Disease Activity Score of 28 joints.

We also studied the relationship between the glycoprotein variables obtained by ^{1}H-NMR and the lipids obtained by traditional biochemistry. Table S-1 of the Supporting Information shows the Spearman correlation coefficients and *p*-values for total cholesterol, low-density lipoprotein cholesterol (LDLc), high-density lipoprotein cholesterol (HDLc), very low-density lipoprotein cholesterol (VLDLc) and total triglycerides (TG). There was a significant but modest correlation between almost all glycoprotein variables and VLDLc, HDLc and TG. In addition, the glycoprotein variables were also compared with the variables of the Liposcale® test (see Table S-2). VLDLc was still the variable most positively and significantly associated with Area GlycA (0.36) and negatively associated with area GlycB (-0.43). **Figure C4.1S** of the Supporting Information shows some scatter plots of the mentioned correlations.

To evaluate the ability of the models based on ^{1}H-NMR glycoprotein and lipoprotein and traditional inflammatory variables to predict the severity of the disease, two PLS-DA models were constructed. The objective of the two models was to evaluate whether the variables of the two models were able to separate individuals in the AR population with a high disease activity based on the DAS28 index. The two models had high RA activity patients (DAS28>80th percentile =4.53) as the y-block input binary variable (prediction block). The first PLS-DA model (Model A) had values of Fibrinogen, CRP, RF and ACPA as the x-block input variables. The second PLS-DA model (Model B) used a GA variable selection method to select the most important of the 5 ^{1}H-NMR-derived glycoprotein variables, 23 variables resulting from the ^{1}H-NMR lipoprotein profile by the

Liposcale® test and the 4 variables in Model A as the x-block input variables. The variables selected by the GA were: Area GlycB, CRP, Fibrinogen, HDL-TG, VLDL-P, small VLDL-P, small HDL-P and LDL diameter (LDL-Z). Both models were auto-scaled and cross-validated using Venetian Blinds. The significance level of Model A was p=0.55x10$^{-0.2}$ (Q2=0.089), while model B was p= 1.1369 x 10^{-08} (Q2= 0.189). The discriminant ability of the models was evaluated using the Area Under the Curve (AUC) of the Receiving Operating Characteristic (ROC) analysis which considered high RA activity patients (above 4.53 on the DAS28 index) as positive cases, and low-moderate RA activity patients (below 4.53) as negative cases. Figure S-3 in the Supporting Information shows the population distribution of DAS28 and a brief description of the population for the two AR groups. The AUC of Model B, which included the traditional variables (Model A) and ^{1}H-NMR-derived parameters, was almost 10% higher than that of Model A (from 0.7 to 0.79 AUC) indicating that Model B was much better at predicting the severity of the disease (see **Figure C4.3**).

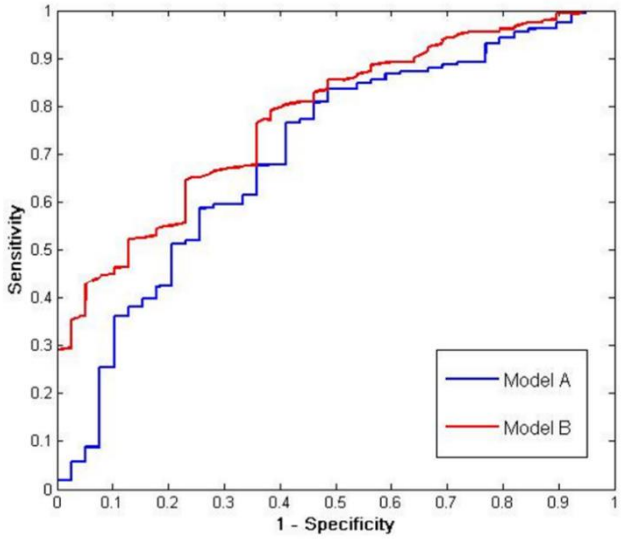


Figure C4.3. ROC curve of cross-validated Model A (blue) and Model B (red) for high/low AR activity. AUC Model A-value= 0.7017; AUC Model B-value= 0.7939.

4.5. Discussion

High-throughput technologies such as ^{1}H-NMR enable circulating plasma or serum glycoproteins to be measured using the specific residues of N-acetylglucosamine, N-acetylgalactosamine and N-acetylneuraminic acid in the branched side chains of glycosylated plasma proteins. In this study we used ^{1}H-NMR to characterize the plasma glycoprotein profile in RA. The results of the univariate study between RA and a control population confirmed expectations that RA patients have a significantly higher values of ^{1}H-NMR glycoprotein than the control group. These results are in line with the studies by J. Ormseth et al. and B. Barlett et al. in which GlycA concentrations were higher in patients with RA than in controls.[27,46] This may be because

RA patients have a permanent inflammatory condition that increases the levels of glycoproteins. In this study, we found that the area of GlycA was significantly higher in RA patients than in controls, which suggests that there was a higher concentration of glycosylated proteins presenting a higher signal associated with acetyl N-acetylgalactosamine and N-acetylglucosamine methyl groups. In addition, the shape factor of GlycA and GlycB indicated that the ratio between the height and width (H/W) of the deconvoluted peaks was significantly higher in patients with RA, which could be explained by the lower glycoprotein aggregation in RA patients' plasma.

In this study we used ^{1}H-NMR to characterize the glycoprotein profile as well as the advanced lipoprotein profile (Liposcale®). This has shown whether the information obtained by ^{1}H-NMR can help from a clinical point of view to evaluate patients with inflammatory processes. Comparable findings were described in the metabolic phenotyping study of Lauridsen et al. [47] where a cohort of patients with RA where measured by ^{1}H-NMR. They report that ^{1}H-NMR plasma metabolic profiles of RA patients (including cholesterol, lactate, acetylated glycoprotein, and lipid signatures) were significantly different from healthy subjects, indicating that the state of inflammation in RA patients is reflected in the ^{1}H-NMR spectra.

Table C4.3 shows that there were significant positive associations between Fibrinogen, ESR and CRP and the 5 ^{1}H-NMR glycoprotein variables. This is in line with other studies carried out in this field. In 2015, Otvos et al. showed that the concentration of GlycA correlated with high-sensitivity CRP and fibrinogen.[25,30] In this

context, the question arises as to whether these glycoproteins are found in high concentrations when there is an acute inflammatory process, such as rheumatoid arthritis, just as CRP, ESR and Fibrinogen are.[48,49] It should be noted that the H/W GlycA and H/W GlycB ratios were undoubtedly key variables throughout this study. They were a solid indicator of the shape of the glycoprotein peak in inflamed patients, capable of distinguishing RA from healthy individuals and also significantly correlated with lipid and inflammatory variables. Correlations were also found with the DAS28, which were very similar to the correlations with the inflammatory variables mentioned above since the DAS28 includes ESR measurements in its formula. Additionally, the glycoprotein variables are significantly associated in a negative way with albumin. This may be explained by the fact that hypoalbuminemia is characteristic of inflammatory states.[15] Furthermore, hypoalbuminemia in also related with juxta-articular erosions or with the incidence of peptic ulcer in RA.[50]

The representation of the ROC curve comparing the two PLS-DA models (Model A and Model B) indicated that by adding the glycoprotein and lipoprotein ^{1}H-NMR variables to classical inflammation parameters such as Fibrinogen, CRP, RF and ACPA we were able to make a better classification model (Model B) of the severity of the disease according to the 80th percentile of the DAS28 from the RA population.

The results of the correlations between glycoproteins and lipid variables in **Table C4.1S** of the Supporting Information show that GlycA and the H/W ratio of GlycA and GlycB parameters correlated

negatively with HDL-cholesterol, which makes sense since HDL-C is related to a cardio-protective function during inflammatory processes that occurs in cardiovascular events.[51–53] Furthermore, with the exception of Area GlycB, all the ^{1}H-NMR glycoprotein variables were significantly and positively correlated with VLDL-cholesterol, which, along with LDL-c, is associated with cardiovascular events. [54] These significant correlations were also noted in table S-2 when we studied the correlations with the variables of the Liposcale® test. In addition, some studies highlight the importance of considering strategies to reduce VLDL-cholesterol as a therapeutic intervention to reduce the residual risk of atherosclerotic cardiovascular disease,[55,56] beyond the well stablished strategy of reducing LDL-c.

In recent years, the advances in knowledge at genetic level (through the study of epigenetic modifications) and at proteomic level (at the level of the study of post-translational modifications) are contributing to a better characterization of the functional diversity of genes and proteins in the pathological states.[57] As we mentioned earlier in this study, glycosylation represents one of the most important and complex of these post-translational modifications. Studying the glycosylation of proteins could allow a more precise knowledge of the mechanisms of the disease and how is the implication of the proteins in the pathogenesis. The study of glycans has been called *glycomics* and it has received increasing interest as a novel tool for identifying markers and potential mediators of disease pathogenesis.[31] As we mentioned earlier, several authors have demonstrated that a serum glycan signal is associated with type 2 diabetes,[58] atherosclerotic CVD,[25,27,28,59] as well as longitudinal CVD and cancer mortality

risk[33,60] highlighting the importance of inflammation as a shared risk pathway. It should be noted that some authors have focused their study in a specific glycoprotein; for example, Menni et al.[61] studied the glycomic profiling of IgG to identify CVD risk; and Väänänen et al.[62] studied YKL-40, a chitinase-like glycoprotein associated with inflammation and tissue remodeling as a potential biomarker of disease activity in patients with early RA. The cited study by Lauridsen et al.[47] highlights that the metabolic profile may provide additional information to the clinical evaluation. In the present study we have evaluated the information provided by glycoprotein variables analyzed by ^{1}H-NMR as non-specific markers of inflammation as part of the current trend to seek new inflammatory markers that improve the limitations of classical inflammation biomarkers.[63–65]

This study confirms that the composite ^{1}H-NMR signal is associated with some clinical parameters of inflammation and the results reported are in line with the studies mentioned above in which the GlycA NMR-base assay has been used. Nevertheless, various points need to be borne in mind about this technique. On the one hand, ^{1}H-NMR is not a selective technique and does not enable individual proteins to be identified and quantified. The areas of GlycB and GlycA that we have obtained correspond to the overall concentration of glycosylated proteins in the plasma. On the other hand, one of its important advantages is that the ^{1}H-NMR spectrum enables a lipoprotein profile (Liposcale®) and a glycoprotein profile to be identified in a single experiment. This is a great advantage because the information is obtained with minimal manipulation of the samples and involves a considerable saving in experimental time. Finally, although

AR patients are being treated, this does not prevent the glycoprotein signals detected by ^{1}H-NMR being sensitive to the level of inflammation, as we have seen in the results. Bearing this in mind, in the future this first exploratory study should carefully study if medication can modify this ^{1}H-NMR spectrum region.

4.6. Conclusions

In this study, we have demonstrated that new markers of inflammation can be characterized using the ^{1}H-NMR signal of glycosylated proteins. In addition, the glycoprotein and lipoprotein ^{1}H-NMR variables, along with classic inflammatory parameters, provide information about the high activity of the disease that is more accurate than if we only consider the classic inflammatory parameters, as is conventional in clinical practice.

4.7. Supporting information

The content of this section is divided into four main parts. First, a table and scatter plots where correlations for each glycoprotein variable detected by ^{1}H-NMR and the lipids obtained by traditional biochemistry (total cholesterol, high density lipoprotein cholesterol, low density lipoprotein cholesterol, very low -density lipoprotein and total triglycerides) are shown. Second, a table listing correlation coefficients for each glycoprotein variable detected by ^{1}H-NMR and the variables resulting from the Liposcale® test. Third, scatter plots of

^{1}H-NMR glycoproteins variables and clinical inflammatory markers. Finally, we included the population distribution of DAS28 and a brief description of the population for the AR groups.

Table C4.1S. Spearman correlation coefficients (r) and p-values (p) for each glycoprotein variable and the lipids obtained by traditional biochemistry.

		Area GlycB	Area GlycA	H/W GlycB	H/W GlycA	Distance GlycB-GlycA
Cho(mg/dl)	r	-0.090	0.001	0.090	0.034	-0.125
	p	$2.56x10^{-01}$	$9.93x10^{-01}$	$2.56x10^{-01}$	$6.78x10^{-01}$	$1.06x10^{-01}$
HDLc(mg/dl)	r	0.165*	-0.224*	-0.031	-0.184*	-0.219*
	p	$3.50x10^{-02}$	$3.55x10^{-03}$	$7.02x10^{-01}$	$1.85x10^{-02}$	$4.29x10^{-03}$
LDLc(mg/dl)	r	-0.120	-0.003	0.080	0.042	-0.051
	p	$1.21x10^{-01}$	$9.78x10^{-01}$	$3.12x10^{-01}$	$6.11x10^{-01}$	$5.47x10^{-01}$
VLDLc(mg/dl)	r	-0.438*	0.449*	0.140	0.318*	0.189*
	p	$4.83x10^{-10}$	$2.05x10^{-10}$	$7.32x10^{-02}$	$1.43x10^{-05}$	$1.53x10^{-02}$
TG (mg/dl)	r	-0.438*	0.444*	0.134	0.312*	0.187*
	p	$4.83x10^{-10}$	$2.91x10^{-10}$	$8.46x10^{-02}$	$2.16x10^{-05}$	$1.64x10^{-02}$

P-values adjusted by FDR. Significant values ($p<0.05$) are marked (*). Cho: Total cholesterol; HDLc: high density lipoprotein cholesterol; LDLc: low density lipoprotein cholesterol; VLDLc: very low-density lipoprotein cholesterol; TG: total triglycerides.

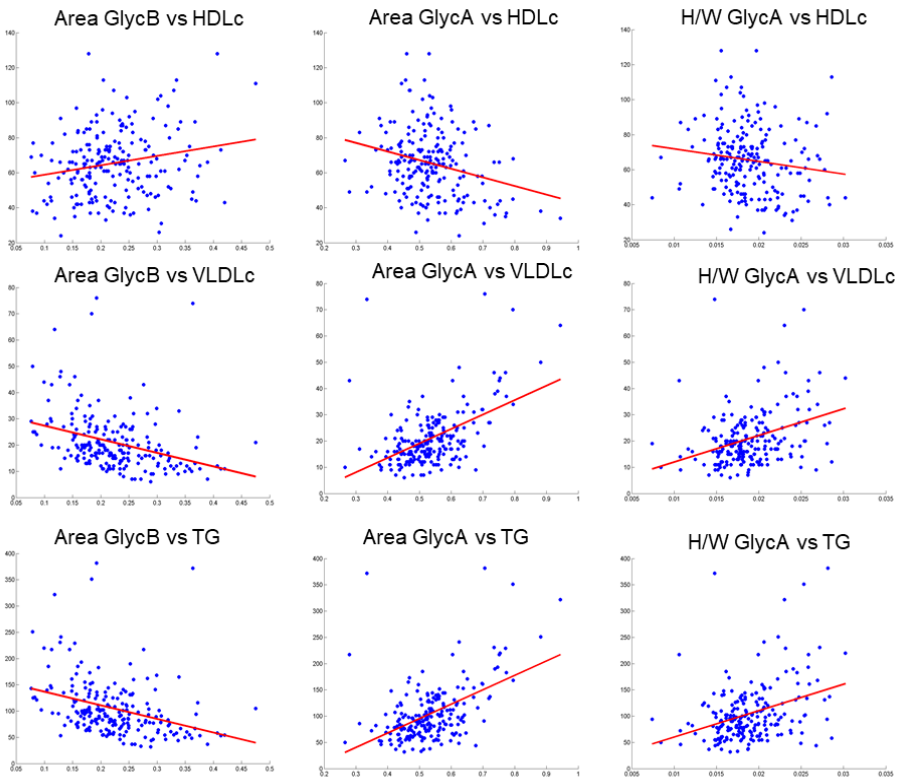


Figure C4.1S: Scatter plots of Lipids variables obtained by traditional biochemistry and glycoprotein variables.

Table C4.2S. Spearman correlation coefficients (r) and p-values (*p*) for each glycoprotein variable and the variables resulting from the Liposcale® test.

		Area GlycB	Area GlycA	H/W GlycB	H/W GlycA	Distance GlycB-GlycA
VLDL-C (mg/dl)	r	-0.381*	0.411*	0.181*	0.295*	0.207*
	p	$2.79x10^{-07}$	$2.37x10^{-08}$	$2.51x10^{-02}$	$1.19x10^{-04}$	$9.87x10^{-03}$
IDL-C (mg/dl)	r	-0.139	0.201*	0.128	0.252*	0.048
	p	$8.20x10^{-02}$	$1.21x10^{-02}$	$1.08x10^{-01}$	$1.23x10^{-03}$	$5.74x10^{-01}$
LDL-C (mg/dl)	r	-0.146	0.011	0.064	0.081	0.046
	p	$7.22x10^{-02}$	$9.04x10^{-01}$	$4.44x10^{-01}$	$3.16x10^{-01}$	$5.91x10^{-01}$
HDL-C (mg/dl)	r	0.076	-0.274*	-0.145	-0.230*	-0.240*
	p	$3.51x10^{-01}$	$4.07x10^{-04}$	$7.25x10^{-02}$	$3.63x10^{-03}$	$2.29x10^{-03}$
VLDL-TG (mg/dl)	r	-0.436*	0.393*	0.157	0.251*	0.167*
	p	$2.32x10^{-09}$	$1.00x10^{-07}$	$5.09x10^{-02}$	$1.27x10^{-03}$	$3.88x10^{-02}$
IDL-TG (mg/dl)	r	-0.153	0.269*	0.179*	0.306*	0.065
	p	$5.69x10^{-02}$	$5.53x10^{-04}$	$2.59x10^{-02}$	$6.42x10^{-05}$	$4.37x10^{-01}$
LDL-TG (mg/dl)	r	0.053	0.142	0.132	0.197*	0.126
	p	$5.44x10^{-01}$	$7.72x10^{-02}$	$9.97x10^{-02}$	$1.40x10^{-02}$	$1.15x10^{-01}$
HDL-TG (mg/dl)	r	-0.161*	0.204*	0.098	0.158	-0.093
	p	$4.71x10^{-02}$	$1.13x10^{-02}$	$2.32x10^{-01}$	$5.09x10^{-02}$	$2.56x10^{-01}$
VLDL-P (nmol/L)	r	-0.434*	0.361*	0.145	0.222*	0.158
	p	$2.89x10^{-09}$	$1.32x10^{-06}$	$7.25x10^{-02}$	$5.16x10^{-03}$	$5.09x10^{-02}$
Large VLDL-P (nmol/L)	r	-0.303*	0.383*	0.168*	0.290*	0.142
	p	$7.34x10^{-05}$	$2.39x10^{-07}$	$3.78x10^{-02}$	$1.64x10^{-04}$	$7.72x10^{-02}$
Medium VLDL-P (nmol/L)	r	-0.363*	0.443*	0.187*	0.336*	0.180*
	p	$1.13x10^{-06}$	$1.24x10^{-09}$	$1.98x10^{-02}$	$9.08x10^{-06}$	$2.59x10^{-02}$
Small VLDL-P (nmol/L)	r	-0.427*	0.329*	0.124	0.191*	0.142
	p	$5.07x10^{-09}$	$1.43x10^{-05}$	$1.20x10^{-01}$	$1.78x10^{-02}$	$7.76x10^{-02}$
LDL-P (nmol/L)	r	-0.144	0.098	0.092	0.157	0.087
	p	$7.33x10^{-02}$	$2.33x10^{-01}$	$2.56x10^{-01}$	$5.10x10^{-02}$	$2.88x10^{-01}$
Large LDL-P (nmol/L)	r	0.003	-0.043	0.003	0.018	-0.024
	p	$9.78x10^{-01}$	$6.11x10^{-01}$	$9.78x10^{-01}$	$8.37x10^{-01}$	$7.89x10^{-01}$
Medium LDL-P (nmol/L)	r	-0.121	0.048	0.139	0.141	0.082
	p	$1.32x10^{-01}$	$5.74x10^{-01}$	$8.24x10^{-02}$	$7.99x10^{-02}$	$3.16x10^{-01}$
Small LDL-P (nmol/L)	r	-0.178	0.161*	0.056	0.189*	0.095
	p	$2.72x10^{-02}$	$4.71x10^{-02}$	$5.14x10^{-01}$	$1.87x10^{-02}$	$2.46x10^{-01}$
HDL-P (µmol/L)	r	-0.067	-0.102	-0.089	-0.108	-0.227*
	p	$4.22x10^{-01}$	$2.10x10^{-01}$	$2.75x10^{-01}$	$1.88x10^{-01}$	$4.09x10^{-03}$
Large HDL-P (µmol/L)	r	-0.165*	0.106	0.039	0.048	0.157
	p	$4.17x10^{-02}$	$1.93x10^{-01}$	$6.41x10^{-01}$	$5.74x10^{-01}$	$5.09x10^{-02}$
	r	0.167*	-0.254*	-0.107	-0.202*	-0.169*

Medium HDL-P (µmol/L)	*p*	$3.91x10^{-02}$	$1.15x10^{-03}$	$1.89x10^{-01}$	$1.21x10^{-02}$	$3.62x10^{-02}$
Small HDL-P (µmol/L)	r	-0.178*	-0.006	-0.044	-0.039	-0.207*
	p	$2.67x10^{-02}$	$9.57x10^{-01}$	$6.09x10^{-01}$	$6.41x10^{-01}$	$9.98x10^{-03}$
VLDL-Z (nm)	r	0.189*	0.021	0.084	0.137	0.002
	p	$1.86x10^{-02}$	$8.12x10^{-01}$	$3.05x10^{-01}$	$8.54x10^{-02}$	$9.78x10^{-01}$
LDL-Z (nm)	r	0.264*	-0.286*	-0.028	-0.199*	-0.105
	p	$6.94x10^{-04}$	$1.95x10^{-04}$	$7.48x10^{-01}$	$1.33x10^{-02}$	$1.99x10^{-01}$
HDL-Z (nm)	r	0.306*	-0.231*	-0.082	-0.155	-0.035
	p	$6.42x10^{-05}$	$3.54x10^{-03}$	$3.16x10^{-01}$	$5.38x10^{-02}$	$6.78x10^{-01}$

P-values adjusted by FDR. Significant values ($p<0.05$) are marked (*).

