Characterization of Glycoprotein and Lipoprotein Profiles of Rheumatoid Arthritis (RA) patiens by ^{1}H-Nuclear Magnetic Resonance Spectroscopy (^{1}H-NMR).

Rocío Fuertes Martín[1], Núria Amigó Grau[1], Joan Carles Vallvé[2], Silvia Paredes[2], Dèlia Taverner[2], Lluis Masana[2]. Xavier Correig Blanchar[3].

[1]Biosfer Teslab,SL; Metabolomics platform, URV; IISPV; CIBERDEM; Reus-Tarragona, Spain.
[2]URLA, Hospital Universitari Sant Joan de Reus; IISPV; CIBERDEM; Reus-Tarragona, Spain.
[3]Metabolomics platform, Universitat Rovira i Virgili, IISPV; CIBERDEM; Reus-Tarragona, Spain.
rfuertes@biosferteslab.com

Background

Rheumatoid Arthritis (RA) is an autoimmune and chronic inflammatory disease associated with a high index of morbidity and mortality by cardiovascular diseases (CVDs)[1]. ^{1}H-NMR is a technique capable of determining the lipoprotein and glycoprotein profiles to characterize dyslipidaemias and to estimate the cardiovascular and to evaluate inflammatory state.

Recent studies have established the importance of glycosylated proteins in important biological processes: cell adhesion, transport, signal transduction and, especially, control of cellular inflammation.

The proton nuclear magnetic resonance spectroscopy (^{1}H-NMR) is emerging as a technique able to detect levels of circulating glycoproteins in a quick and accurate way [2].