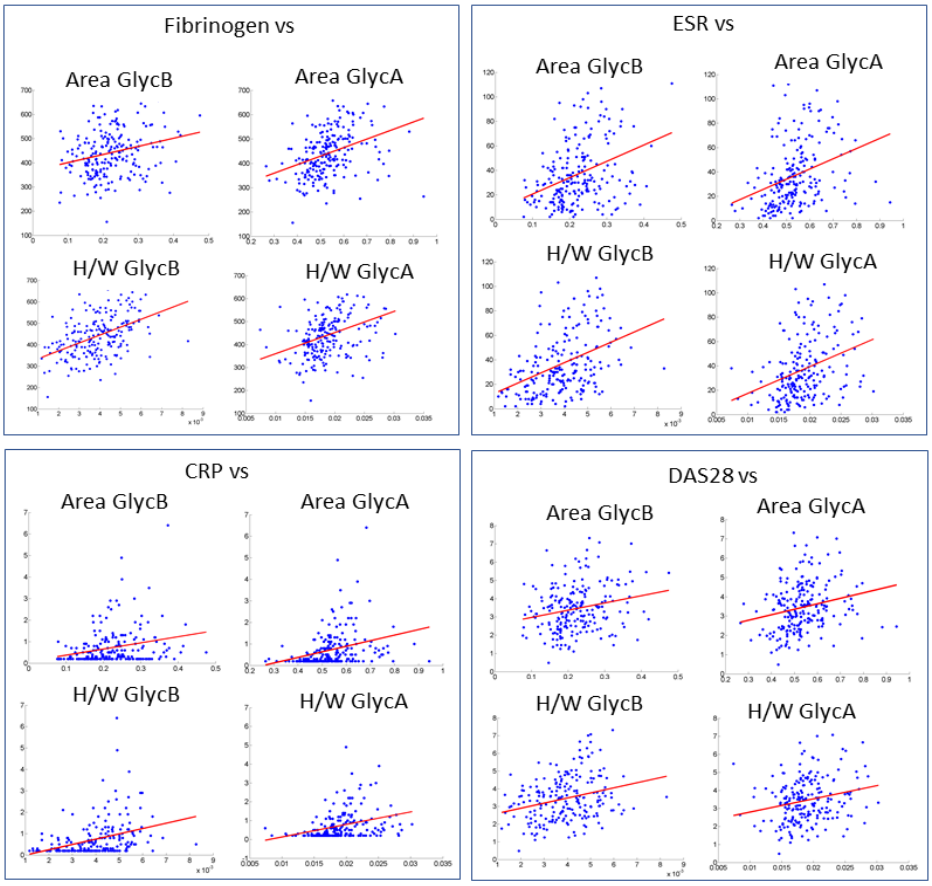


Figure C4.2S. Scatter plots of ^{1}H-NMR glycoproteins variables and clinical inflammatory markers.

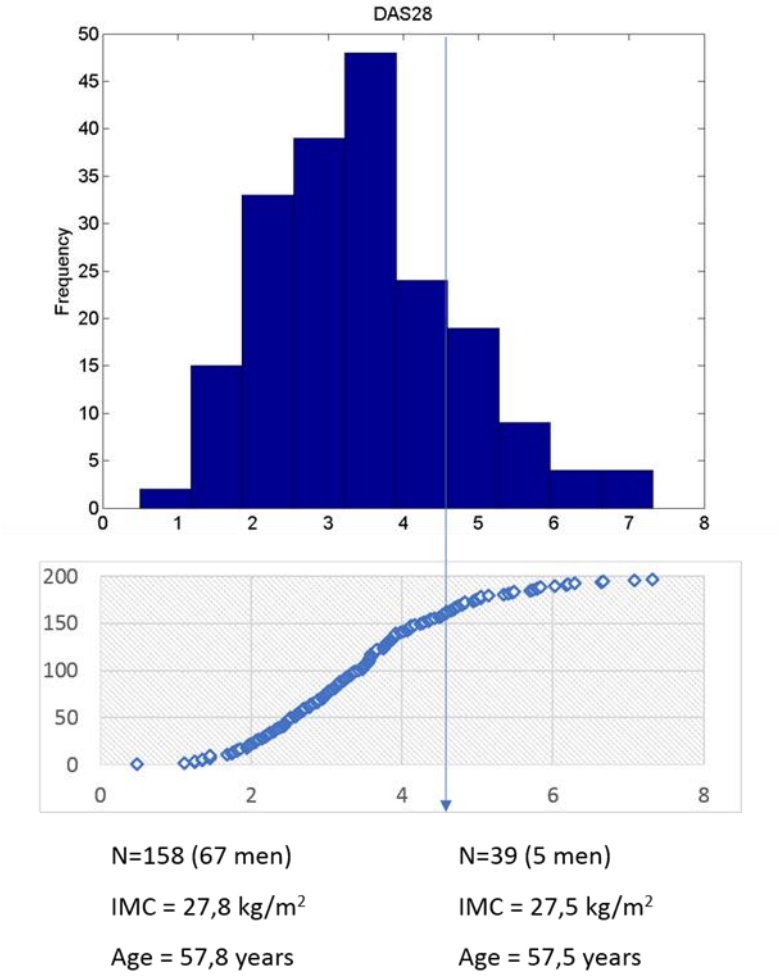


Figure C4.3S. Population distribution of DAS28 and a brief description of the population for the AR groups.

4.8. References

(1) McInnes, I. B.; Schett, G. The Pathogenesis of Rheumatoid Arthritis. *N. Engl. J. Med.* **2011**, *365* (23), 2205–2219.

(2) van de Stadt, L. A.; van der Horst, A. R.; de Koning, M. H. M. T.; Bos, W. H.; Wolbink, G. J.; van de Stadt, R. J.; Pruijn, G. J. M.; Dijkmans, B. A. C.; van Schaardenburg, D.; Hamann, D. The Extent of the Anti-Citrullinated Protein Antibody Repertoire Is Associated with Arthritis Development in Patients with Seropositive Arthralgia. *Ann. Rheum. Dis.* **2011**, *70* (1), 128–133.

(3) Steiner, G.; Urowitz, M. B. Lipid Profiles in Patients with Rheumatoid Arthritis: Mechanisms and the Impact of Treatment. *Semin. Arthritis Rheum.* **2009**, *38* (5), 372–381.

(4) Alamanos, Y.; Drosos, A. A. Epidemiology of Adult Rheumatoid Arthritis. *Autoimmun. Rev.* **2005**, *4* (3), 130–136.

(5) Gabriel, S. E. Why Do People with Rheumatoid Arthritis Still Die Prematurely? *Ann. Rheum. Dis.* **2008**, *67* (Suppl 3), iii30-iii34.

(6) Gabriel, S. E.; Crowson, C. S.; Kremers, H. M.; Doran, M. F.; Turesson, C.; O'Fallon, W. M.; Matteson, E. L. Survival in Rheumatoid Arthritis: A Population-Based Analysis of Trends over 40 Years. *Arthritis Rheum.* **2003**, *48* (1), 54–58.

(7) Castañeda, S.; Martín-Martínez, M. A.; González-Juanatey, C.; Llorca, J.; García-Yébenes, M. J.; Pérez-Vicente, S.; Sánchez-Costa, J. T.; Díaz-Gonzalez, F.; González-Gay, M. A. Cardiovascular Morbidity and Associated Risk Factors in Spanish Patients with Chronic Inflammatory Rheumatic Diseases Attending Rheumatology Clinics: Baseline Data of the CARMA Project. *Semin. Arthritis Rheum.* **2015**, *44* (6), 618–626.

(8) Stevens, R. J.; Douglas, K. M. J.; Saratzis, A. N.; Kitas, G. D. Inflammation and Atherosclerosis in Rheumatoid Arthritis. *Expert Rev. Mol. Med.* **2005**, *7* (07), 1–24.

(9) Choy, E. Understanding the Dynamics: Pathways Involved in the Pathogenesis of Rheumatoid Arthritis. *Rheumatology* **2012**, *51* (suppl 5), v3–v11.

(10) Amos, R. S.; Constable, T. J.; Crockson, R. A.; Crockson, A. P.; McConkey, B. Rheumatoid Arthritis: Relation of Serum C-Reactive Protein and Erythrocyte Sedimentation Rates to Radiographic Changes. *BMJ* **1977**, *1* (6055), 195–197.

(11) Prevoo, M. L. L.; Van'T Hof, M. A.; Kuper, H. H.; Van Leeuwen, M. A.; Van De Putte, L. B. A.; Van Riel, P. L. C. M. Modified Disease Activity Scores That Include Twenty-Eight-Joint Counts Development and Validation in a Prospective Longitudinal Study of Patients with Rheumatoid Arthritis. *Arthritis Rheum.* **1995**, *38* (1), 44–48.

(12) Gruys, E.; Toussaint, M. J. M.; Niewold, T. A.; Koopmans, S. J. Acute Phase Reaction and Acute Phase Proteins. *J. Zhejiang Univ. Sci. B* **2005**, *6* (11), 1045–1056.

(13) Lyons, J. J.; Milner, J. D.; Rosenzweig, S. D. Glycans Instructing Immunity: The Emerging Role of Altered Glycosylation in Clinical Immunology. *Front. Pediatr.* **2015**, *3*, 54.

(14) Marth, J. D.; Grewal, P. K. Mammalian Glycosylation in Immunity. *Nat. Rev. Immunol.* **2008**, *8* (11), 874–887.

(15) Gabay, C.; Kushner, I. Acute-Phase Proteins and Other Systemic Responses to Inflammation. *N. Engl. J. Med.* **1999**, *340* (6), 448–454.

(16) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **2006**, *126* (5), 855–867.

(17) van Kooyk, Y.; Rabinovich, G. a. Protein-Glycan Interactions in the Control of Innate and Adaptive Immune Responses. *Nat. Immunol.* **2008**, *9* (6), 593–601.

(18) Gornik, O.; Lauc, G. Glycosylation of Serum Proteins in Inflammatory Diseases. *Dis. Markers* **2008**, *25* (4–5), 267–278.

(19) McCarthy, C.; Saldova, R.; Wormald, M. R.; Rudd, P. M.; McElvaney, N. G.; Reeves, E. P. The Role and Importance of Glycosylation of Acute Phase Proteins with Focus on Alpha-1 Antitrypsin in Acute and Chronic Inflammatory Conditions. *J. Proteome Res.* **2014**, *13* (7), 3131–3143.

(20) Hart, G. W.; Copeland, R. J. Glycomics Hits the Big Time. *Cell* **2010**, *143* (5), 672–676.

(21) Connelly, M. A.; Gruppen, E. G.; Wolak-Dinsmore, J.; Matyus, S. P.; Riphagen, I. J.; Shalaurova, I.; Bakker, S. J. L.; Otvos, J. D.; Dullaart, R. P. F. GlycA, a Marker of Acute Phase Glycoproteins, and the Risk of Incident Type 2 Diabetes Mellitus: PREVEND Study. *Clin. Chim. Acta* **2016**, *452*, 10–17.

(22) Connelly, M. A.; Gruppen, E. G.; Otvos, J. D.; Dullaart, R. P. F. Inflammatory Glycoproteins in Cardiometabolic Disorders, Autoimmune Diseases and Cancer. *Clin. Chim. Acta* **2016**, *459*, 177–186.

(23) Soininen, P.; Kangas, A. J.; Würtz, P.; Tukiainen, T.; Tynkkynen, T.; Laatikainen, R.; Järvelin, M.-R.; Kähönen, M.; Lehtimäki, T.; Viikari, J.; et al. High-Throughput Serum NMR Metabonomics for Cost-Effective Holistic Studies on Systemic Metabolism. *Analyst* **2009**, *134* (9), 1781.

(24) Soininen, P.; Haarala, J.; Vepsäläinen, J.; Niemitz, M.; Laatikainen, R. Strategies for Organic Impurity Quantification by 1H NMR Spectroscopy: Constrained Total-Line-Shape Fitting. *Anal. Chim. Acta* **2005**, *542* (2), 178–185.

(25) Akinkuolie, A. O.; Buring, J. E.; Ridker, P. M.; Mora, S. A Novel Protein Glycan Biomarker and Future Cardiovascular Disease Events. *J. Am. Heart Assoc.* **2014**, *3* (5), e001221.

(26) Bell, J. D.; Brown, J. C. C.; Nicholson, J. K.; Sadler, P. J. Assignment of Resonances for 'Acute-Phase' Glycoproteins in High Resolution Proton NMR Spectra of Human Blood Plasma. *FEBS Lett.* **1987**, *215* (2), 311–315.

(27) Gruppen, E. G.; Riphagen, I. J.; Connelly, M. A.; Otvos, J. D.; Bakker, S. J. L.; Dullaart, R. P. F. GlycA, a Pro-Inflammatory Glycoprotein Biomarker, and Incident Cardiovascular Disease: Relationship with C-Reactive Protein and Renal Function. *PLoS One* **2015**, *10* (9), e0139057.

(28) Otvos, J. D.; Guyton, J. R.; Connelly, M. A.; Akapame, S.; Bittner, V.; Kopecky, S. L.; Lacy, M.; Marcovina, S. M.; Muhlestein, J. B.; Boden, W. E. Relations of GlycA and Lipoprotein Particle Subspecies with Cardiovascular Events and Mortality: A Post Hoc Analysis of the AIM-HIGH Trial. *J. Clin. Lipidol.* **2018**, *12* (2), 348–355.e2.

(29) Lawler, P. R.; Akinkuolie, A.; Buring, J.; Ridker, P.; Glynn, R.; Mora, S. A Novel Biomarker of Circulating Glycoproteins and Cardiovascular and All-Cause Mortality among 39,521 Initially Healthy Adults. *J. Am. Coll. Cardiol.* **2015**, *65* (10), A1358.

(30) Otvos, J. D.; Shalaurova, I.; Wolak-Dinsmore, J.; Connelly, M. A.; Mackey, R. H.; Stein, J. H.; Tracy, R. P. GlycA: A Composite Nuclear Magnetic Resonance Biomarker of Systemic Inflammation. *Clin. Chem.* **2015**, *61:5*, 714–723.

(31) Lawler, P. R.; Mora, S. Glycosylation Signatures of Inflammation Identify Cardiovascular Risk. *Circ. Res.* **2016**, *119* (11), 1154–1156.

(32) Akinkuolie, A. O.; Pradhan, A. D.; Buring, J. E.; Ridker, P. M.;

Mora, S. Novel Protein Glycan Side-Chain Biomarker and Risk of Incident Type 2 Diabetes Mellitus. *Arterioscler. Thromb. Vasc. Biol.* **2015**, *35* (6).

(33) Bertini, I.; Cacciatore, S.; Jensen, B. V.; Schou, J. V.; Johansen, J. S.; Kruhoffer, M.; Luchinat, C.; Nielsen, D. L.; Turano, P. Metabolomic NMR Fingerprinting to Identify and Predict Survival of Patients with Metastatic Colorectal Cancer. *Cancer Res.* **2012**, *72* (1), 356–364.

(34) Bairaktari, E.; Hatzidimou, K.; Tzallas, C.; Vini, M.; Katsaraki, A.; Tselepis, A.; Elisaf, M.; Tsolas, O. Estimation of LDL Cholesterol Based on the Friedewald Formula and on Apo B Levels. *Clin. Biochem.* **2000**, *33* (7), 549–555.

(35) Wu, P. S. C.; Otting, G. Rapid Pulse Length Determination in High-Resolution NMR. *J. Magn. Reson.* **2005**, *176* (1), 115–119.

(36) Nicholson, J. K.; Foxall, P. J. D.; Spraul, M.; Farrant, R. D.; Lindon, J. C. 750 MHz 1H and 1H-13C NMR Spectroscopy of Human Blood Plasma. *Anal. Chem.* **1995**, *67* (5), 793–811.

(37) Mallol, R.; Amigó, N.; Rodríguez, M. A.; Heras, M.; Vinaixa, M.; Plana, N.; Rock, E.; Ribalta, J.; Yanes, O.; Masana, L.; et al. Liposcale: A Novel Advanced Lipoprotein Test Based on 2D Diffusion-Ordered 1H NMR Spectroscopy. *J. Lipid Res.* **2015**, *56* (3), 737–746.

(38) Jeyarajah, E. J.; Cromwell, W. C.; Otvos, J. D. Lipoprotein Particle Analysis by Nuclear Magnetic Resonance Spectroscopy. *Clin. Lab. Med.* **2006**, *26* (4), 847–870.

(39) Rosenson, R. S.; Brewer, H. B.; Chapman, M. J.; Fazio, S.; Hussain, M. M.; Kontush, A.; Krauss, R. M.; Otvos, J. D.; Remaley, A. T.; Schaefer, E. J. HDL Measures, Particle Heterogeneity, Proposed Nomenclature, and Relation to Atherosclerotic Cardiovascular Events. *Clin. Chem.* **2011**, *57* (3),

392–410.

(40) Mallol, R.; Rodríguez, M. A.; Heras, M.; Vinaixa, M.; Plana, N.; Masana, L.; Morris, G. A.; Correig, X. Particle Size Measurement of Lipoprotein Fractions Using Diffusion-Ordered NMR Spectroscopy. *Anal. Bioanal. Chem.* **2012**, *402* (7), 2407–2415.

(41) Schumaker, V. N.; Phillips, M. L.; Chatterton, J. E. Apolipoprotein B and Low-Density Lipoprotein Structure: Implications for Biosynthesis of Triglyceride-Rich Lipoproteins. *Adv. Protein Chem.* **1994**, *45*, 205–248.

(42) Mallol, R.; Rodríguez, M. A.; Heras, M.; Vinaixa, M.; Cañellas, N.; Brezmes, J.; Plana, N.; Masana, L.; Correig, X. Surface Fitting of 2D Diffusion-Edited 1H NMR Spectroscopy Data for the Characterisation of Human Plasma Lipoproteins. *Metabolomics* **2011**, *7* (4), 572–582.

(43) Yekutieli, D.; Benjamini, Y. Resampling-Based False Discovery Rate Controlling Multiple Test Procedures for Correlated Test Statistics. *J. Stat. Plan. Inference* **1999**, *82* (1–2), 171–196.

(44) Fransen, J.; Stucki, G.; van Riel, P. L. C. M. Rheumatoid Arthritis Measures: Disease Activity Score (DAS), Disease Activity Score-28 (DAS28), Rapid Assessment of Disease Activity in Rheumatology (RADAR), and Rheumatoid Arthritis Disease Activity Index (RADAI). *Arthritis Rheum.* **2003**, *49* (S5), S214–S224.

(45) Chipperfield, A. J. The MATLAB Genetic Algorithm Toolbox. In *IEE Colloquium on Applied Control Techniques Using MATLAB*; IEE, 1995; Vol. 1995, pp 10–10.

(46) Bartlett, D. B.; Connelly, M. A.; AbouAssi, H.; Bateman, L. A.; Tune, K. N.; Huebner, J. L.; Kraus, V. B.; Winegar, D. A.; Otvos, J. D.; Kraus, W. E.; et al. A Novel Inflammatory Biomarker,

GlycA, Associates with Disease Activity in Rheumatoid Arthritis and Cardio-Metabolic Risk in BMI-Matched Controls. *Arthritis Res. Ther.* **2016**, *18* (1), 86.

(47) Lauridsen, M. B.; Bliddal, H.; Christensen, R.; Danneskiold-Samsøe, B.; Bennett, R.; Keun, H.; Lindon, J. C.; Nicholson, J. K.; Dorff, M. H.; Jaroszewski, J. W.; et al. ^{1}H NMR Spectroscopy-Based Interventional Metabolic Phenotyping: A Cohort Study of Rheumatoid Arthritis Patients. *J. Proteome Res.* **2010**, *9* (9), 4545–4553.

(48) Choy, E. H. S.; Panayi, G. S. Cytokine Pathways and Joint Inflammation in Rheumatoid Arthritis. *N. Engl. J. Med.* **2001**, *344* (12), 907–916.

(49) Dawes, P. T.; Fowler, P. D.; Clarke, S.; Fisher, J.; Lawton, A.; Shadforth, M. F. Rheumatoid Arthritis: Treatment Which Controls the C-Reactive Protein and Erythrocyte Sedimentation Rate Reduces Radiological Progression. *Rheumatology* **1986**, *25* (1), 44–49.

(50) Niwa, Y.; Iio, A.; Niwa, G.; Sakane, T.; Tsunematsu, T.; Kanoh, T. Serum Albumin Metabolism in Rheumatic Diseases: Relationship to Corticosteroids and Peptic Ulcer. *J. Clin. Lab. Immunol.* **1990**, *31* (1), 11–16.

(51) Barter, P. J.; Puranik, R.; Rye, K.-A. New Insights into the Role of HDL as an Anti-Inflammatory Agent in the Prevention of Cardiovascular Disease. *Curr. Cardiol. Rep.* **2007**, *9* (6), 493–498.

(52) Barter, P.; Gotto, A. M.; LaRosa, J. C.; Maroni, J.; Szarek, M.; Grundy, S. M.; Kastelein, J. J. P.; Bittner, V.; Fruchart, J.-C. HDL Cholesterol, Very Low Levels of LDL Cholesterol, and Cardiovascular Events. *N. Engl. J. Med.* **2007**, *357* (13), 1301–1310.

(53) Navab, M.; Reddy, S. T.; Van Lenten, B. J.; Fogelman, A. M. HDL and Cardiovascular Disease: Atherogenic and Atheroprotective Mechanisms. *Nat. Rev. Cardiol.* **2011**, *8* (4), 222–232.

(54) MacDougall, E. D.; Kramer, F.; Polinsky, P.; Barnhart, S.; Askari, B.; Johansson, F.; Varon, R.; Rosenfeld, M. E.; Oka, K.; Chan, L.; et al. Aggressive Very Low-Density Lipoprotein (VLDL) and LDL Lowering by Gene Transfer of the VLDL Receptor Combined with a Low-Fat Diet Regimen Induces Regression and Reduces Macrophage Content in Advanced Atherosclerotic Lesions in LDL Receptor-Deficient Mice. *Am. J. Pathol.* **2006**, *168* (6), 2064–2073.

(55) Lawler, P. R.; Akinkuolie, A. O.; Chu, A. Y.; Shah, S. H.; Kraus, W. E.; Craig, D.; Padmanabhan, L.; Glynn, R. J.; Ridker, P. M.; Chasman, D. I.; et al. Atherogenic Lipoprotein Determinants of Cardiovascular Disease and Residual Risk among Individuals with Low Low-Density Lipoprotein Cholesterol. *J. Am. Heart Assoc.* **2017**, *6* (7), e005549.

(56) Lawler, P. R.; Akinkuolie, A. O.; Harada, P.; Glynn, R. J.; Chasman, D. I.; Ridker, P. M.; Mora, S. Residual Risk of Atherosclerotic Cardiovascular Events in Relation to Reductions in Very-Low-Density Lipoproteins. *J. Am. Heart Assoc.* **2017**, *6* (12), e007402.

(57) Lawler, P. R. Glycomics and Cardiovascular Disease. *Circ. Res.* **2018**, *122* (11), 1488–1490.

(58) Akinkuolie, A. O.; Pradhan, A. D.; Buring, J. E.; Ridker, P. M.; Mora, S. Novel Protein Glycan Side-Chain Biomarker and Risk of Incident Type 2 Diabetes MellitusSignificance. *Arterioscler. Thromb. Vasc. Biol.* **2015**, *35* (6), 1544–1550.

(59) McGarrah, R. W.; Kelly, J. P.; Craig, D. M.; Haynes, C.; Jessee, R. C.; Huffman, K. M.; Kraus, W. E.; Shah, S. H. A Novel

Protein Glycan-Derived Inflammation Biomarker Independently Predicts Cardiovascular Disease and Modifies the Association of HDL Subclasses with Mortality. *Clin. Chem.* **2017**, *63* (1), 288–296.

(60) Lawler, P. R.; Akinkuolie, A. O.; Chandler, P. D.; Moorthy, M. V.; Vandenburgh, M. J.; Schaumberg, D. A.; Lee, I.-M.; Glynn, R. J.; Ridker, P. M.; Buring, J. E.; et al. Circulating N-Linked Glycoprotein Acetyls and Longitudinal Mortality RiskNovelty and Significance. *Circ. Res.* **2016**, *118* (7), 1106–1115.

(61) Menni, C.; Gudelj, I.; Macdonald-Dunlop, E.; Mangino, M.; Zierer, J.; Bešić, E.; Joshi, P. K.; Trbojević-Akmačić, I.; Chowienczyk, P. J.; Spector, T. D.; et al. Glycosylation Profile of Immunoglobulin G Is Cross-Sectionally Associated With Cardiovascular Disease Risk Score and Subclinical Atherosclerosis in Two Independent Cohorts. *Circ. Res.* **2018**, *122* (11), 1555–1564.

(62) Väänänen, T.; Vuolteenaho, K.; Kautiainen, H.; Nieminen, R.; Möttönen, T.; Hannonen, P.; Korpela, M.; Kauppi, M. J.; Laiho, K.; Kaipiainen-Seppänen, O.; et al. Glycoprotein YKL-40: A Potential Biomarker of Disease Activity in Rheumatoid Arthritis during Intensive Treatment with CsDMARDs and Infliximab. Evidence from the Randomised Controlled NEO-RACo Trial. *PLoS One* **2017**, *12* (8), e0183294.

(63) Rietzschel, E.; De Buyzere, M. High-Sensitive C-Reactive Protein: Universal Prognostic and Causative Biomarker in Heart Disease? *Biomark. Med.* **2012**, *6* (1), 19–34.

(64) Gerszten, R. E.; Wang, T. J. The Search for New Cardiovascular Biomarkers. *Nature* **2008**, *451* (7181), 949–952.

(65) Hoefer, I. E.; Steffens, S.; Ala-Korpela, M.; Bäck, M.; Badimon, L.; Bochaton-Piallat, M.-L.; Boulanger, C. M.; Caligiuri, G.; Dimmeler, S.; Egido, J.; et al. Novel Methodologies for

Biomarker Discovery in Atherosclerosis. *Eur. Heart J.* **2015**, *36* (39), 2635–2642.

CHAPTER 5

5. Influence of sexual steroids and obesity on glycoprotein profiles

5.1. Abstract

The polycystic ovary syndrome (PCOS) is a common endocrine disorder affecting women in reproductive age. Obesity and low-grade chronic inflammation are frequently associated with PCOS. Recently, proton nuclear magnetic resonance (^{1}H-NMR)-derived glycoprotein profiles have emerged as potential biomarkers that reflect systemic inflammation in type 2 diabetes, obesity and other pathological processes. The aim of this work is to study plasma glycoprotein profiles as metabolic/inflammatory biomarkers that links PCOS with inflammation and obesity.

We evaluated five glycoprotein variables, namely GlycA, GlycB and GlycF and the height to width (H/W) ratio of GlycA and GlycB, by using 1H-NMR spectroscopy in non-obese and obese women with PCOS and compared them with male and female control subjects.

Obese subjects presented higher GlycA, GlycB and GlycF areas and higher H/W GlycA and GlycB ratios than their non-obese counterparts, regardless of sex and PCOS. Patients with PCOS also presented greater H/W ratios of GlycA and GlycB compared with the female and male control groups. All glycoprotein variables were associated with hs-CRP, showing different correlations among group of subjects.

Even though additional research is required to confirm our findings, the results of the current exploratory investigation pave the way toward the use of ^{1}H-NMR-derived glycoprotein profiles as

diagnostic markers to predict chronic low-grade inflammation in patients with PCOS.

5.2. Introduction

Low-grade chronic inflammation is involved in the pathogenesis of certain metabolic disorders such as obesity, abdominal adiposity, metabolic syndrome, type 2 diabetes and polycystic ovary syndrome (PCOS).[1,2] Adipose tissue dysfunction, resulting into the secretion of proinflammatory adipokines, cytokines and other mediators, is known to play a major role in metabolic disorders yet precise mechanisms remain not fully established.[3] PCOS is possibly the most common endocrine and metabolic disorder in women of reproductive age. Some PCOS phenotypes are characterized by adipose tissue dysfunction, visceral obesity, and insulin resistance (IR) being androgen excess its pivotal pathogenic mechanism.[2]

C-reactive protein (CRP), tumor necrosis factor-α (TNF-α), interleukin-6 (IL-6) and leptin have been conventionally used as circulating inflammatory markers.[4-6] Even though an increase in these inflammatory molecules has been reported in women with PCOS, it is unclear whether or not this finding is specific of PCOS or is the result of the frequent association of this syndrome with obesity.[7-9] In addition, circulating inflammatory markers may be subjected to dramatic and rapid intra-individual fluctuations, inter-individual variability and differences in the sensitivity of the assays used for their

measurement,[10,11] making the interpretation of previous results on the issue particularly difficult.

Glycoproteins play a key role in inflammatory processes.[10,11] Inflammation is characterized by protein structural changes that modify the number of antennary branches [rich in N-acetylglucosamine/N-acetylgalactosamine (GlycA) and N-acetylneuraminic acid (GlycB)] resulting into circulating glycoprotein profiles that may be used as potential early biomarkers of inflammation.[13] Proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy has demonstrated to be a quick and accurate analytical technique able to quantify metabolites, lipoproteins, and glycoproteins.[14,15] After the seminal work by Bell et al. in 1987,[16] several studies have reported the role of GlycA as a proinflammatory biomarker associated with disorders such as type 2 diabetes,[17] colorectal cancer and cardiovascular events,[18] among others.

At present, the influence of obesity, sex and sex hormones on the glycosylation changes related to inflammation is unclear. Hence, we aimed to delineate the glycoprotein profiles of non-obese and obese women with PCOS comparing them with non-obese and obese female and male controls. These comprehensive profiles included the quantification of the acetyl groups of N-acetylglucosamine and N-acetylgalactosamine (GlycA) and N-acetylneuraminic acid (GlycB) of the main circulating glycoproteins, as well as height/weight (H/W) parameters derived from the shape of the GlycA and GlycB signals, since higher and narrower signals have been related to inflammation.[14]

5.3. Material and Methods

5.3.1 Subjects

We studied a population of 53 young adults, including 17 patients with PCOS, 17 control women and 19 control men. The subjects included in each subgroup were selected in order to be similar in terms of age and body mass index (BMI), and were grouped into non-obese (BMI < 30 kg/m2, n = 28) and obese (BMI ≥ 30 kg/m2, n = 25) subgroups. The diagnosis of PCOS required the presence of clinical and/or biochemical hyperandrogenism, menstrual dysfunction, and exclusion of other disorders.[19] Control women showed no evidence of hyperandrogenism or ovulatory dysfunction. Before enrollment, the participants had no history of obesity-associated comorbidities including disorders of glucose tolerance, hypertension, sleep apnea, cardiovascular disease, or male hypogonadism. Smokers were excluded. All subjects provided informed written consent and the study was approved by the local ethics committee.

5.3.2. Phenotyping of subjects

All individuals underwent a comprehensive clinical, anthropometric and physical evaluation. In addition, they followed the same diet for a few days before the start of the study. Samples (serum and plasma) were obtained after a 12-h overnight fasting (during the follicular phase of the menstrual cycle in women). Samples were immediately assayed or frozen at -80°C until biochemical and ^{1}H-NMR analysis. These samples were used for the measurement of total testosterone (T), total estradiol (E2), sex hormone-binding globulin (SHBG), glucose, insulin and high-sensitivity C-reactive protein (hs-

CRP). The technical characteristics of the assays have been described in detail elsewhere.[20,21] Fasting glucose and insulin levels were used for homeostasis model assessment of insulin resistance (HOMA-IR).[22] In addition, the composite insulin sensitivity index (ISI) was estimated from the glucose and insulin concentrations measured during a standard 75 g oral glucose test.[23]

5.2.3. Glycoprotein profiling

Before ^{1}H-NMR analysis, 200 µL of serum were diluted with 50 µl deuterated water and 300 µl of 50 mM phosphate buffer solution (PBS) at pH 7.4. 1H-NMR spectra were recorded at 310 K on a Bruker Avance III 600 spectrometer (Bruker BioSciences Española S.A., Rivas Vaciamadrid, Madrid, Spain) operating at a proton frequency of 600.20 MHz (14.1 T) following previously reported procedures.[14] Briefly, we analyzed the region of the ^{1}H-NMR spectrum where the glycoproteins resonate (2.15-1.90 ppm) using several functions. In this study, we included a new function termed GlycF, between GlycA and GlycB. For each function, we determined the total area (proportional to concentration), height, position and bandwidth. The area of GlycA provided the concentration of acetyl groups of protein-bound N-acetylglucosamine and N-acetylgalactosamine, and the area of GlycB those of N-acetylneuraminic acid. GlycF area arises from the concentration of the acetyl groups of N-acetylglucosamine, N-acetylgalactosamine and N-acetylneuraminic acid unbound to proteins (free fraction). H/W ratios of GlycA and GlycB translate the shape of the peaks present in the ^{1}H-NMR spectra into information based on height – which is related to concentration – and width –which is related

to the flexibility and the aggregation of the molecules generating the signal.[14] In order to provide optimal line shape fits of the GlycA and GlycB signals, the bandwidth of GlycF is constrained in the deconvolution model and thus, only the area is of analytical interest but not the H/W ratio.

5.2.4. Statistical analysis

Phenotyp variables are expressed as mean ± standard deviation (SD). Normality of continuous variables was assessed by the Kolmogorov-Smirnov method and logarithmic transformation was applied to ensure a normal distribution when appropriate. Two-way univariate general linear models (GLMs) were used to determine the influence of group of subjects, obesity and their interaction on hormonal and metabolic variables. The least significant difference post hoc test was used for multiple comparisons among group of subjects. The relationships between continuous variables were assessed by Pearson's or Spearman's correlation analysis as appropriate. SPSS Statistic 15.0 (SPSS Inc., Chicago, IL) was used for data analysis setting alpha at the 0.05 level.

The chemometric methods evaluated in this work have been widely described in the literature.[14] Briefly, multivariate statistical analyses were computed in MATLAB, Ver. 7.10.0 using PLS-Toolbox, Ver. 5.2.2 (Eigenvector Research Inc., Manson, WA, United States). We applied principal component analysis (PCA) as an exploratory data method in order to detect groups of data and outliers.[24] Partial Least Square Discriminant Analysis (PLS-DA) was used as supervised classification method. While PCA ignores the information

regarding the class labels of the samples, PLS-DA relates the X matrix (experimental data) and the Y matrix (classes of samples) in order to find the maximum discrimination between classes and the maximum covariance between the X and Y matrices simultaneously.[25]

5.3. Results

5.3.1. Effects of group and obesity on phenotype

The demographic, clinical and biochemical characteristics of the population being studied are summarized in **Table C5.1**. As expected from the study design, no differences were found between patients with PCOS and female and male controls in terms of age and BMI.

Table C5.1. Clinical, biochemical/metabolic and hormonal characteristics of participants.

	Control women		Patients with PCOS		Control men		Group	Obesity	Interaction
	Non-obese (n = 9)	Obese (n = 8)	Non-obese (n = 9)	Obese (n = 8)	Non-obese (n = 10)	Obese (n = 9)	*P*	*P*	*P*
Age (years)	26 ± 5	27 ± 6	24 ± 8	30 ± 4	24 ± 5	25 ± 4	0.342	0.101	0.319
BMI (kg/m^2)	23 ± 2	36 ± 4	24 ± 2	37 ± 5	23 ± 2	34 ± 3	0.226	< 0.001	0.782
Waist to hip ratio	0.75 ± 0.08	0.83 ± 0.12	0.73 ± 0.05	0.85 ± 0.06	0.83 ± 0.04	0.90 ± 0.05	0.002 [B,C]	< 0.001	0.436
Hirsutism	1.4 ± 1.3	1.8 ± 0.4	9.7 ± 4.5	9.3 ± 4.5	-	-	< 0.001 [A]	0.738	0.876
Total testosterone (nmol/l)	1.5 ± 0.3	2.0 ± 0.5	2.4 ± 0.7	2.4 ± 1.1	18.5 ± 3.3	12.2 ± 3.5	0.001 [A,B,C]	0.776	0.196
Free testosterone (pmol/l)	21 ± 7	31 ± 8	36 ± 12	45 ± 24	450 ± 104	464 ± 94	< 0.001 [A,B,C]	0.024	0.265
Total estradiol (pmol/l)	149 ± 63	276 ± 200	182 ± 201	149 ± 49	68 ± 16	94 ± 26	< 0.001 [B,C]	0.024	0.422
Free estradiol (pmol/l)	2.7 ± 1.1	5.3 ± 3.0	3.6 ± 4.3	3.4 ± 1.4	1.8 ± 0.5	2.6 ± 0.7	0.010 [B]	0.003	0.520
Ratio FT/FE$_2$	8.6 ± 0.9	7.6 ± 1.6	17.4 ± 4.0	14.0 ± 1.9	263.4 ± 22.6	186.4 ± 12.7	< 0.001 [A,B,C]	0.119	0.750
SHBG (nmol/l)	56 ± 25	43 ± 14	50 ± 21	32 ± 13	27 ± 10	20 ± 6	< 0.001 [B,C]	0.008	0.568
Fasting insulin (pmol/l)	55.2 ± 20.0	77.0 ± 21.1	52.3 ± 28.8	94.4 ± 24.9	40.2 ± 10.6	74.5 ± 27.4	0.123	< 0.001	0.440

Fasting glucose (mmol/l)	4.7 ± 0.4	5.3 ± 0.4	4.5 ± 0.5	4.8 ± 0.5	4.9 ± 0.5	5.2 ± 0.4	0.011 [A,C]	0.001	0.408
HOMA-IR	1.6 ± 0.6	2.6 ± 0.7	1.5 ± 0.9	2.9 ± 0.7	1.3 ± 0.4	2.5 ± 1.0	0.424	< 0.001	0.742
Insulin sensitivity index	6.6 ± 2.9	3.3 ± 1.2	8.1 ± 4.7	3.5 ± 1.4	7.3 ± 2.8	3.8 ± 1.6	0.640	< 0.001	0.909
hs-CRP (nmol/L)	27 ± 22	38 ± 28	20 ± 21	65 ± 74	31 ± 25	31 ± 12	0.998	0.009	0.207

Abbreviations: BMI, body mass index; FE2, free estradiol; FT, free testosterone; HOMA-IR, homeostasis model assessment of insulin resistance; hs-CRP, high sensitive C-reactive protein; SHBG, sex hormone-binding globulin. Data are means ± SD. The effects of group and obesity on continuous variables were analyzed by a two-way GLM after applying logarithmic transformation when needed. A: $p < 0.05$ for the differences between control women and women with PCOS. B: $p < 0.05$ for the differences between control women and men. C: $p < 0.05$ for the differences between women with PCOS and men.