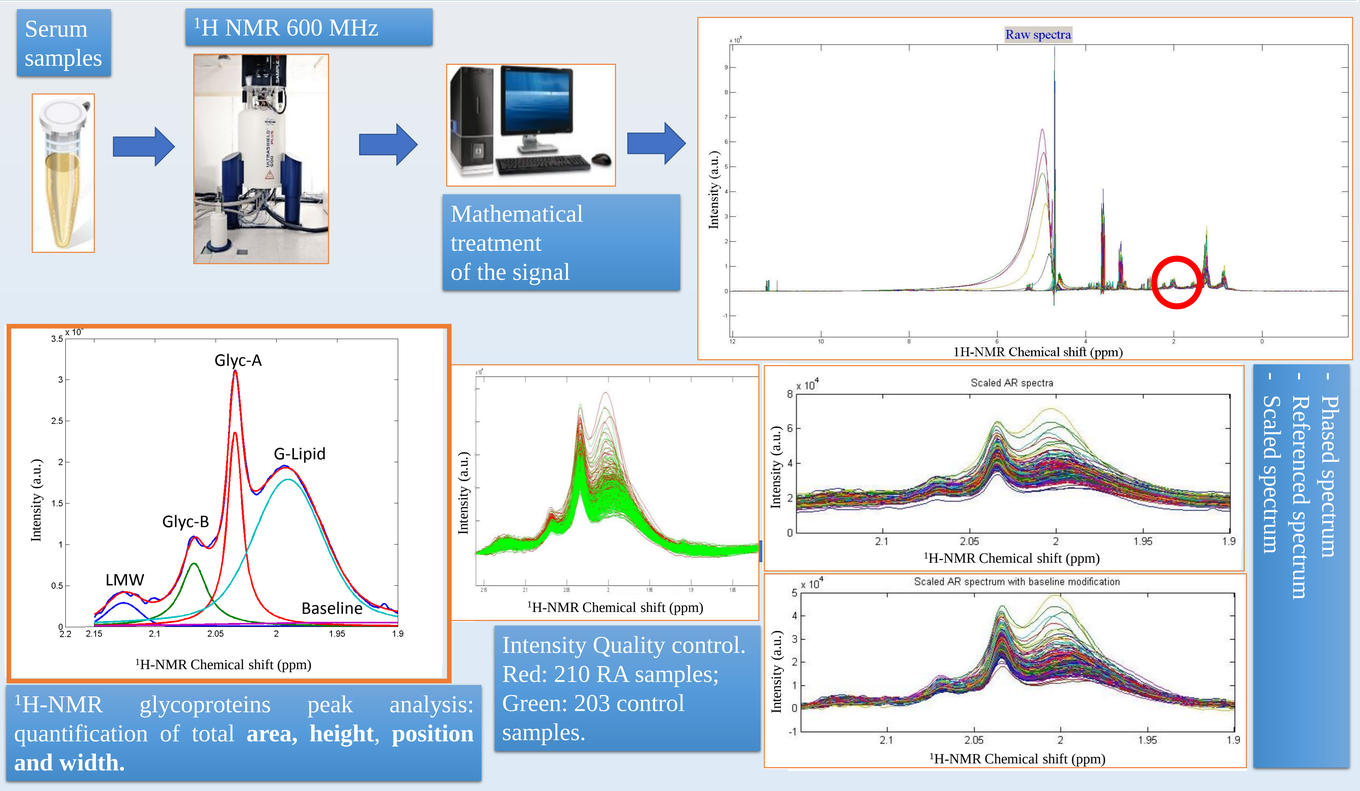
Aim

This study aims to characterize by ^{1}H-NMR the plasma glycoprotein profile of patients with RA versus healthy individuals and to identify patterns indicating the severity of the disease.

Methodology

Serum samples of 214 RA patients and 203 healthy individuals matched by sex, age and body mass index were analysed by ^{1}H-NMR. The glycoprotein and lipoprotein [3] profile was characterized by the study of the ^{1}H-NMR spectrum of both groups. A mathematical treatment of the signal was conducted as follows: the raw spectra was fitted with five analytic Lorentzian/Gaussian functions reproducing the peak associated with the glycoproteins (LMW, Glyc-B, Glyc-A, and Baseline) based on their chemical shift. From each of these functions the total area and height (proportional to the concentration), 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. For each function we calculated the derived parameter H/W=Height/Width to capture the shape of the peaks.

On the other hand, the traditional inflammatory markers such as C-reactive protein (CRP), fibrinogen, glycosylated haemoglobin (HbA1c) and the erythrocyte sedimentation rate (ESR) were determined. Lipid variables such as total cholesterol, LDL (low density lipoprotein) cholesterol, HDL (high density lipoprotein) cholesterol, very low density lipoproteins (VLDL), total triglicerides (TG), apoprotein A1 (ApoA1) and apoprotein B (ApoB) were also determined by traditional biochemistry methods.



^{1}H-NMR glycoproteins peak analysis: quantification of total **area, height, position and width.**

- Univariate statistical analysis of the glycoprotein variables was conducted to identify differences between the RA patients group and the healthy individuals group.
- Associations between the lipid variables and inflammatory markers determined by biochemistry methods and the ^{1}H-NMR glycoprotein profile were also studied.
- A Partial Least Squares-Discriminant Analysis (PLS-DA) was conducted to identify the characteristic glycoprotein profiles associated with the AR disease and the healthy group.
- Different PLS-DA models were evaluated to build the best predictor model for the highest AR disease severity individuals according to the DAS28 index scale (>75th Percentile) by using the traditional inflammatory markers and ^{1}H-NMR parameters (glycoprotein and lipoprotein) as input variables.

Results

We found statistically significant differences in several ^{1}H-NMR glycoproteins peak variables highlighting the values of *total area*, *Heigth/width Glyc-A* and *Heigth/width Glyc-B* variables between the AR patients and the control group (p values $9.17e^{-15}$, $1.15e^{-09}$ and $1.87e^{-09}$ respectively) (Figure 1). The PLS-DA could perfectly discriminate the AR and healthy groups (Figure 2). Figure 3 depicts the contribution of each variable (loadings) on the maximized variance between the two groups.

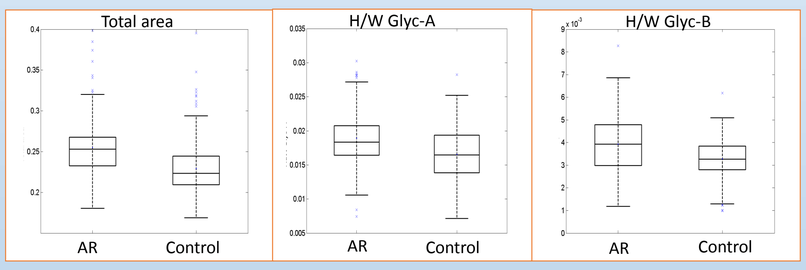


Figure 1: Boxplots of the total area of total area, Heigth/width Glyc-A and Heigth/width Glyc-B variables.