Regardless of the group of subjects, obesity was characterized by increased circulating concentrations of free T, total and free E2, hs-CRP, fasting insulin and glucose and larger HOMA-IR, and by decreased SHBG and ISI values, compared with non-obese subjects. Patients with PCOS presented both clinical and biochemical hyperandrogenism characterized by increased hirsutism score, and total and free testosterone levels compared with control women. Besides their higher total T and free T levels, men also had higher waist circumference (WC), waist to hip ratio (WHR), and free T/free E2 ratio compared to both groups of women. Conversely, patients with PCOS and control women showed higher circulating E2 and SHBG concentrations than men. We did not observe any statistically significant interactions between group of subjects and obesity on these variables.

5.3.2. Glycoprotein profile: exploratory data analysis

To assess similarities and differences among subject in terms of their glycoprotein profiles, we first applied a PCA using GlycA, GlycB and GlycF areas and H/W ratios of GlycA and GlycB as input variables. We defined our input matrix X, containing 53 observation row vectors of five variables each. **Figure C5.1** shows the glycoprotein profile of all participants plotted according to the values of the PCA scores and loadings. In particular, the scores of the second PC are presented against the first PC. The first two PCs accumulated 95.7% of the total variance. We observed that most of the scores of obese subjects lied within the first and fourth quadrants whereas non-obese

scores preferentially lied within the second and third; suggesting that the glycoprotein profile is influenced by obesity.

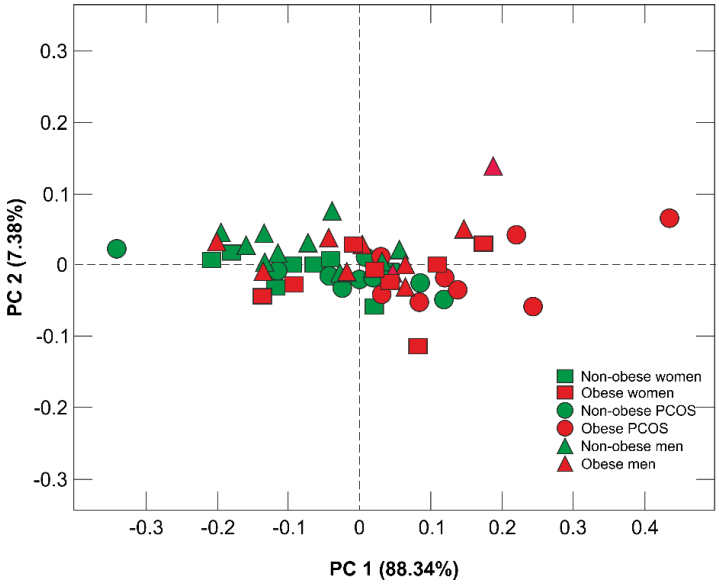


Figure C5.1. Principal Component Analysis scatter plot scores, considering all subjects as a whole (n=53). The variability in the data is mainly explained by the first Principal Component (PC). PCOS, polycystic ovary syndrome.

5.3.3. Effects of obesity on the glycoprotein profiles

Based on the results of the PCA and keeping X as input matrix, we applied a PLS-DA model to optimize the separation between non-obese and obese subjects. Two classes were taken into account to create the target Y matrix. Panel A of **Figure C5.2** shows the score plot of the PLS-DA model considering the first two latent variables (LVs). Compared with the PCA model, PLS-DA provided a better

discrimination between non-obese and obese individuals, regardless of the group of subjects. PLS-DA model was cross validated in order to avoid the risk of over-fitting and to increase the predictive ability by using the venetian blinds method.

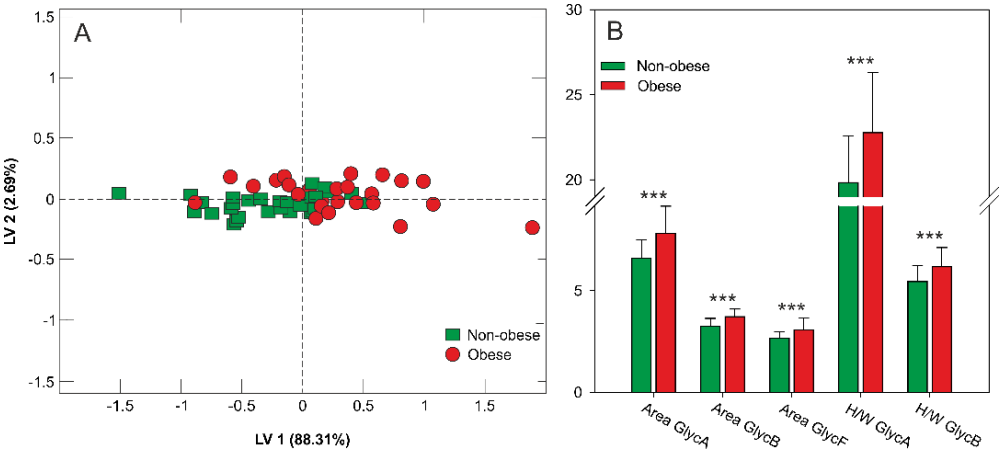


Figure C5.2. Panel A, Partial Least Square Discriminant Analysis scores of non-obese and obese subjects. Panel B, univariate GLM glycoprotein analysis. Data are mean ± SEM. $^{***}p < 0.001$ for differences between non-obese and obese participants, regardless of the group of subjects. H/W, height to width ratio; LV, latent variable.

The results of the univariate analysis of glycoprotein are detailed in **Table C5.2,** which describes not only the influence of group of subjects and obesity independently, but also their interaction. Regardless of sex and PCOS, obese subjects presented with larger GlycA, GlycB, GlycF areas and higher H/W ratios of GlycA and GlycB compared with their non-obese counterparts (**Figure C5.2**, panel B). Accordingly, considering all subjects as a whole, all glycoprotein variables correlated positively BMI, WC, HOMA-IR and

hs-CRP and negatively with insulin sensitivity index (**Supporting information Table C5S.1**).

Table C5.2. Univariate analysis of glycoprotein variables

	Control women		Patients with PCOS		Control men		Group	Obesity	Interaction
	Nonobese (n = 9)	Obese (n = 8)	Nonobese (n = 9)	Obese (n = 8)	Nonobese (n = 10)	Obese (n =9)	*P*	*P*	*P*
Area GlycA	6.54 ± 0.93	7.35 ± 1.24	6.86 ± 1.00	8.62 ± 1.39	6.39 ± 0.81	7.53 ± 1.15	0.059	< 0.001	0.442
Area GlycB	3.28 ± 0.32	3.69 ± 0.33	3.40 ± 0.50	4.00 ± 0.33	3.16 ± 0.31	3.56 ± 0.35	0.028 [B]	< 0.001	0.668
Area GlycF	2.57 ± 0.23	2.80 ± 0.43	2.65 ± 0.42	3.32 ± 0.64	2.72 ± 0.29	3.08 ± 0.59	0.140	0.001	0.334
H/W GlycA	19.63 ± 2.32	21.48 ± 2.70	20.78 ± 3.47	25.49 ± 3.70	19.26 ± 2.39	21.57 ± 2.98	0.014 [A,B]	0.001	0.322
H/W GlycB	5.43 ± 0.69	5.86 ± 0.78	5.73 ± 1.09	6.99 ± 0.86	5.26 ± 0.50	5.79 ± 0.75	0.007 [A,B]	0.001	0.248

Abbreviations: H/W, height to width ratio.

Glyc areas are expressed as arbitrary units. Data are means ± SD. The effects of group and obesity on continuous variables were analyzed by a two-way GLM after applying logarithmic transformation when needed. A: $p < 0.05$ for the differences between control women and women with PCOS. B: $p < 0.05$ for the differences between women with PCOS and men.

5.3.4. Differences between the groups of subjects in the glycoprotein profilescoprotein profiles

We applied a second PLS-DA model to discriminate among patients with PCOS, control women, and control men. In this case, three classes were taken into account to create the target Y matrix corresponding to each group. This model was also cross-validated using the venetian blinds method. The PLS-DA score plot showed certain separation between women with PCOS and female and male controls regardless of obesity (**Figure C5.3, Panel A**).

The results of univariate analysis of the glycoprotein profiles showed that the H/W ratios of GlycA and GlycB were significantly higher in women with PCOS compared with the male and female control groups (**Figure C5.3, panel B**). A similar behavior was observed for GlycA and GlycB areas, reaching statistical significance for GlycB, where patients with PCOS presented a larger area compared with men **(Table C2.2 and Figure C5.3, panel B**). Interestingly, we observed different correlations when subjects were classified according to group. All glycoprotein variables correlated negatively with ISI in control women and men, yet not in the PCOS subgroup. Strong correlations for HOMA-IR were observed in men and to a lesser extent in PCOS patients being negligible in control women. Finally, the correlation observed for hs-CRP persisted only in the PCOS and men subgroups (**Supporting information Table C5S.2**).

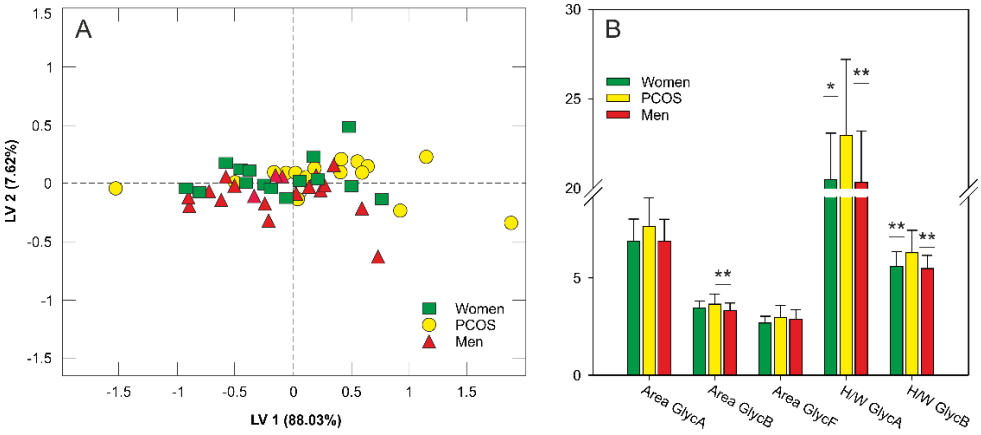


Figure C5.3. Panel A, Partial Least Square Discriminant Analysis scores of women with PCOS, control women and control men. Panel B, univariate GLM glycoprotein analysis. Data are mean ± SEM. *p < 0.05 and **p < 0.01 for differences between women with PCOS and female and male controls, regardless of obesity. H/W, height to width ratio; LV, latent variable; PCOS, polycystic ovary syndrome.

5.4. Discussion

Our present results indicate that circulating glycoprotein variables of otherwise healthy young adults are influenced mostly by obesity and, to a lesser extent, by the presence of PCOS in women.

Obesity associated with increases in both the areas and H/W ratios of all the glycoprotein profiles studied here. To date, the effect of obesity on circulating glycoproteins has not been extensively examined and most studies measured only GlycA. GlycA levels were consistently associated with obesity in different populations including premenopausal women,[26] pregnant women[27] and adolescents.[28,29]

Moreover, strong correlations between GlycA and surrogate indexes of obesity, insulin resistance and inflammation have been reported in both adolescent and adult populations, including correlations with BMI,[17,28,30-32] the leptin/adiponectin ratio[31] and abdominal fat amount.[33] The intimate association between obesity and GlycA levels is further reinforced by the fact that GlycA levels decreased after interventions directed towards reducing BMI such as lifestyle changes[29,30] or bariatric surgery.[26]

In agreement, obese subjects in our series presented with higher GlycA levels compared their non-obese counterparts, and GlycA levels correlated positively with surrogate indexes of obesity, abdominal adiposity, insulin resistance and with serum hs-CRP concentrations, usually associated with low-grade chronic inflammation.

Data regarding other glycoprotein variables are scarce. To the best of our knowledge, there are only two research articles that proposed GlycB as a novel inflammatory marker associated with insulin resistance and adiposity, even though such associations appear to be weaker than those observed for GlycA.[14,34] Our results provide novel evidence indicating that the association of glycoproteins with obesity is not restricted to GlycA but also extends to GlycB and GlycF, suggesting that all the glycoproteins studied here might be considered potential biomarkers of the chronic inflammatory process associated with obesity.

Moreover, our present results may facilitate the understanding of the contribution of obesity to the metabolic dysfunction associated

with PCOS. Our current interpretation is that PCOS associates abdominal adiposity, adipose tissue dysfunction and central (hepatic) insulin resistance per se, and that obesity aggravates these traits extending the occurrence of insulin resistance to peripheral tissues such as muscle and adipose tissue.[2,35,36] Low-grade chronic inflammation arising from the dysfunctional adipose tissue of these women is among the major pathophysiological mechanisms that mediate the effects of obesity on metabolic dysfunction[35,37] explaining the association of PCOS with proinflammatory genotypes and phenotypes already reported.[7,38]

Even considering the major impact of obesity on the glycoprotein profiles, the similar pattern found in patients with PCOS – increased H/W ratios of GlycA and GlycB, and increased GlycB and almost GlycA areas compared with control groups – might translate the low-grade chronic inflammatory process associated with this syndrome, as has been proposed for other inflammatory conditions.[14]

In the present study we did not find a specific influence of PCOS on serum hs-CRP levels, in conceptual agreement with earlier reports from our group conducted in Caucasian women from Spain,[9] even though our own meta-analysis indicated that, in general, circulating hs-CRP concentrations are increased in patients with PCOS regardless of obesity.[7] This discrepancy might be related to the fact that the women with PCOS in our studies were young and did not suffer from serious metabolic comorbidities. However, the association of PCOS with increased H/W ratios of GlycA and GlycB and GlycB area suggests the presence of a low-grade chronic inflammatory state in our

PCOS patients, further supported by the specific correlations of GlycA and GlycF areas with hs-CRP only in women with PCOS. The robustness of ^{1}H-NMR spectroscopy and the relatively small intra-individual fluctuation and inter-individual variability in glycoprotein variables may contribute to this technique outperforming conventional measurements of circulating hs-CRP as a surrogate marker of low-grade chronic inflammation.

Nevertheless, our study is not free from certain limitations. The sample size of the subgroups was relatively small; thus, our study may have been underpowered to detect small differences. Also, ^{1}H-NMR is not a selective technique and does not enable individual proteins to be identified and quantified. The areas of the studied glycoproteins correspond to the overall concentration of serum glycosylated proteins (mainly α1-acid glycoprotein, α1-antitrypsin, haptoglobin, transferrin and immunoglobulins).[16] Of note, changes in several of these circulating glycoproteins have been also associated with PCOS according to non-targeted proteomic approaches.[39] Among the strengths of our experimental design we might highlight the rather homogeneous in terms of age and percentage of obesity and mostly healthy population studied here, the quality of the analytical procedures used, and the recommendation made to all subjects of following the same diet for a few days before sampling.

In conclusion, the present study indicates that circulating ^{1}H-NMR-derived glycoprotein profiles suggestive of low-grade chronic inflammation are associated with obesity and PCOS in young adults. Whether or not these glycoprotein profiles are useful as markers of

low-grade chronic inflammation in these disorders requires further evaluation.

5.5. Supporting information

This section includes a study of the correlations between glycoproteins and phenotypic variables considering first all subjects as a whole and then, according to group of subjects.

Table C5S.1. Correlations between glycoproteins and phenotypic variables considering all subjects as a whole.

	BMI		WC		WHR		HOMA-IR		ISI		hs-CRP	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Area GlycA	0.566	<0.001	0.553	<0.001	0.315	0.022	0.471	<0.001	-0.500	<0.001	0.405	0.003
Area GlycB	0.590	<0.001	0.515	<0.001	-	-	0.449	0.001	-0.546	<0.001	0.423	0.002
Area GlycF	0.508	<0.001	0.530	<0.001	0.344	0.012	0.403	0.003	-0.398	0.003	0.372	0.006
H/W GlycA	0.486	<0.001	0.467	<0.001	-	-	0.415	0.002	-0.463	<0.001	0.447	0.001
H/W GlycB	0.489	<0.001	0.409	0.002	-	-	0.363	0.008	-0.395	0.003	0.457	0.001

Abbreviations: BMI, body mass index; H/W, height to width ratio; HOMA-IR homeostasis model assessment of insulin resistance; hs-CRP, high sensitivity C-reactive protein; ISI, composite insulin sensitivity index; WC, waist circumference; WHR, waist to hip ratio. Data were submitted to Pearson's correlation analysis. Glycoprotein areas are expressed as arbitrary units.

Table C5S.2. Correlations between glycoprotein and phenotypic variables according to group of subjects.

	ISI		**HOMA-IR**		**hs-CRP**	
Women	r	p	r	p	r	p
Area GlycA	-0.642	0.005	-	-	-	-
Area GlycB	-0.753	<0.001	-	-	-	-
Area GlycF	-0.57	0.017	0.53	0.029	-	-
H/W GlycA	-0.685	0.002	-	-	-	-
H/W GlycB	-0.531	0.028	-	-	-	-
PCOS						
Area GlycA	-	-	0.49	0.046	0.558	0.02
Area GlycB	-	-	0.523	0.031	-	-
Area GlycF	-	-	-	-	0.6	0.011
H/W GlycA	-	-	0.488	0.047	0.564	0.018
H/W GlycB	-	-	0.498	0.042	0.571	0.017
Men						
Area GlycA	-0.705	0.001	0.786	<0.001	-	-
Area GlycB	-0.72	0.001	0.641	0.003	0.487	0.035
Area GlycF	-0.511	0.025	0.644	0.003	-	-
H/W GlycA	-0.667	0.002	0.661	0.002	0.551	0.014
H/W GlycB	-0.601	0.007	0.577	0.01	0.565	0.012

Abbreviations: BMI, body mass index; HOMA-IR, homeostasis model assessment of insulin resistance; hs-CRP, high sensitivity C-reactive protein; H/W, height to width ratio. Data were submitted to Pearson's correlation analysis.

5.6. References

(1) Fernandez-Real JM, Ricart W. Insulin resistance and chronic cardiovascular inflammatory syndrome. Endocr Rev 2003;24:278-301.

(2) Escobar-Morreale HF. Polycystic ovary syndrome: definition, aetiology, diagnosis and treatment. Nat Rev Endocrinol 2018;14:270-84.

(3) Jung UJ, Choi M-S. Obesity and its metabolic complications: the role of adipokines and the relationship between obesity, inflammation, insulin resistance, dyslipidemia and nonalcoholic fatty liver disease. Int J Mol Sci 2014;15:6184-223.

(4) Dandona P, Weinstock R, Thusu K, Abdel-Rahman E, Aljada A, Wadden T. Tumor necrosis factor-alpha in sera of obese patients: fall with weight loss. J Clin Endocrinol Metab 1998;83:2907-10.

(5) Grassmann S, Wirsching J, Eichelmann F, Aleksandrova K. Association between peripheral adipokines and inflammation markers: A systematic review and meta-analysis. Obesity (Silver Spring, Md) 2017;25:1776-85.

(6) Rincon M. Interleukin-6: from an inflammatory marker to a target for inflammatory diseases. Trends Immunol 2012;33:571-7.

(7) Escobar-Morreale HF, Luque-Ramirez M, Gonzalez F. Circulating inflammatory markers in polycystic ovary syndrome: a systematic review and metaanalysis. Fertil Steril 2010;95:1048-58.e1-2.

(8) Peng Z, Sun Y, Lv X, Zhang H, Liu C, Dai S. Interleukin-6 levels in women with polycystic ovary syndrome: A systematic review and meta-analysis. PLoS ONE 2016;11:e0148531.

(9) Escobar-Morreale HF, Villuendas G, Botella-Carretero JI, Sancho J, San Millan JL. Obesity, and not insulin resistance, is the major determinant of serum inflammatory cardiovascular risk markers in pre-menopausal women. Diabetologia 2003;46:625-33.

(10) Browning LM, Krebs JD, Jebb SA. Discrimination ratio analysis of inflammatory markers: implications for the study of inflammation in chronic disease. Metabolism 2004;53:899-903.

(11) Gabay C, Kushner I. Acute-phase proteins and other systemic responses to inflammation. New Engl J Med 1999;340:448-54.

(12) Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. Cell 2006;126:855-67.

(13) Gornik O, Lauc G. Glycosylation of serum proteins in inflammatory diseases. Dis Markers 2008;25:267-78.

(14) Fuertes-Martin R, Taverner D, Vallve JC, Paredes S, Masana L, Correig Blanchar X, Amigo Grau N. Characterization of (1)H NMR plasma glycoproteins as a new strategy to identify inflammatory patterns in rheumatoid arthritis. J Proteome Res 2018;17:3730-9.

(15) Mallol R, Amigo N, Rodriguez MA, Heras M, Vinaixa M, Plana N, et al. Liposcale: a novel advanced lipoprotein test based on 2D diffusion-ordered 1H NMR spectroscopy. J Lipid Res 2015;56:737-46.

(16) Bell JD, Brown JC, Nicholson JK, Sadler PJ. Assignment of resonances for 'acute-phase' glycoproteins in high resolution proton NMR spectra of human blood plasma. FEBS Letters 1987;215:311-5.

(17) Connelly MA, Gruppen EG, Wolak-Dinsmore J, Matyus SP, Riphagen IJ, Shalaurova I, et al. GlycA, a marker of acute phase glycoproteins, and the risk of incident type 2 diabetes mellitus: PREVEND study. Clin Chim Acta 2010;452:10-7.

(18) Akinkuolie AO, Buring JE, Ridker PM, Mora S. A novel protein glycan biomarker and future cardiovascular disease events. J Am Heart Assoc 2014;3:e001221.

(19) Zawadzki JK, Dunaif A. Diagnostic criteria for polycystic ovary syndrome: Towards a rational approach. In: Dunaif A, Givens JR, Haseltine FP, Merriam GR, eds. Polycystic ovary syndrome, Vol. 4. Boston: Blackwell Scientific Publications, 1992:377-84.

(20) Ankarberg-Lindgren C, Norjavaara E. Sensitive RIA measures testosterone concentrations in prepubertal and pubertal children comparable to tandem mass spectrometry. Scand J Clin Lab Invest 2015;75:341-4.

(21) Escobar-Morreale HF, Sanchon R, San Millan JL. A prospective study of the prevalence of nonclassical congenital adrenal hyperplasia among women presenting with hyperandrogenic symptoms and signs. J Clin Endocrinol Metab 2008;93:527-33.

(22) Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412-9.

(23) Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22:1462-70.

(24) Bro R, Smilde AK. Principal component analysis. Anal Methods 2014;6:2812-31.

(25) Ballabio D, Consonni V. Classification tools in chemistry. Part 1: linear models. PLS-DA. Anal Methods 2013;5:3790-8.

(26) Manmadhan A, Lin BX, Zhong J, Parikh M, Berger JS, Fisher EA, Heffron SP. Elevated GlycA in severe obesity is normalized by bariatric surgery. Diabetes Obes Metab 2018;21:178-82.

(27) Houttu N, Mokkala K, Laitinen K. Overweight and obesity status in pregnant women are related to intestinal microbiota and serum metabolic and inflammatory profiles. Clin Nutr 2018;37:1955-66.

(28) Jago R, Drews KL, Otvos JD, Willi SM, Buse JB. Novel measures of inflammation and insulin resistance are related to

obesity and fitness in a diverse sample of 11-14 year olds: The HEALTHY Study. Int J Obes (2005) 2016;40:1157-63.

(29) Olson ML, Renteria-Mexia A, Connelly MA, Vega-Lopez S, Soltero EG, Konopken YP, et al. Decreased GlycA after lifestyle intervention among obese, prediabetic adolescent Latinos. J Clin Lipidol 2018;13:186-93.

(30) Bartlett DB, Slentz CA, Connelly MA, Piner LW, Willis LH, Bateman LA, et al. Association of the composite inflammatory biomarker GlycA, with exercise-induced changes in body habitus in men and women with prediabetes. Oxid Med Cell Longev 2017;2017:5608287.

(31) Dullaart RP, Gruppen EG, Connelly MA, Lefrandt JD. A pro-inflammatory glycoprotein biomarker is associated with lower bilirubin in metabolic syndrome. Clin Biochem 2002;48:1045-7.

(32) Gruppen EG, Connelly MA, Otvos JD, Bakker SJ, Dullaart RP. A novel protein glycan biomarker and LCAT activity in metabolic syndrome. Eur J Clin Invest 2015;45:850-9.

(33) Bogl LH, Kaye SM, Ramo JT, Kangas AJ, Soininen P, Hakkarainen A, et al. Abdominal obesity and circulating metabolites: A twin study approach. Metabolism 2016;65:111-21.

(34) Lorenzo C, Festa A, Hanley AJ, Rewers MJ, Escalante A, Haffner SM. Novel protein glycan-derived markers of systemic inflammation and C-reactive protein in relation to glycemia, insulin resistance, and insulin secretion. Diabetes care 2017;40:375-82.

(35) Escobar-Morreale HF, San Millan JL. Abdominal adiposity and the polycystic ovary syndrome. Trends Endocrinol Metab 2007;18:266-72.

(36) Escobar-Morreale HF, Samino S, Insenser M, Vinaixa M, Luque-Ramirez M, Lasuncion MA, Correig X. Metabolic heterogeneity in polycystic ovary syndrome is determined by obesity: plasma metabolomic approach using GC-MS. Clin Chem 2012;58:999-1009.

(37) Escobar-Morreale HF, Alvarez-Blasco F, Botella-Carretero JI, Luque-Ramirez M. The striking similarities in the metabolic associations of female androgen excess and male androgen deficiency. Hum Reprod 2014;29:2083-91.

(38) Escobar-Morreale HF, Luque-Ramirez M, San Millan JL. The molecular-genetic basis of functional hyperandrogenism and the polycystic ovary syndrome. Endocr Rev 2005;26:251-82.

(39) Insenser M, Martinez-Garcia MA, Montes R, San-Millan JL, Escobar-Morreale HF. Proteomic analysis of plasma in the polycystic ovary syndrome identifies novel markers involved in iron metabolism, acute-phase response, and inflammation. J Clin Endocrinol Metab 2010;95:3863-70.

CHAPTER 5

CHAPTER 6

6. Glycoprotein profile measured by ^{1}H-Nuclear Magnetic Resonance in patients with diabetes

6.1. Abstract

Patients with type 2 diabetes mellitus (T2DM) and atherogenic dyslipidaemia (AD) are at higher risk of developing cardiovascular diseases (CVDs) so interest in discovering inflammation biomarkers as indicators of processes related to CVD progression is increasing. This study has two aims: to characterize the plasma glycoprotein profile of a cohort of 504 participants including patients with and without T2DM and/or AD and controls by using ^{1}H Nuclear Magnetic Resonance Spectroscopy (^{1}H-NMR); and to study the associations between the glycoprotein profile and other lipid and clinical variables in these populations.

We characterized plasma glycoprotein profiles of T2DM patients with and without AD by using ^{1}H-NMR. We quantified the two peaks associated with the concentration of plasma glycoproteins (GlycA and GlycB) and their height/width ratios (H/W GlycA and H/W GlycB), as higher and narrower signals have been related to inflammation. We also quantified GlycF, the signal of which is proportional to the concentration of the acetyl groups of free N-acetylglucosamine, N-acetylgalactosamine and N-acetylneuraminic in the samples. The lipoprotein profile was also determined by ^{1}H-NMR spectroscopy (Liposcale®). Standard clinical and anthropometric measurements were taken. Several PLSDA models were developed to study differences between the study groups.

A pro-atherogenic lipoprotein pattern that included reduced HDL-C levels, increased small dense LDL and HDL particles, and

elevated TG levels was significantly associated with glycoprotein variables. Along the same lines, glycoprotein values in the diagnostic groups were significantly different from those in the CT groups. AD and DM conditions together contribute to a positive and significant synergetic effect on the GlycA area (<0.05) and the H/W ratios of GlycA (<0.01) and GlycB (<0.05). By adding the new glycoprotein variables to the traditionally used marker of inflammation C-reactive protein (CRP), the AUC increased sharply for classification models between the CT group and the rest (0.68 to 0.84), patients with and without dyslipidaemia (0.54 to 0.86), and between patients with and without diabetes (0.55 to 0.75).

^{1}H-NMR-derived glycoproteins can be used as possible markers of the degree of inflammation associated with diabetes and dyslipidemia. More studies are required to evaluate in detail the mechanism underlying glycoprotein variation in AD and DM.

6.2. Introduction

Type 2 diabetes (T2DM) is one of the most common diseases and has been growing exponentially in recent years.[1] Patients with diabetes are known to have a 2 to 4 times greater risk of cardiovascular diseases (CVD) such as stroke and death from coronary heart disease than non-diabetic subjects.[2] Most patients with diabetes also have dyslipidemia, which increases the risk of CVD.[3] The most common pattern of dyslipidemia in T2DM patients is atherogenic dyslipidemia (AD) which is characterized by elevated total triglycerides (TG) and

small LDL particles, and decreased HDL cholesterol levels[4,5] with insulin resistance (IR) being the main pathophysiological machanism.[3] On the other hand, T2DM is considered a chronic subclinical inflammatory state which also contributes to vascular complications.[6,7]

C-reactive protein (CRP) may help to refine global risk assessment for coronary heart disease (CHD), particularly among persons who are at intermediate risk on the basis of traditional risk factors alone. However, it has also been shown to be prone to fluctuations and an inconsistent marker for predicting CVD in patients with chronic inflammatory diseases at an individual level.[8] For some years now, there has been increasing interest in inflammation biomarkers, which play a major role in the onset and progression of CVD.[9] Several studies have shown that glycoproteins play a key role in inflammatory and pathological processes such as the genesis and/or progression of CVDs and are also related to numerous pathologies, in particular cardiovascular risk factors such as diabetes mellitus.[10–13] It has been reported that the signal in the ^{1}H-NMR spectrum of glycated proteins is produced by the -COCH3 acetyl groups of N-acetylglucosamine and N-acetylgalactosamine (GlycA) and N-acetylneuraminic acid (GlycB).[14,15] Unlike common biomarkers of inflammation such as CRP or inflammatory cytokines, GlycA is a composite biomarker that integrates the protein levels and glycation states of several of the most abundant acute-phase proteins in serum (α1-acid glycoprotein, haptoglobin, α1-antitrypsin, α1-antichymotrypsin, and transferrin).[14,16] This provides a more stable measure of systemic inflammation with less intra-individual variability

for GlycA than for high sensitivity CRP.[14] Although GlycA, and to a lesser extent GlycB, is the variable that has most been studied in terms of glycoproteins determined by ^{1}H-NMR, other new variables have recently been reported such as the derived parameters H/W GlycA and H/W GlycB, which are related to the shape of the signal (higher and narrower signals have been related to inflammation).[15]

In this study we characterized the plasma ^{1}H-NMR glycoprotein profile of T2DM patients with and without AD, we explored its association with their NMR lipoprotein profile, and finally we compared the ability of these new NMR based on inflammatory markers to discriminate specific patterns between study groups other than those commonly used in clinical settings such as CRP.

6.3. Methodology

6.3.1. Design and study subject

For this cross-sectional study we recruited 504 individuals attending the vascular medicine and metabolism unit of the University Hospital Sant Joan de Reus for disorders associated with the lipid metabolism. Type 2 diabetes was diagnosed using standard clinical criteria. Atherogenic dyslipidemia was defined as TG > 150 mg/dl and HDL-c < 40 or 50 mg/dl for males and females, respectively. Subjects with chronic lung, renal or liver disease, cancer, or any other serious disease were excluded. Patients on lipid-lowering drugs underwent a 6-week wash-out period (8 weeks if they were on fibrates). Anamnesis, anthropometric, and physical examination data were recorded.

The entire population was classified into four groups depending on whether they had T2DM (DM) and/or AD; diabetes and dyslipidaemia [DM (+) AD (+), n=129], AD [DM (-) AD (+), n=38], DM [DM (+) AD (-), n=222]. A total of 115 patients had neither T2DM nor AD and they were referred to as the control population [CT, n=115]. The study was approved by the Ethical and Clinical Investigation Committee of the Pere Virgili Institute for Health Research (IISPV), and all participants signed the written consent form.

6.3.2. Clinical and standard biochemical variables

Participants had their clinical history and their anthropometric data recorded and were subject to a physical examination. A blood sample was obtained in a fasting state and deeply frozen plasma and serum aliquots were stored in our research institute's biobank until use. Standard biochemical determinations were performed using standard biochemical methods. If the patients were on lipid lowering therapy, the blood sample was obtained after a six-week washout period.

6.3.3. Glycoprotein and lipoprotein profiling by ^{1}H-NMR

For glycoprotein profiling, plasma samples were analysed by ^{1}H-NMR. Serum samples (200 µL) were previously diluted with 50 µl deuterated water and 300 µl of 50 mM phosphate buffer solution (PBS) at pH 7.4 consisting of 30.70 Na_2HPO_4 mM and 19.30 NaH_2PO_4 mM before NMR analysis. ^{1}H-NMR spectra were recorded at 310 K on a Bruker Avance III 600 spectrometer operating at a proton frequency of 600.20 MHz (14.1 T). One-dimensional ^{1}H NMR pulse experiments were carried out including nuclear Overhauser effect spectroscopy

(NOESY) and longitudinal eddy-current delay (LED) pulses. To measure the glycoprotein variables we analysed the glycoprotein region of the ^{1}H-NMR spectra which is between 2.19 and 1.90 ppm and represented the resonance of the N-acetyl methyl groups of the N-acetylglucosamine, N-acetylgalactosamine and N-acetylneuraminic acid moieties on the carbohydrate portions of circulating glycoproteins such as α1-acid glycoprotein, haptoglobin, α1-antitrypsin, α1-antichymotrypsin, and transferrin.[14–16] We measured the previously reported[15] glycoprotein variables including areas of GlycB (Area GlycB), GlycA (Area GlycA) and height/width ratios of GlycB and GlycA. The areas of GlycA and GlycB provide the acetyl group concentrations of protein-bound N-acetylglucosamine, N-acetylgalactosamine and N-acetylneuraminic acid in plasma. In each case the form of the function (H/W ratio) depends on the height, which is related to the concentration, and the width, which is related to the flexibility and the aggregation of the molecules generating the signal. This study describes a complementary new analytical function called GlycF, the area of which reflects the acetyl groups of free N-acetylglucosamine, N-acetylgalactosamine and N-acetylneuraminic acid in the sample (that's to say, they do not bind to proteins).

We also obtained the lipoprotein profile using the NMR test Liposcale® (CE), as previously reported.[17] This protocol evaluates the lipid concentrations (i.e., triglycerides and cholesterol), size and particle number of three different classes of lipoproteins (VLDL, LDL and HDL), as well as the particle number of nine subclasses (large, medium and small VLDL, LDL and HDL).

6.3.4. Statistical methods

To study the distribution of the clinical variables of the study groups, we used the one-factor Anova with Bonferroni post-hoc correction or the $\chi 2$ test depending on whether the variable was continuous or categorical. Spearman correlation coefficients were obtained between the glycoprotein variables, the clinical variables and the lipoprotein-related variables. A logarithmic transformation and an outlier's test were performed for all those dependent variables that, according to the Lilliefors test, did not follow a normal distribution. A Wilcoxon-Mann-Whitney test was used to capture the differences in the glycoprotein profile and CRP of each of the four groups. Two-way Anova was performed to examine the influence of the AD and DM conditions of each glycoprotein variable.

Multivariate data analysis

We used linear regression models adjusted for age, gender and BMI, DM, HDL-C, LDL-C and TG. Partial least squares discriminant analysis (PLSDA) models were used as a supervised classification method between study groups. PLS-DA relates the $\boldsymbol{X}$ matrix (experimental data) and the $\boldsymbol{Y}$ matrix (classes of samples) to find the maximum discrimination between classes (groups of study) and the maximum covariance between the $\boldsymbol{X}$ and $\boldsymbol{Y}$ matrices simultaneously.[18] The area under the curve (AUC) was used to evaluate the capacity of the glycoprotein and CRP variables to distinguish between the two groups. All the models were auto scaled and cross-validated by the Venetian Blinds cross validation method.

IBM SPSS® statistics ver. 22 was used to study the clinical statistical distribution between groups and the adjustment of the linear regression models. Multivariate data analyses were computed in MATLAB, ver. 7.10.0 using PLS-Toolbox, ver. 5.2.2 (Eigenvector Research).

6.4. Results

Table C6.1 shows the demographic, clinical and biochemical characteristics of the study population. Significant differences were observed between groups in terms of age, and between CT and the other groups in terms of body mass index (BMI), CRP and Hb1AC. Lipid variables were generally higher in the study groups than in the CT group, except for HDL-C, which was higher in the CT group.

6.4.1. Associations of ^{1}H-NMR-derived glycoprotein variables with the clinical and the Liposcale® test variables

Table C6.2 shows the relationship between the five ^{1}H-NMR glycoprotein variables and continuous clinical variables such as BMI, blood glucose, PCR, HbA1c and Liposcale®. A similar pattern was observed in all five variables. We found no significant relationship between age and the concentration of glycated proteins in plasma. However, there was a positive and significant association between the concentration of glycated proteins and BMI, CRP and blood glucose. For HbA1c, the trend was similar but only significant with H/W GlycA ($p<0.05$) and H/W GlycB ($p<0.01$) ratios. As far as the relationship with lipidic Liposcale® variables is concerned, we found that total

abundance of GlycA, GlycB and GlycF and shape of GlycA and GlycB were clearly associated with a proatherogenic pattern, which included a negative and significant association of the five glycoprotein parameters and HDL cholesterol (HDL-C) levels. The same applied to medium HDL particle concentration and to the size of VLDL, HDL and LDL particles.

The contribution of the clinical variables to each glycoprotein variable was studied using three linear regression models with the following equations:

i) $$\boldsymbol{X} = \beta_0 + \beta_1 Age + \beta_2 Gender + \beta_3 BMI$$

ii) $$\boldsymbol{X} = \beta_0 + \beta_1 Age + \beta_2 Gender + \beta_3 BMI + \beta_4 Diabetes$$

iii) $$\boldsymbol{X} = \beta_0 + \beta_1 Age + \beta_2 Gender + \beta_3 BMI + \beta_4 Diabetes + \beta_5 HDL\_C + \beta_6 LDL\_C + \beta_7 TG$$

The independent variables considered were age, gender and BMI for model A (**i**); age, gender, BMI and diabetes for model B (**ii**); and age, gender, BMI, diabetes LDL-C, HDL-C and total triglycerides (TG) for model C (**iii**). The dependent variable (X) was each of the 5 ^{1}H-NMR glycoprotein variables for each of the three models. The beta coefficients (β) determine the explanatory variable that has most weight in the explanation of each glycoprotein variable (see **table C6.3**).

As can be seen in **table C6.3**, some of the independent variables considered in these models significantly affected each glycoprotein

variable. In model A, all the glycoprotein variables were significantly affected by BMI. Moreover, both the area and H/W ratio of GlycB were affected by age. In model B, all the glycoprotein variables were equally significantly and positively affected by BMI and diabetes.

Finally, in model C, after correcting for lipid variables, we observed that all glycoprotein variables were significantly and negatively affected by HDL-C, and significantly and positively by TG. However, some differences were observed between the glycoproteins in this model; the area and H/W ratio of GlycB was also significantly affected by gender, BMI and diabetes; the area and H/W ratio of GlycF was affected by gender, diabetes, LDL-C and TG; the area of GlycA was only significantly affected by the three lipid variables and diabetes and the H/W ratio of GlycA was affected by gender and BMI.