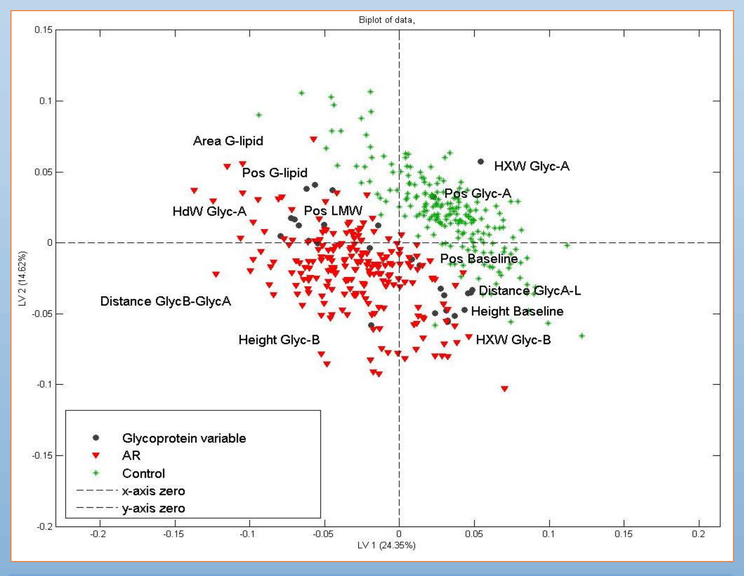


Figure 2: Biplot with the 28 ^{1}H-NMR glycoprotein variables calculated values for the control group in green and for the AR group in red.

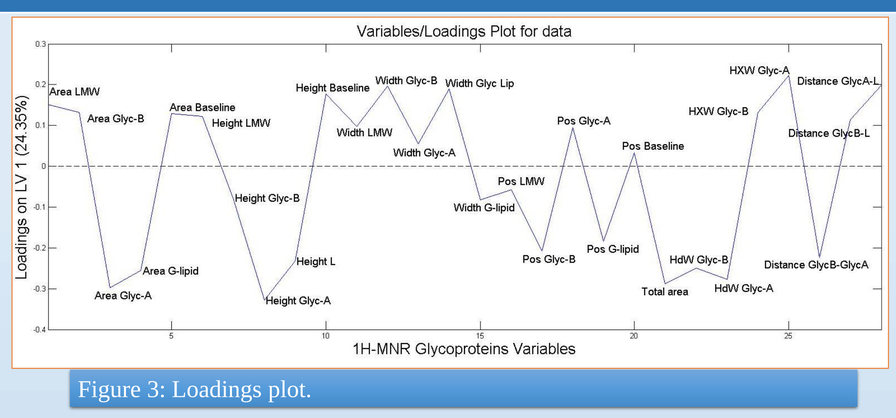


Figure 3: Loadings plot.

In addition, the concentration of total cholesterol, VLDL and triglycerides were significantly associated with the concentration of glycoprotein parameters. The ESR, PCR and fibrinogen were positively associated with the concentration of glycoproteins (Table 1).