Table C6.1: Distribution and clinical characteristic of groups.

	AD (+) DM (+) n=129	DM (-) AD (+) n= 38	DM (+) AD (-) n=222	CT n= 115	*p*
Age (years)	58 ± 14	54 ± 15	65 ± 10.75	58 ± 13.50	0.00** (A-B, A-D, B-C, C-D)
Gender (% male)	50	33.3	49.5	61.4	0.02*
BMI (kg/m)	31.85 ± 5.87	31.76 ± 2.86	30.20 ± 6.22	26.90 ± 4.49	0.00 ** (A-D, B-D, C-D)
Smoking (%)	25.3	25	18	4.7	0.00**
HBP (%)	70.1	47.2	74.5	16.1	0.00**
Total-C	215.37 ± 70.81	223.84 ± 47.54	197.19 ± 66.92	206.38 ± 43	0.00**(A-C, B-C, C-D)
LDL-C	117.89 ± 59.25	125.14 ± 46.16	111.57 ± 46.27	107.79 ± 46.14	>0.05
HDL-C	41.04 ± 15.75	40.63 ± 16.53	51.54 ± 18.84	56.16 ± 22.93	0.00**(A-C, A-D, B-C, B-D)
TG	215.56 ±127.62	217.10 ± 134.97	111.35 ± 57	71.48 ± 42.84	0.00**(A-C, A-D, B-C)
CRP (mg/L)	2.89 ± 2.15	2.09 ± 2.51	2.06 ± 1.64	2.25 ± 0.90	0.00 **(A-D, C-D)
Glucose	158.50 ± 57	114 ± 22	151.50 ± 62.25	101.50 ± 18	0.00**(A-B, A-D, B-C, C-D)
Hb1AC	6.70 ± 1.50	5.20 ± 0.60	6.90 ± 1.67	2.97 ± 2.80	0.00** (A-B, A-D, B-C, B-D, C-D)

Abbreviations: BMI, body mass index; AHT, arterial hypertension; CVA, cerebrovascular accident; CRP, C-reactive protein; Hb1AC, glycated hemoglobin; A, AD(+)DM(+); B, AD(+)DM(-); C, AD(-)DM(+); D, CT. A) DM (+) DA (+), B) DM (-) AD (+), C) DM (+) AD (-), D) CT. The median and the interquartile range are indicated for continuous variables and the distribution percentage is indicated for categorical variables. *P*-values are corrected by the Bonferroni method. For categorical variables the chi-squared *p*-value is shown. * is marked for $p<0.05$ and ** for $p<0.01$.

Table C6.2: Relationship between the five ^{1}H-NMR glycoprotein variables and continuous clinical variables

	Area GlycB		Area GlycF		Area GlycA		H/W GlycB		H/W GlycA	
	r	*p-value*	r	*p-value*	r	*p-value*	r	*p-value*	r	*p-value*
Clinical variables										
Age	-0.03	0.46	-0.01	0.84	-0.05	0.31	0.02	0.75	-0.02	0.66
BMI	0.30	0.00**	0.30	0.00**	0.34	0.00**	0.29	0.00**	0.32	0.00**
Glucose_cobas	0.38	0.00**	0.44	0.00**	0.42	0.00**	0.39	0.00**	0.44	0.00**
PCR	0.37	0.00**	0.32	0.00**	0.31	0.00**	0.42	0.00**	0.39	0.00**
HbA1c	0.10	0.17	0.15	0.05	0.09	0.25	0.28	0.00**	0.16	0.04*
Liposcale ® variables										
VLDL-C	0.61	0.00**	0.76	0.00**	0.88	0.00**	0.42	0.00**	0.74	0.00**
IDL-C	0.49	0.00**	0.48	0.00**	0.55	0.00**	0.31	0.00**	0.49	0.00**
LDL-C	0.04	0.30	0.08	0.04*	0.13	0.00**	-0.02	0.66	0.10	0.02*
HDL-C	-0.40	0.00**	-0.47	0.00**	-0.56	0.00**	-0.34	0.00**	-0.49	0.00**
VLDL-TG	0.60	0.00**	0.79	0.00**	0.89	0.00**	0.44	0.00**	0.75	0.00**
IDL-TG	0.59	0.00**	0.61	0.00**	0.70	0.00**	0.40	0.00**	0.63	0.00**
LDL-TG	0.41	0.00**	0.35	0.00**	0.42	0.00**	0.25	0.00**	0.39	0.00**
HDL-TG	0.42	0.00**	0.52	0.00**	0.58	0.00**	0.24	0.00**	0.48	0.00**
VLDL-P (nmol/L)	0.60	0.00**	0.79	0.00**	0.89	0.00**	0.44	0.00**	0.75	0.00**
Large VLDL-P (nmol/L)	0.59	0.00**	0.76	0.00**	0.88	0.00**	0.44	0.00**	0.75	0.00**
Medium VLDL-P (nmol/L)	0.61	0.00**	0.78	0.00**	0.90	0.00**	0.44	0.00**	0.75	0.00**
Small VLDL-P (nmol/L)	0.60	0.00**	0.79	0.00**	0.89	0.00**	0.44	0.00**	0.74	0.00**
LDL-P (nmol/L)	0.14	0.00**	0.19	0.00**	0.23	0.00**	0.06	0.16	0.20	0.00**
Large LDL-P (nmol/L)	0.03	0.53	-0.01	0.89	0.06	0.17	-0.06	0.12	0.03	0.54
Medium LDL-P (nmol/L)	0.01	0.88	0.03	0.47	0.07	0.10	-0.03	0.50	0.06	0.15
Small LDL-P (nmol/L)	0.24	0.00**	0.31	0.00**	0.35	0.00**	0.14	0.00**	0.31	0.00**
HDL-P (µmol/L)	-0.09	0.03*	-0.06	0.17	-0.09	0.03*	-0.14	0.00**	-0.10	0.02*
Large HDL-P (µmol/L)	0.19	0.00**	0.17	0.00**	0.27	0.00**	0.06	0.16	0.17	0.00**
Medium HDL-P (µmol/L)	-0.36	0.00**	-0.46	0.00**	-0.55	0.00**	-0.31	0.00**	-0.48	0.00**
Small HDL-P (µmol/L)	0.10	0.01*	0.21	0.00**	0.23	0.00**	0.00	0.96	0.17	0.00**
VLDL-Z (nm)	-0.27	0.00**	-0.44	0.00**	-0.40	0.00**	-0.21	0.00**	-0.33	0.00**
LDL-Z (nm)	-0.38	0.00**	-0.53	0.00**	-0.51	0.00**	-0.35	0.00**	-0.47	0.00**
HDL-Z (nm)	-0.45	0.00**	-0.62	0.00**	-0.74	0.00**	-0.34	0.00**	-0.61	0.00**
Non-HDL-P (nmol/L)	0.29	0.00**	0.37	0.00**	0.43	0.00**	0.17	0.00**	0.38	0.00**
Total-P/HDL-P	0.28	0.00**	0.32	0.00**	0.39	0.00**	0.21	0.00**	0.35	0.00**
LDL-P/HDL-P	0.17	0.00**	0.19	0.00**	0.25	0.00**	0.13	0.00**	0.23	0.00**

Spearman correlation coefficients (r) and Spearman p-values (p) are shown. Significance is marked (* for p<0.05 and ** for p<0.01). Green indicates positive associations, while orange indicates opposite associations. For Age, IMC and Glucose_cobas n=455; for CRP n=396; for HbA1c n=172; and for the other variables n=504.

6.4.2 Analysis of glycoproteins and CRP in the study population groups

Significant differences were detected in glycoprotein variables and CRP between the study groups (see **Figure C6.1**). The first thing to note is the difference in the dispersion of the variables. While the glycoprotein variables had a narrow distribution, CRP clearly had a wider dispersion. In all the variables there were significant differences between the control group and the others. For Area GlycB and CRP, there were only differences between the three diagnostic groups and CT. Area GlycF shows the same differences as Area GlycB but there were also differences between [DM (-) AD (+)] and [DM (+) AD (-)]. Area GlycA and the ratios H/W GlycB and H/W GlycA show significant differences between all the groups. No variable was significantly different between the groups [DM (+) AD (+)] and [DM (-) AD (+)].

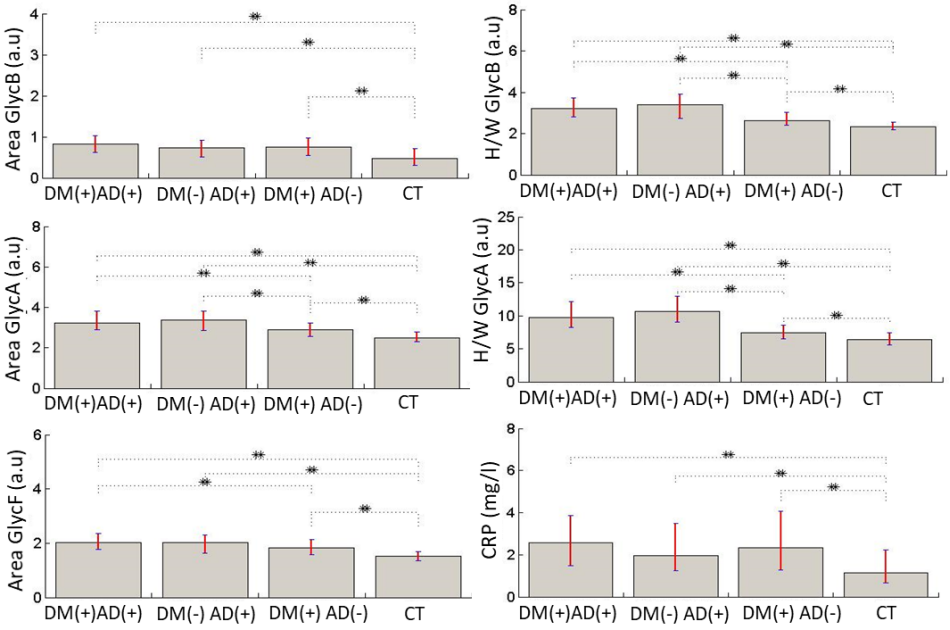


Figure C6.1: Univariant analysis of ^{1}H-NMR glycoprotein variables and CRP between groups. Median and interquartile range (Iqr) are represented. Iqr is indicated in red. The Wilcoxon−Mann−Whitney test has been used to calculate significance (*p<0.05, ** p<0.01).

Neither the GlycB area nor CRP were able to discriminate between the condition of having diabetes or dyslipidaemia. GlycF distinguished between having dyslipidaemia or not in a population with diabetes. The other variables were also capable of detecting significant differences between the group that only has diabetes and the group that only has AD. These data indicate that the dominant condition is AD, and that the glycoprotein variables are much more affected when there is AD than diabetes.

To examine the influence of AD and DM on each glycoprotein variable and CRP, we analysed the main effect of each independent condition and its interaction. **Table C6.4** shows the two-way Anova results for AD, DM and the interaction between them for each variable. AD was significantly influenced by increasing the value of all the glycoprotein and CRP variables. DM significantly affected the value of Area GlycB, H/W GlycB, H/W GlycA and CRP. However, the interaction between both effects (AD and DM) was only significant for Area GlycA ($p<0.05$), H/W GlycA ($p<0.01$) and H/W GlycB($p<0.05$), which showed that there was a synergistic effect between the two conditions.

Table C6.3: Covariance-adjusted linear regression models

			DEPENDENT VARIABLE									
			Area GlycB		Area GlycF		Area GlycA		H/W GlycB		H/W GlycA	
			β	*p*	β	*P*	β	*p*	β	*p*	β	*p*
INDEPENDENT VARIABLE	**Age**	*Model A*	0.09	0.03*	0.06	0.15	0.05	0.22	0.09	0.04*	0.07	0.10
		Model B	0.02	0.58	-0.02	0.60	0.00	0.96	-0.03	0.46	-0.02	0.65
		Model C	0.02	0.61	-0.04	0.15	-0.02	0.37	-0.02	0.55	-0.03	0.32
	Gender	*Model A*	-0.02	0.64	0.00	0.86	0.02	0.56	-0.02	0.57	0.00	0.92
		Model B	-0.04	0.33	-0.03	0.41	0.00	0.87	-0.06	0.12	-0.03	0.43
		Model C	-0.08	0.04*	-0.07	0.02*	-0.04	0.05	-0.10	0.01*	-0.07	0.02*
	BMI	*Model A*	0.34	0.00**	0.28	0.00**	0.32	0.00**	0.34	0.00**	0.35	0.00**
		Model B	0.26	0.00**	0.19	0.00**	0.26	0.00**	0.21	0.00**	0.26	0.00**
		Model C	0.14	0.00**	0.01	0.71	0.04	0.08	0.13	0.00**	0.08	0.02*
	Diabetes	*Model B*	0.20	0.00**	0.27	0.00**	0.17	0.00**	0.37	0.00**	0.28	0.00**
		Model C	0.16	0.00**	0.21	0.00**	0.10	0.00**	0.35	0.00**	0.23	0.00**
	HDL-C	*Model C*	-0.15	0.00**	-0.06	0.11	-0.14	0.00**	-0.17	0.00**	-0.13	0.00**
	LDL-C	*Model C*	0.06	0.10	0.09	0.00**	0.10	0.00**	0.04	0.33	0.13	0.00**
	TG	*Model C*	0.35	0.00**	0.67	0.00**	0.75	0.00**	0.22	0.01*	0.5	0.00**

The beta coefficients (β) and Anova p-value (*p*) are represented. Significance is marked (* for p<0.05 and ** for p<0.01). Abbreviations: BMI, body mass index; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; TG, total triglycerides; AHT, arterial hypertension; CVA, cerebral vascular accident.

Table C6.4: Two-way analysis of variance

	(AD+) vs (AD-)	(DM+) vs (DM-)	(AD+) vs (DM+)
Area GlycB	0.00**	0.01*	0.59
Area GlycF	0.00**	0.68	0.17
Area GlycA	0.00**	0.30	0.02*
H/W GlycB	0.00**	0.00*	0.02*
H/W GlycA	0.00**	0.00**	0.00**
CRP	0.02*	0.00**	0.27

p-values for a balanced two-way ANOVA. Significance is marked (* for $p<0.05$ and ** for $p<0.01$).

To refine the separation between groups, three different PLSDA models were built (see **figure C6.2**) with the five glycoprotein variables and CRP for matrix ***X***. Matrix ***Y*** was a vector of two classes that includes the whole population. Model 1 shows CT versus the rest of the study population; Model 2 shows individuals with DA versus individuals without DM; Model 3 classifies individuals with or without DM. For each of these models, six ROC curves were evaluated: one for CRP alone and the other five for CRP with one of the glycoprotein variables added one at a time. The biplot in **Figure C6.2A** shows a clear separation between CT and the rest of the study population. **Figure C6.2B** shows a good separation between patients with and without dyslipidaemia if glycoprotein variables were considered, and a less clear separation if only CRP was considered. **Figure C6.2C** shows a weaker separation between the patients with and without diabetes.

In all cases, if the ^{1}H-NMR glycoprotein variables were taken into account the separation between groups greatly improved. The AUC of the CRP ROC curve in the three models was much smaller

than the AUC when the glycoprotein variables were added (0.68 vs 0.84 for model 1; 0.54 vs 0.86 for model 2; and 0.55 vs 0.75 for model 3).

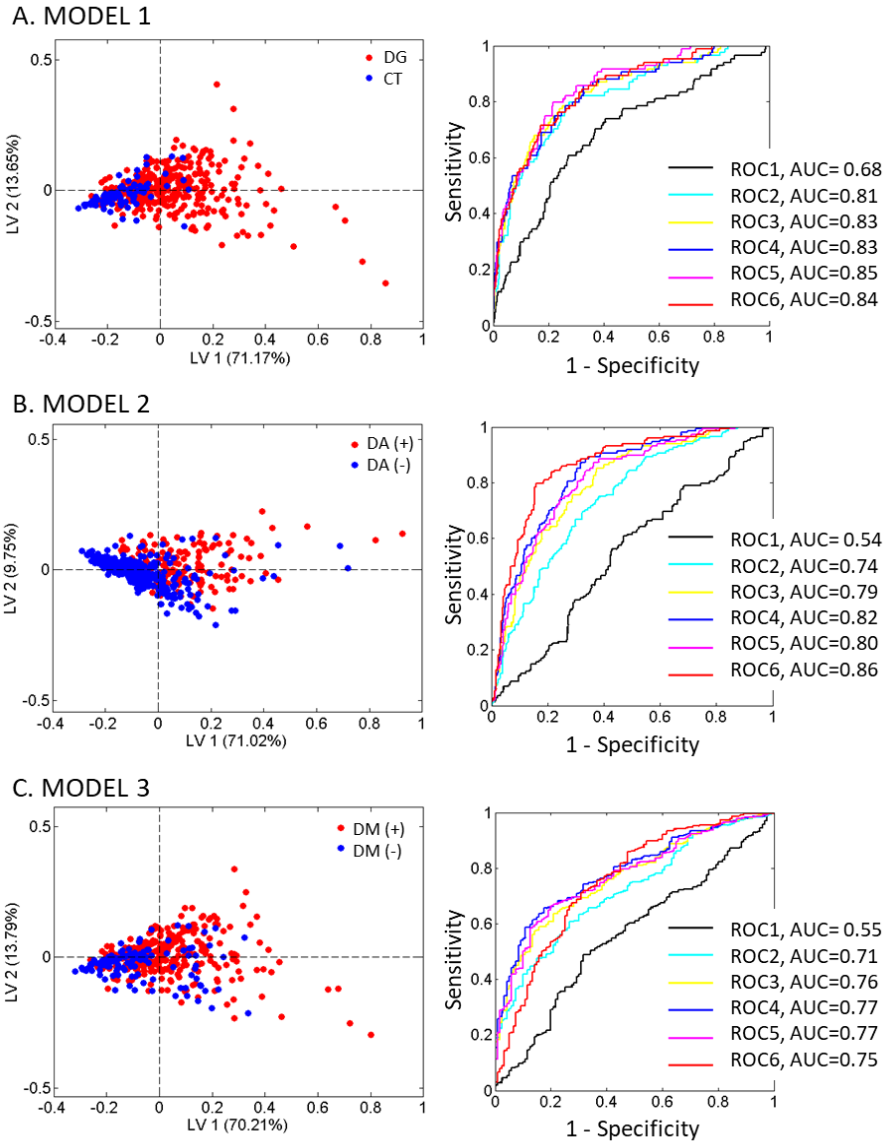


Figure C6.2. *Abbreviations:* DG, diagnostic groups; CT, control groups; AD, atherogenic dyslipidaemia; DM, diabetes mellitus. Biplots and ROC curves of PLSDA models for matrix X with the 5 ^{1}H-NMR

glycoprotein variables and CRP. [ROC1: CRP; ROC2: CRP + Area GlycB; ROC3: CRP + Area GlycB + Area GlycA; ROC4: CRP + Area GlycB + Area GlycF + Area GlycA; ROC5: CRP + Area GlycB + Area GlycF + Area GlycA + H/W GlycB; ROC6: CRP + Area GlycB + Area GlycF + Area GlycA + H/W GlycB + H/W GlycA].

6.5. Discussion

Diabetes is a chronic disease associated with circulating levels of inflammatory proteins. The enhanced synthesis of proinflammatory cytokines and acute-phase proteins characterizes the early stages of T2DM and increases as the disease progresses.[6] Many studies have suggested that the intra-individual reliability and variability of CRP is an inflammatory marker in the prediction and follow-up of some diseases.[19,20] In this study, we have characterized the ^{1}H-NMR plasma glycoproteins of T2DM patients with and without AD, we have explored their relation with lipoproteins and clinical variables, and we have evaluated them as new emerging inflammatory markers of disease in comparison to CRP.

Our study showed that the glycoprotein profile was positively associated with being female, BMI, blood glucose, CRP and a proatherogenic lipoprotein pattern with increased levels of LDL-C and TG. These results are in line with those of some studies in which GlycA has been associated with CRP and a high BMI.[21–25] The association with HbA1c, however, was not so clear. There appeared to be a significant association only with the ratios H/W of GlycA and GlycB. However, HbA1c was significantly higher in the three study groups

than in the control group and even higher in the DM+ groups. These results suggest that the inflammatory pathways for HbA1c and serum/plasma glycoproteins are different and determined by ^{1}H-NMR. More studies are needed to establish the bases of the inflammatory pathways of these blood proteins.

The significant negative associations found between the five glycoprotein variables and HDL-C and the size of VLDL, HDL and LDL particles and the high correlation with TG are in line with the literature, where diabetes is related to reduced HDL-C levels, a predominance of small dense LDL particles, and elevated TG levels.[28] In addition, this study has found a significant positive association between glycoproteins and small HDL, the most abundant subclass in patients with T2DM.[29] These results suggest that an increase in glycosylated proteins is consistent with the characteristic lipid pattern of both T2DM and AD. In addition, linear regression models showed that both the area and H/W ratio of GlycB was largely associated with high TG concentrations and low HDL-C concentrations, while the area and H/W ratio of GlycA were also associated with high concentrations of LDL-C.

The significant differences in the glycoprotein values of each of the diagnostic groups versus the CT group were expected. However, the CRP showed greater dispersion among individuals from each group than the ^{1}H-NMR glycoprotein variables, which confirms the higher intra-variability mentioned in the literature.[19,20] It appears that AD contributes to an increase in the five glycoprotein variables, while DM contributes to an increase in the GlycB area and the two variable H/W

ratios, but not in the GlycA area. The GlycA area was associated with AD but not with DM. The two conditions together contribute to a positive and significant synergic effect on the GlycA area and the H/W ratios of GlycA and GlycB.

PLSDA models suggest that using the new glycoprotein variables in conjunction with common inflammation markers such as CRP maximizes the sensitivity and specificity of classifying DM and AD, pathological states linked with inflammatory processes. A set of inflammatory markers could be used to develop effective and novel strategies for disease intervention.

These results are in line with those of several studies in the literature which report an increased concentration of GlycA in patients with T2DM.[23,30–33] The mechanisms by which glycoproteins parameters are increased in diabetes are not well understood. α1-glycoprotein, one of the circulating glycoproteins detected by ^{1}H-NMR, is known to be a marker of glycemia because prolonged inflammation decreases glucose tolerance.[30] The close relationship of insulin sensitivity with GlycA, but not with GlycB, has also been described.[34]

Nevertheless, our study is not free from certain limitations. The sample size and the distribution of some variables are not homogeneous between groups. However, the models have been corrected in an attempt to minimize this variability. In addition, we have not been able to provide data on other clinical complications of diabetes, because this was beyond the scope of this article. Among the

strengths of this study are the quality and novelty of the analytical procedures used. And from a single experiment with ^{1}H-NMR we can obtain both the glycoprotein variables and the advanced lipid profile, which leads to lower cost and less experimental variability.

In conclusion, the present study suggests that ^{1}H-NMR glycoproteins could be used as markers of the degree of inflammation associated with diabetes and dyslipidaemia, as they have been shown to have lower inter-individual variability and to improve predictive outcomes of CRP in patients with and without diabetes and in patients with and without dyslipidaemia. Future studies will be necessary to understand the mechanism underlying inflammation in AD and diabetes.

6.6. References

(1) Maffi, P.; Secchi, A. The Burden of Diabetes: Emerging Data. In *Developments in ophthalmology*; 2017; Vol. 60, pp 1–5.

(2) Newman, J. D.; Schwartzbard, A. Z.; Weintraub, H. S.; Goldberg, I. J.; Berger, J. S. Primary Prevention of Cardiovascular Disease in Diabetes Mellitus. *J. Am. Coll. Cardiol.* **2017**, *70* (7), 883–893.

(3) Manoria, P. C.; Chopra, H. K.; Parashar, S. K.; Dutta, A. L.; Pinto, B.; Mullasari, A.; Prajapati, S. The Nuances of Atherogenic Dyslipidemia in Diabetes: Focus on Triglycerides and Current Management Strategies. *Indian Heart J.* **2013**, *65* (6), 683–690.

(4) Haffner, S. M.; American Diabetes Association. Management of

Dyslipidemia in Adults with Diabetes. *Diabetes Care* **2003**, *26 Suppl 1* (suppl 1), S83-6.

(5) Mooradian, A. D. Dyslipidemia in Type 2 Diabetes Mellitus. *Nat. Rev. Endocrinol.* **2009**, *5* (3), 150–159.

(6) Badawi, A.; Klip, A.; Haddad, P.; Cole, D. E.; Bailo, B. G.; El-Sohemy, A.; Karmali, M. Type 2 Diabetes Mellitus and Inflammation: Prospects for Biomarkers of Risk and Nutritional Intervention. *Diabetes. Metab. Syndr. Obes.* **2010**, *3*, 173–186.

(7) Karstoft, K.; Pedersen, B. K. Exercise and Type 2 Diabetes: Focus on Metabolism and Inflammation. *Immunol. Cell Biol.* **2016**, *94* (2), 146–150.

(8) Ritchie, S. C.; Würtz, P.; Nath, A. P.; Abraham, G.; Havulinna, A. S.; Fearnley, L. G.; Sarin, A.-P.; Kangas, A. J.; Soininen, P.; Aalto, K.; et al. The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection. *Cell Syst.* **2015**, *1* (4), 293–301.

(9) Ziegler, D. Type 2 Diabetes as an Inflammatory Cardiovascular Disorder. *Curr. Mol. Med.* **2005**, *5* (3), 309–322.

(10) Duncan, B. B.; Schmidt, M. I.; Pankow, J. S.; Ballantyne, C. M.; Couper, D.; Vigo, A.; Hoogeveen, R.; Folsom, A. R.; Heiss, G.; Atherosclerosis Risk in Communities Study. Low-Grade Systemic Inflammation and the Development of Type 2 Diabetes: The Atherosclerosis Risk in Communities Study. *Diabetes* **2003**, *52* (7), 1799–1805.

(11) Hotamisligil, G. S. Inflammation and Metabolic Disorders. *Nature* **2006**, *444* (7121), 860–867.

(12) Akinkuolie, A. O.; Buring, J. E.; Ridker, P. M.; Mora, S. A Novel Protein Glycan Biomarker and Future Cardiovascular Disease Events. *J. Am. Heart Assoc.* **2014**, *3* (5), e001221.

(13) McGarrah, R.; Craig, D.; Haynes, C.; Dowdy, Z. E.; Shah, S.; Kraus, W. GlycA, a Novel Biomarker of Systemic Inflammation,

Improves Cardiovascular Risk Prediction in a High-Risk Coronary Catheterization Cohort. *J. Am. Coll. Cardiol.* **2015**, *65* (10), A1606.

(14) Otvos, J. D.; Shalaurova, I.; Wolak-Dinsmore, J.; Connelly, M. A.; Mackey, R. H.; Stein, J. H.; Tracy, R. P. GlycA: A Composite Nuclear Magnetic Resonance Biomarker of Systemic Inflammation. *Clin. Chem.* **2015**, *61* (5), 714–723.

(15) Fuertes-Martín, R.; Taverner, D.; Vallvé, J.-C.; Paredes, S.; Masana, L.; Correig Blanchar, X.; Amigó Grau, N. Characterization of 1 H NMR Plasma Glycoproteins as a New Strategy To Identify Inflammatory Patterns in Rheumatoid Arthritis. *J. Proteome Res.* **2018**, *17* (11), 3730–3739.

(16) Bell, J. D.; Brown, J. C. C.; Nicholson, J. K.; Sadler, P. J. Assignment of Resonances for 'Acute-Phase' Glycoproteins in High Resolution Proton NMR Spectra of Human Blood Plasma. *FEBS Lett.* **1987**, *215* (2), 311–315.

(17) Mallol, R.; Amigó, N.; Rodríguez, M. A.; Heras, M.; Vinaixa, M.; Plana, N.; Rock, E.; Ribalta, J.; Yanes, O.; Masana, L.; et al. Liposcale: A Novel Advanced Lipoprotein Test Based on 2D Diffusion-Ordered ^{1} H NMR Spectroscopy. *J. Lipid Res.* **2015**, *56* (3), 737–746.

(18) Ballabio, D.; Consonni, V. Classification Tools in Chemistry. Part 1: Linear Models. PLS-DA. *Anal. Methods* **2013**, *5* (16), 3790.

(19) Danesh, J.; Wheeler, J. G.; Hirschfield, G. M.; Eda, S.; Eiriksdottir, G.; Rumley, A.; Lowe, G. D. O.; Pepys, M. B.; Gudnason, V. C-Reactive Protein and Other Circulating Markers of Inflammation in the Prediction of Coronary Heart Disease. *N. Engl. J. Med.* **2004**, *350* (14), 1387–1397.

(20) Hardikar, S.; Song, X.; Kratz, M.; Anderson, G. L.; Blount, P. L.; Reid, B. J.; Vaughan, T. L.; White, E. Intraindividual Variability over Time in Plasma Biomarkers of Inflammation

and Effects of Long-Term Storage. *Cancer Causes Control* **2014**, *25* (8), 969–976.

(21) Dierckx, T.; Verstockt, B.; Vermeire, S.; van Weyenbergh, J. GlycA, a Nuclear Magnetic Resonance Spectroscopy Measure for Protein Glycosylation, Is a Viable Biomarker for Disease Activity in IBD. *J. Crohn's Colitis* **2018**, 1–6.

(22) Gruppen, E. G.; Connelly, M. A.; Vart, P.; Otvos, J. D.; Bakker, S. J.; Dullaart, R. P. GlycA, a Novel Proinflammatory Glycoprotein Biomarker, and High-Sensitivity C-Reactive Protein Are Inversely Associated with Sodium Intake after Controlling for Adiposity: The Prevention of Renal and Vascular End-Stage Disease Study. *Am. J. Clin. Nutr.* **2016**, *104* (2), 415–422.

(23) Gruppen, E. G.; Connelly, M. A.; Dullaart, R. P. F. Higher Circulating GlycA, a pro-Inflammatory Glycoprotein Biomarker, Relates to Lipoprotein-Associated Phospholipase A2 Mass in Nondiabetic Subjects but Not in Diabetic or Metabolic Syndrome Subjects. *J. Clin. Lipidol.* **2016**, *10* (3), 512–518.

(24) Dullaart, R. P. F.; Gruppen, E. G.; Connelly, M. A.; Lefrandt, J. D. A Pro-Inflammatory Glycoprotein Biomarker Is Associated with Lower Bilirubin in Metabolic Syndrome. *Clin. Biochem.* **2015**, *48* (16–17), 1045–1047.

(25) Dullaart, R. P. F.; Gruppen, E. G.; Connelly, M. A.; Otvos, J. D.; Lefrandt, J. D. GlycA, a Biomarker of Inflammatory Glycoproteins, Is More Closely Related to the Leptin/Adiponectin Ratio than to Glucose Tolerance Status. *Clin. Biochem.* **2015**, *48* (12), 811–814.

(26) WHO. Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus: Abbreviated Report of a WHO Consultation. *Geneva World Heal. Organ.* **2011**.

(27) Morgner, F.; Lecointre, A.; Charbonnière, L. J.; Löhmannsröben, H.-G. Detecting Free Hemoglobin in Blood Plasma and Serum

with Luminescent Terbium Complexes. *Phys. Chem. Chem. Phys.* **2015**, *17* (3), 1740–1745.

(28) Krauss, R. M. Lipids and Lipoproteins in Patients with Type 2 Diabetes. *Diabetes Care* **2004**, *27* (6), 1496–1504.

(29) Amigó, N.; Mallol, R.; Heras, M.; Martínez-Hervás, S.; Blanco-Vaca, F.; Escolà-Gil, J. C.; Plana, N.; Yanes, Ó.; Masana, L.; Correig, X. Lipoprotein Hydrophobic Core Lipids Are Partially Extruded to Surface in Smaller HDL: "Herniated" HDL, a Common Feature in Diabetes. *Sci. Rep.* **2016**, *6* (1), 19249.

(30) Wurtz, P.; Tiainen, M.; Makinen, V.-P.; Kangas, A. J.; Soininen, P.; Saltevo, J.; Keinanen-Kiukaanniemi, S.; Mantyselka, P.; Lehtimaki, T.; Laakso, M.; et al. Circulating Metabolite Predictors of Glycemia in Middle-Aged Men and Women. *Diabetes Care* **2012**, *35* (8), 1749–1756.

(31) Akinkuolie, A. O.; Pradhan, A. D.; Buring, J. E.; Ridker, P. M.; Mora, S. Novel Protein Glycan Side-Chain Biomarker and Risk of Incident Type 2 Diabetes Mellitus. *Arterioscler. Thromb. Vasc. Biol.* **2015**, *35* (6), 1544–1550.

(32) Dungan, K.; Binkley, P.; Osei, K. GlycA Is a Novel Marker of Inflammation Among Non-Critically Ill Hospitalized Patients with Type 2 Diabetes. *Inflammation* **2015**, *38* (3), 1357–1363.

(33) Connelly, M. A.; Shimizu, C.; Winegar, D. A.; Shalaurova, I.; Pourfarzib, R.; Otvos, J. D.; Kanegaye, J. T.; Tremoulet, A. H.; Burns, J. C. Differences in GlycA and Lipoprotein Particle Parameters May Help Distinguish Acute Kawasaki Disease from Other Febrile Illnesses in Children. *BMC Pediatr.* **2016**, *16* (1).

(34) Lorenzo, C.; Festa, A.; Hanley, A. J.; Rewers, M. J.; Escalante, A.; Haffner, S. M. Novel Protein Glycan–Derived Markers of Systemic Inflammation and C-Reactive Protein in Relation to Glycemia, Insulin Resistance, and Insulin Secretion. *Diabetes Care* **2017**, *40* (3), 375–382.

CHAPTER 7

7. General discussion and conclusions

7.1. Glycoprotein quantification, a further step in NMR applications

Nuclear Magnetic Resonance (NMR) spectroscopy enables us to obtain a comprehensive analytical profile including the most abundant chemical components such as lipoproteins, aminoacids, organic acids, carbohydrates and albumin found in biological samples.[1,2] In this thesis, we focused on the quantification of glycoprotein acetyls in serum or plasma samples, which becomes a further step in the advancement of NMR applications, especially considering that they can be quantified from the same experiment used for lipids, which means great cost savings and maximum profitability. Another major advantage associated with the study of serum or plasma with NMR is that the measurements can often be made with minimal sample preparation, even the process can be automated with sample preparation robots such as the Automated Liquid Handling Gilson robot, which has been implemented in recent years in our company. The reliability of a high throughput technique such as NMR and the ease of sample preparation in turn mean that the coefficient of variation associated with inter- and intra-laboratory NMR experiments is typically very low and biological differences between individuals are normally greater than the variations induced by experimental factors.[3]

For glycoprotein analysis, other classic techniques such as Lectin analysis,[4] DNA sequencing equipment-Fluorophore Assisted Carbohydrate Electrophoresis (DSA-FACE),[5] Chromatographic methods such as high performance liquid chromatography (HPLC)[6] or Hydrophilic interaction liquid chromatography (HILIC),[4] or MALDI-

TOF analysis,[4] among others, are usually targeted techniques that only allow to characterize specific glycated proteins. Mass spectrometry (MS) is another common high throughput technology but the analysis of protein-derived circulating glycans is still difficult due to the high complexity of glycan linkage and branching.[7] However, with NMR, although we do not characterize the specific glycated protein, we are able to determine a quantification of global glycosylation status, specifically of the most abundant acute-phase proteins in the blood, which are elevated during inflammatory processes,[8] so it becomes an advantageous technique for quantifying a global inflammation profile.

In particular, the measurement of glycoproteins by NMR is not as established as the measurement of lipoproteins, which has been a revolution in the last decade, but it is likely to be used in the very close future because of the above-mentioned advantages.

In **Chapter 2**, we present our methodology for the characterization of glycoproteins based on NMR spectroscopy, as well as the industrial development, which has been made possible thanks to the cooperation between the Rovira i Virgili university and the spin off company Biosfer Teslab, whose main product is the Liposcale® test, a test based on the characterization of lipoproteins through NMR spectroscopy. We assessed the robustness, reliability and reproducibility of the methodology, as well as the effect of sample storage conditions and sample interday assays. The performed experiments indicated a very low variability in the results of glycoprotein quantification.

There are currently two NMR glycoprotein analysis methods developed by other research groups that have been applied in a wide variety of studies, based on curve fitting methods.[1,8,9] Other approaches quantifying glycoproteins generally rely on the integration of their NMR signal, which is proportional to the concentration. However, this approach is very sensitive to baseline distortions and it is not recommended for overlapped peaks. It's worth to say that, as we have seen in the thesis, the GlycA peak is very close to a lipid response peak. If a peak deconvolution approach is not used, the generated error in the GlycA concentration can be very high.

The main difference of our methodology with respect to those previously mentioned is that we obtain some parameters that have not been previously described by using the LED pulse spectra, while others use regression methods and the NOESY pulse. Normally, studies in the literature refer to GlycA and GlycB, which report the concentration of the acetyl groups of N-Acetylglucosamine (GlcNAc) and N-Acetylgalactosamine (GalNAc), and sialic acid or N-Acetilneuraminic acid (Neu5Ac) monosaccharides bound to protein, respectively. The methodology for glycoprotein profiling developed in the present thesis is also based on a line-shape fitting approach but in LED spectra, with the difference that provides not only GlycA and GlycB area. We quantify a new parameter that we have called ratio H/W of GlycA and GlycB related to the shape of the NMR peaks defined by the ratio between the height - which is related to concentration - and the width - which is related to the flexibility and the aggregation of the molecules generating the signal. These new parameters have showed higher clinical relevance throughout the thesis studies. In addition, we also

quantify another new parameter that we have called GlycF, which is associated with the concentration of the three free monosaccharides (GalNAc, GlcNAc and Neu5Ac) in the sample without binding to protein. However, more studies are required in the future to establish the specificity of each of them beyond their increase in inflammatory or metabolic diseases.

In order to prepare NMR-based on glycoprotein profiling applications for clinical use, one aspect to consider in the near future is a comparative study between the different quantification methodologies. The standardization of both, NMR measurements and the glycoprotein parameters, will allow the comparison of different cohorts by providing new knowledge about glycosylation patterns.

7.2. Perspectives on NMR-based on glycoprotein quantification into clinical practice

Inflammation is the basis for many autoimmune and chronic low grade inflammatory diseases.[7] The inflammatory response leads to an increase in acute phase proteins (APP) such as tumour necrosis factor (TNF-α), interleukin-1 and 6 (IL-1 and IL-6), among others.[10] Many APPs are heavily glycosylated[11] and they are predominantly synthesized and secreted by hepatocytes. During an acute phase response, glycosylation of proteins is affected and the carbohydrate structures of many of N-linked inflammatory glycoproteins branch out.[12–14] These branched glycans are rich in GlcNAc, GalNAc, sialic acid and fucose residues in a myriad of different arrangements, contributing to the potential diversity of glycan structures.[12,13,15–17]

Given the diversity in the number of glycoproteins in biological fluids, as well as the unique changes that can occur in some diseases and not others, it is likely that there is a large amount of information that has not yet been extracted from glycoproteins for clinical use. The state of art presented in **Chapter 3** reflects the growing interest in circulating glycoproteins quantification by ^{1}H-NMR spectroscopy and the related clinical studies. GlycA and GlycB parameters have been studied in a wide variety of diseases such as tumours and cancer, [18–23] metabolic diseases, [24–26] cardiovascular risk, [27–33] chronic inflammatory diseases, [34–36] among others. All of these studies support the hypothesis that GlycA is a robust marker of systemic inflammation. GlycB, however, is believed to be associated also with inflammatory diseases but with less strength than GlycA. The results of the review chapter open the doors to a future incorporation of these markers for prediction and monitoring of several diseases.

In this thesis, we have studied the clinical relevance of the glycoprotein parameters obtained by our new NMR analysis method in three different cohorts with chronic inflammatory disease such as rheumatoid arthritis and metabolic diseases such as diabetes, obesity and PCOS. In **chapter 4** we have characterized the plasma glycoprotein profile of a cohort of patients with rheumatoid arthritis (RA) versus healthy individuals. RA patients presented a significant increase in the GlycA associated area compared with the control group. They also presented significantly higher H/W GlycA and GlycB ratios than the control population. In a prediction PLS-DA model we calculated the receiver operating curve (ROC) in order to evaluate the classificatory power of our glycoprotein and lipoprotein (Liposcale)

^{1}H-NMR profile to classify the study population in low or high disease activity degree. The results demonstrated that our glycoprotein and lipoprotein (Liposcale) ^{1}H-NMR variables, along with classic inflammatory parameters such as Fibrinogen, C-reactive protein (CRP), rheumatoid factor (RF) and Anti-Citrullinated Peptide Antibodies (ACPA), provide information about the high activity of the disease that is more accurate than if we only consider the classic inflammatory parameters, widely used in clinical practice.[37]

Chapter 5 evaluates the glycoprotein profiles of non-obese and obese women with polycystic ovary syndrome (PCOS) comparing them with non-obese and obese female and male controls. Even though additional research is required to confirm our findings, the results of the current exploratory investigation pave the way toward the use of ^{1}H-NMR-derived glycoprotein profiles as diagnostic markers to predict chronic low-grade inflammation in patients with PCOS. Obese subjects presented higher GlycA, GlycB and GlycF areas and higher H/W GlycA and GlycB ratios than their non-obese counterparts, regardless of sex and PCOS. Patients with PCOS also presented greater H/W ratios of GlycA and GlycB compared with the female and male control groups.

Finally, **Chapter 6** describes our latest study reported in this thesis in which we characterized plasma glycoprotein profiles of patients with type 2 diabetes (T2DM) with and without atherogenic dyslipidaemia (AD).The results shown that a pro-atherogenic lipoprotein pattern that included reduced HDL-C levels, increased small dense LDL and HDL particles, and elevated TG levels was

significantly associated with glycoprotein variables. Along the same lines, all glycoprotein variables values in the diagnostic groups were significantly higher than those in the control groups. AD and T2DM conditions together contribute to a positive and significant synergetic effect on the GlycA area and the H/W ratios of GlycA and GlycB. In the last part of this study, we calculated three PLS-DA models in order to evaluate if glycoprotein variables added some discrimination power of to CRP, the universal clinical variable used commonly for inflammation assessment; the first model is between the control population and the rest of the diagnostic groups; the second one between the group that presented DA and the group that did not; and finally, the third one between the group that presented DM and the group that did not. The results shown that by adding the new glycoprotein variables to the traditionally used marker of inflammation CRP, the AUC increased sharply for the three models suggesting a higher classification power than the CRP.

All the diseases contemplated in these chapters have a low-grade chronic inflammation base. Therefore, the study of new markers of inflammation such as glycoproteins can be very useful in clinical, as has been demonstrated in the results.

In the studies that we have carried out in chapters 5 and 6, the cooperation of two factors has been evaluated, AD and DM in chapter 5, and PCOS and obesity in chapter 6. The existence of both factors at the same time supposes an increase in the concentration values of glycoproteins. Although more studies should be conducted to understand the inflammatory processes that trigger AD, DM, PCOS

and obesity, we believe that there may be a greater amount of protein glycation when two factors cooperate at once, due to the increased concentration of acetyl groups that we have detected by using our NMR method.

Special mention should be made of the added value that our new glycoprotein profile brings to the traditionally ones in clinics. Chapter 4 clearly shows how glycoproteins, this time together with some lipoprotein variables (both obtained with a single NMR experiment), far surpass the classifying power of the traditional CRP, RF and ACPA evaluating RA patients with a greater or lesser degree of the disease activity.

At the same time, **chapter 6** clearly shows a marked improvement in diagnostic models with the progressive addition of glycoprotein variables with respect to CRP. Although the CRP has been considered as the main inflammatory biomarker widely studied since its discovery in 1930 , it has been shown to be prone to fluctuations due to causes of different origin.[38] Glycoproteins acetyls NMR signal becomes a measure that reflects the integrated concentrations of several of the most abundant glycoproteins in the circulation, more than specific molecular markers of inflammation,. In relation to this, we hypothesize that glycoprotein acetyls concentrations may provide a more stable measure of low-grade systemic inflammation and respond more uniformly to diverse inflammatory stimuli than individual inflammatory reactants. Our presented results are in line with the results exposed at chapter 3 about the state of the art of NMR glycoproteins. Taking this into account, the

long-term application of NMR-based glycoprotein analysis in medical research is obviously an encouraging example for the epidemiological and clinical prospects of NMR-based on technologies.

Other clinical issues to consider in the near future is the usefulness of these markers to measure the effect of pharmacological treatments. This aspect is not well studied, but it is a challenge for the clinic to achieve reliable markers to measure the effect of drugs in order to improve treatment guidelines. If the glycoprotein profile, which appear to be more stable than traditional inflammatory markers, can be modified by the effect of a drug, it would open a new window to become a clinically relevant biomarker to monitor the severity of the disease and the treatment efficacy.

7.3. Future Industrial development

Additional developments toward automated high-throughput applications of quantitative NMR metabolomics are planned in the near future. The present thesis has been partly developed in the facilities of Biosfer Teslab. It directly opens the door to the introduction of the quantification of glycoproteins in multiple and diverse research studies in collaboration with different national and international entities with the aim in the future of entering the clinic, expanding services beyond the Liposcale test. This fact has a fundamental advantage; NMR experiments to quantify glycoproteins are the same for characterization of lipoprotein and lipid subclasses. Therefore, the stored original NMR data allow retrospective analysis of the glycoprotein profile in all

archived studies. These data and the continuous analyses carried out in the company will probably speed up the performance of new studies in order to clarify and define the role of each of the parameters we have defined in our method for the first time (GlycB, GlycA and GlycF areas and GlycA and GlycB H/W ratios) to establish behaviour patterns.

Altogether, the results presented in this thesis demonstrate that the use of NMR is ideal for the routine characterization of glycoproteins in the research, the clinical and the epidemiological frameworks, due to the simplicity of the preparation of the sample, the reliability and the robustness of the ^{1}H-NMR spectroscopy and the possibility of automatization. In addition, the same measurement protocol is used for lipoprotein characterization. The results presented in this thesis and others epidemiological and clinical studies that will be carried out soon, must show clear evidence that glycoproteins quantified by NMR are robust and provide enough additional biological and predictive value to be incorporated into new clinical protocols.

7.4. References

(1) Soininen, P.; Kangas, A. J.; W€, P.; Tukiainen, T.; Tynkkynen, T.; Laatikainen, R.; J€, M.-R.; K€ Ah€ Onen, M.; Lehtim€, T.; Viikari, J.; et al. High-Throughput Serum NMR Metabonomics for Cost-Effective Holistic Studies on Systemic Metabolism †. *Analyst* **2009**, *134* (9), 1781–1785.

(2) Nicholson, J. K.; Foxall, P. J. D.; Spraul, M.; Farrant, R. D.; Lindon, J. C. 750 MHz 1H and 1H-13C NMR Spectroscopy of

Human Blood Plasma. *Anal. Chem.* **1995**, *67* (5), 793–811.

(3) Lenz, E. M.; Bright, J.; Wilson, I. D.; Morgan, S. R.; Nash, A. F. P. A 1H NMR-Based Metabonomic Study of Urine and Plasma Samples Obtained from Healthy Human Subjects. *J. Pharm. Biomed. Anal.* **2003**, *33* (5), 1103–1115.

(4) Pilobello, K. T.; Mahal, L. K. Lectin Microarrays for Glycoprotein Analysis. *Methods Mol. Biol.* **2007**, *385*, 193–203.

(5) Vanhooren, V.; Laroy, W.; Libert, C.; Chen, C. N-Glycan Profiling in the Study of Human Aging. *Biogerontology* **2008**, *9* (5), 351–356.

(6) Gornik, O.; Lauc, G. Glycosylation of Serum Proteins in Inflammatory Diseases. *Dis. Markers* **2008**, *25* (4–5), 267–278.

(7) Connelly, M. A.; Gruppen, E. G.; Otvos, J. D.; Dullaart, R. P. F. Inflammatory Glycoproteins in Cardiometabolic Disorders, Autoimmune Diseases and Cancer. *Clin. Chim. Acta* **2016**, *459*, 177–186.

(8) Otvos, J. D.; Shalaurova, I.; Wolak-Dinsmore, J.; Connelly, M. A.; Mackey, R. H.; Stein, J. H.; Tracy, R. P. GlycA: A Composite Nuclear Magnetic Resonance Biomarker of Systemic Inflammation. *Clin. Chem.* **2015**, *61* (5), 714–723.

(9) Ala-Korpela, M. 1H NMR Spectroscopy of Human Blood Plasma. *Prog. Nucl. Magn. Reson. Spectrosc.* **1995**, *27* (5–6), 475–554.

(10) Jain, S.; Gautam, V.; Naseem, S. Acute-Phase Proteins: As Diagnostic Tool. *J. Pharm. Bioallied Sci.* **2011**, *3* (1), 118–127.

(11) Dempsey, E.; Rudd, P. M. Acute Phase Glycoproteins: Bystanders or Participants in Carcinogenesis? *Ann. N. Y. Acad. Sci.* **2012**, *1253* (1), 122–132.

(12) van Dijk, W.; Turner, G. A.; Mackiewicz, A. Changes in Glycosylation of Acute-Phase Proteins in Health and Disease: Occurrence, Regulation and Function. *Glycoconj. J.* **1994**, *1* (1), 5–14.

(13) Ceciliani, F.; Pocacqua, V. The Acute Phase Protein Alpha1-Acid Glycoprotein: A Model for Altered Glycosylation during Diseases. *Curr. Protein Pept. Sci.* **2007**, *8* (1), 91–108.

(14) Gornik, O.; Lauc, G. Glycosylation of Serum Proteins in Inflammatory Diseases. *Dis. Markers* **2008**, *25* (4–5), 267–278.

(15) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **2006**, *126* (5), 855–867.

(16) McCarthy, C.; Saldova, R.; Wormald, M. R.; Rudd, P. M.; McElvaney, N. G.; Reeves, E. P. The Role and Importance of Glycosylation of Acute Phase Proteins with Focus on Alpha-1 Antitrypsin in Acute and Chronic Inflammatory Conditions. *J. Proteome Res.* **2014**, *13* (7), 3131–3143.

(17) Hart, G. W.; Copeland, R. J. Glycomics Hits the Big Time. *Cell* **2010**, *143* (5), 672–676.

(18) Kolwijck, E.; Engelke, U. F.; van der Graaf, M.; Heerschap, A.; Blom, H. J.; Hadfoune, M.; Buurman, W. A.; Massuger, L. F.; Wevers, R. A. *N* -Acetyl Resonances in *in Vivo* and *in Vitro* NMR Spectroscopy of Cystic Ovarian Tumors. *NMR Biomed.* **2009**, n/a-n/a.

(19) Bell, J. D.; Brown, J. C. C.; Nicholson, J. K.; Sadler, P. J. Assignment of Resonances for 'Acute-Phase' Glycoproteins in High Resolution Proton NMR Spectra of Human Blood Plasma. *FEBS Lett.* **1987**, *215* (2), 311–315.

(20) Hasim, A.; Ma, H.; Mamtimin, B.; Abudula, A.; Niyaz, M.; Zhang, L.; Anwer, J.; Sheyhidin, I. Revealing the Metabonomic

Variation of EC Using 1H-NMR Spectroscopy and Its Association with the Clinicopathological Characteristics. *Mol. Biol. Rep.* **2012**, *39* (9), 8955–8964.

(21) Jobard, E.; Pontoizeau, C.; Blaise, B. J.; Bachelot, T.; Elena-Herrmann, B.; Trédan, O. A Serum Nuclear Magnetic Resonance-Based Metabolomic Signature of Advanced Metastatic Human Breast Cancer. *Cancer Lett.* **2014**, *343* (1), 33–41.

(22) Deja, S.; Porebska, I.; Kowal, A.; Zabek, A.; Barg, W.; Pawelczyk, K.; Stanimirova, I.; Daszykowski, M.; Korzeniewska, A.; Jankowska, R.; et al. Metabolomics Provide New Insights on Lung Cancer Staging and Discrimination from Chronic Obstructive Pulmonary Disease. *J. Pharm. Biomed. Anal.* **2014**, *100*, 369–380.

(23) Suman, S.; Sharma, R. K.; Kumar, V.; Sinha, N.; Shukla, Y. Metabolic Fingerprinting in Breast Cancer Stages through 1H NMR Spectroscopy-Based Metabolomic Analysis of Plasma. *J. Pharm. Biomed. Anal.* **2018**, *160*, 38–45.

(24) Dullaart, R. P. F.; Gruppen, E. G.; Connelly, M. A.; Otvos, J. D.; Lefrandt, J. D. GlycA, a Biomarker of Inflammatory Glycoproteins, Is More Closely Related to the Leptin/Adiponectin Ratio than to Glucose Tolerance Status. *Clin. Biochem.* **2015**, *48* (12), 811–814.

(25) Gruppen, E. G.; Connelly, M. A.; Otvos, J. D.; Bakker, S. J. L.; Dullaart, R. P. F. A Novel Protein Glycan Biomarker and LCAT Activity in Metabolic Syndrome. *Eur. J. Clin. Invest.* **2015**, *45* (8), 850–859.

(26) Gruppen, E. G.; Connelly, M. A.; Vart, P.; Otvos, J. D.; Bakker, S. J.; Dullaart, R. P. GlycA, a Novel Proinflammatory Glycoprotein Biomarker, and High-Sensitivity C-Reactive Protein Are Inversely Associated with Sodium Intake after

Controlling for Adiposity: The Prevention of Renal and Vascular End-Stage Disease Study. *Am. J. Clin. Nutr.* **2016**, *104* (2), 415–422.

(27) Connelly, M. A.; Shimizu, C.; Winegar, D. A.; Shalaurova, I.; Pourfarzib, R.; Otvos, J. D.; Kanegaye, J. T.; Tremoulet, A. H.; Burns, J. C. Differences in GlycA and Lipoprotein Particle Parameters May Help Distinguish Acute Kawasaki Disease from Other Febrile Illnesses in Children. *BMC Pediatr.* **2016**, *16* (1).

(28) Chandler, P. D.; Akinkuolie, A. O.; Tobias, D. K.; Lawler, P. R.; Li, C.; Moorthy, M. V.; Wang, L.; Duprez, D. A.; Jacobs, D. R.; Glynn, R. J.; et al. Association of N-Linked Glycoprotein Acetyls and Colorectal Cancer Incidence and Mortality. *PLoS One* **2016**, *11* (11), e0165615.

(29) Duprez, D. A.; Otvos, J.; Sanchez, O. A.; Mackey, R. H.; Tracy, R.; Jacobs, D. R. Comparison of the Predictive Value of GlycA and Other Biomarkers of Inflammation for Total Death, Incident Cardiovascular Events, Noncardiovascular and Noncancer Inflammatory-Related Events, and Total Cancer Events. *Clin. Chem.* **2016**, *62* (7), 1020–1031.

(30) Akinkuolie, A. O.; Pradhan, A. D.; Buring, J. E.; Ridker, P. M.; Mora, S. Novel Protein Glycan Side-Chain Biomarker and Risk of Incident Type 2 Diabetes Mellitus. *Arterioscler. Thromb. Vasc. Biol.* **2015**, *35* (6), 1544–1550.

(31) Akinkuolie, A. O.; Buring, J. E.; Ridker, P. M.; Mora, S. A Novel Protein Glycan Biomarker and Future Cardiovascular Disease Events. *J. Am. Heart Assoc.* **2014**, *3* (5), e001221.

(32) Fischer, K.; Kettunen, J.; Würtz, P.; Haller, T.; Havulinna, A. S.; Kangas, A. J.; Soininen, P.; Esko, T.; Tammesoo, M.-L.; Mägi, R.; et al. Biomarker Profiling by Nuclear Magnetic Resonance Spectroscopy for the Prediction of All-Cause Mortality: An Observational Study of 17,345 Persons. *PLoS Med.* **2014**, *11* (2),

e1001606.

(33) Muhlestein, J. B.; May, H.; Winegar, D.; Rollo, J.; Connelly, M.; Otvos, J.; Anderson, J. GlycA and GlycB, Novel NMR Biomarkers of Inflammation, Strongly Predict Future Cardiovascular Events, but Not the Presence of Coronary Artery Disease (CAD), among Patients Undergoing Coronary Angiography: The Intermountain Heart Collaborative Study. *J. Am. Coll. Cardiol.* **2014**, *63* (12), A1389.

(34) Fuertes-Martín, R.; Taverner, D.; Vallvé, J.-C.; Paredes, S.; Masana, L.; Correig Blanchar, X.; Amigó Grau, N. Characterization of ^{1} H NMR Plasma Glycoproteins as a New Strategy To Identify Inflammatory Patterns in Rheumatoid Arthritis. *J. Proteome Res.* **2018**, *17* (11), 3730–3739.

(35) Bartlett, D. B.; Connelly, M. A.; AbouAssi, H.; Bateman, L. A.; Tune, K. N.; Huebner, J. L.; Kraus, V. B.; Winegar, D. A.; Otvos, J. D.; Kraus, W. E.; et al. A Novel Inflammatory Biomarker, GlycA, Associates with Disease Activity in Rheumatoid Arthritis and Cardio-Metabolic Risk in BMI-Matched Controls. *Arthritis Res. Ther.* **2016**, *18* (1), 86.

(36) Ormseth, M. J.; Chung, C. P.; Oeser, A. M.; Connelly, M. A.; Sokka, T.; Raggi, P.; Solus, J. F.; Otvos, J. D.; Stein, C. M. Utility of a Novel Inflammatory Marker, GlycA, for Assessment of Rheumatoid Arthritis Disease Activity and Coronary Atherosclerosis. *Arthritis Res. Ther.* **2015**, *17* (1), 117.

(37) Fuertes-Martín, R.; Taverner, D.; Vallvé, J.-C.; Paredes, S.; Masana, L.; Correig Blanchar, X.; Amigó Grau, N. Characterization of 1 H NMR Plasma Glycoproteins as a New Strategy To Identify Inflammatory Patterns in Rheumatoid Arthritis. *J. Proteome Res.* **2018**, *17* (11), 3730–3739.

(38) Lloyd-Jones, D. M.; Liu, K.; Tian, L.; Greenland, P. Narrative Review: Assessment of C-Reactive Protein in Risk Prediction for

Cardiovascular Disease. *Ann. Intern. Med.* **2006**, *145* (1), 35–42.

(39) Maachi, M.; Piéroni, L.; Bruckert, E.; Jardel, C.; Fellahi, S.; Hainque, B.; Capeau, J.; Bastard, J.-P. Systemic Low-Grade Inflammation Is Related to Both Circulating and Adipose Tissue TNFα, Leptin and IL-6 Levels in Obese Women. *Int. J. Obes.* **2004**, *28* (8), 993–997.

(40) Bogdański, P.; Chyrek, R.; Pupek-Musialik, D.; Jabłecka, A. Evaluation of Selected Acute Phase Proteins in Patients with Metabolic Syndrome. *Pol. Merkur. Lekarski* **2006**, *21* (121), 12–14.

UNIVERSITAT ROVIRA I VIRGILI
1H-NMR GLYCOPROTEIN ANALYSIS: AN ADVANCED APPROACH FOR INFLAMMATORY DISEASES DIAGNOSIS
Rocio Fuertes Martín

UNIVERSITAT ROVIRA I VIRGILI
1H-NMR GLYCOPROTEIN ANALYSIS: AN ADVANCED APPROACH FOR INFLAMMATORY DISEASES DIAGNOSIS
Rocio Fuertes Martín

UNIVERSITAT ROVIRA I VIRGILI
1H-NMR GLYCOPROTEIN ANALYSIS: AN ADVANCED APPROACH FOR INFLAMMATORY DISEASES DIAGNOSIS
Rocio Fuertes Martín

UNIVERSITAT ROVIRA I VIRGILI
1H-NMR GLYCOPROTEIN ANALYSIS: AN ADVANCED APPROACH FOR INFLAMMATORY DISEASES DIAGNOSIS
Rocio Fuertes Martín

UNIVERSITAT ROVIRA i VIRGILI