	Col	HDLc	LDLc	VLDL	TG	ESR	PCR	Fibrinogen
Area LMW	**-0.19**	**0.32**	**-0.20**	**-0.58**	**-0.58**	0.13	**0.21**	0.13
Area Glyc-B	-0.06	**0.21**	-0.07	**-0.34**	**-0.34**	**0.37**	**0.25**	**0.25**
Area Glyc-A	-0.02	**-0.26**	-0.04	**0.50**	**0.50**	**0.33**	**0.31**	**0.36**
Area G-lipid	**0.56**	-0.04	**0.42**	**0.73**	**0.73**	-0.04	-0.08	-0.06
Area Baseline	**-0.14**	0.12	-0.10	**-0.35**	**-0.35**	-0.07	-0.05	0.00
Height LMW	**-0.14**	**0.24**	-0.13	**-0.49**	**-0.50**	0.07	**0.17**	0.10
Height Glyc-B	0.00	0.12	-0.01	**-0.17**	**-0.17**	**0.57**	**0.48**	**0.53**
Height Glyc-A	0.07	**-0.18**	0.03	**0.47**	**0.46**	**0.47**	**0.43**	**0.49**
Height G-lipid	**0.57**	-0.07	**0.40**	**0.85**	**0.84**	-0.03	-0.09	-0.07
Height Baseline	-0.14	**0.18**	-0.13	-0.42	-0.42	-0.07	-0.04	-0.04
Width LMW	-0.20	**0.26**	-0.22	-0.49	-0.49	0.10	0.05	0.03
Width Glyc-B	0.04	0.07	0.04	-0.10	-0.10	**-0.14**	-0.12	**-0.22**
Width Glyc-A	0.08	-0.01	0.06	0.09	0.09	0.08	0.05	0.06
Width G-lipid	0.11	**0.20**	**0.15**	**-0.40**	**-0.40**	0.03	-0.03	-0.08
Width Baseline	-0.03	0.07	-0.08	0.00	0.00	-0.01	0.07	0.06
Pos LMW	0.11	**-0.04**	0.12	0.09	0.09	-0.01	-0.04	0.06
pos glyc-B	**-0.29**	-0.17	**-0.19**	**-0.10**	**-0.10**	0.17	0.20	0.28
Pos Glyc-A	**0.14**	-0.04	**0.11**	**0.22**	**0.22**	-0.25	-0.16	**-0.27**
Pos G-lipid	**0.30**	-0.52	**0.39**	**0.77**	**0.77**	-0.02	0.00	-0.01
Pos Baseline	0.02	0.02	0.03	-0.07	-0.06	-0.01	-0.03	0.03
Total area	**0.56**	-0.02	**0.40**	**0.75**	**0.74**	**0.13**	0.06	0.11
H/W Glyc-B	0.08	-0.02	0.07	0.08	0.08	**0.40**	**0.36**	**0.46**
H/W Glyc-A	0.03	**-0.14**	0.00	**0.35**	**0.34**	**0.32**	**0.30**	**0.36**
HXW Glyc-B	0.07	0.12	0.05	-0.13	-0.13	**0.16**	**0.14**	0.08
HXW Glyc-A	0.03	**-0.14**	0.00	**0.35**	**0.34**	**0.32**	**0.30**	**0.36**
Distance GlycB-GlycA	**-0.32**	**-0.14**	**-0.22**	**-0.18**	**-0.19**	**0.26**	**0.25**	**0.37**
Distance GlycB-Glipid	**-0.38**	**0.45**	**-0.43**	**-0.78**	**-0.78**	0.07	0.06	0.09
Distance GlycA-Glipid	**-0.29**	**0.52**	**-0.38**	**-0.76**	**-0.75**	-0.01	-0.02	-0.03

Table 1: Correlation coefficients for each of the variables. Significant values ($p<0.05$) are underlined and bold.

The multivariate study showed that including glycoproteins to traditional inflammation parameters improved the activity classification and severity of RA being the area under the cross validated ROC curve 0.79 (Figure 4) while only with the traditional inflammatory markers was 0.74 (Figure 5).

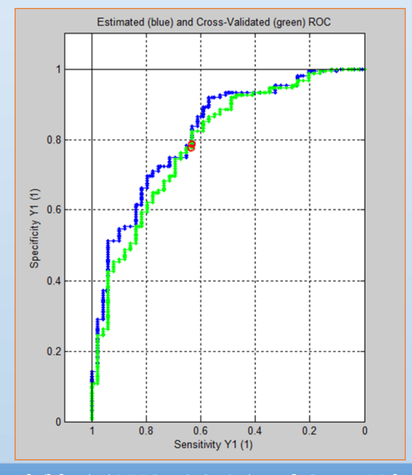


Figure 4: Estimated (blue) (AUC= 0.8191) and Cross-Validated (green) (AUC= 0.7921) ROC curve of prediction model for DAS28 including 13 ^{1}H-NMR selected variables.

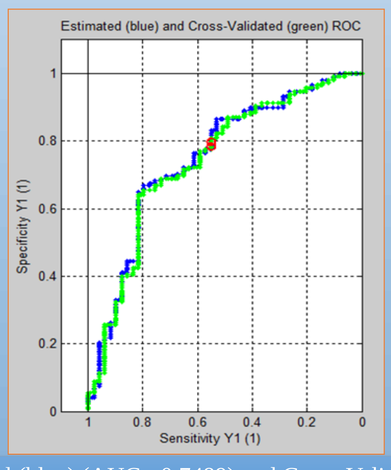


Figure 5: Estimated (blue) (AUC= 0.7489) and Cross-Validated (green) (AUC= 0.7360) ROC curve of prediction model for DAS28 including the 3 traditional inflammatory markers (ESR, PCR and Fibrinogen).

Conclusions

The ^{1}H-NMR, is a useful technique to identify an atherogenic / pro-inflammatory profile in patients with RA. The glycoprotein and lipoprotein profiles extracted from ^{1}H-NMR spectra, along with the classic inflammatory parameters, provide more accurate information about the cardiometabolic status of RA disease as well as severity and activity of the disease.

References

[1] Daniel H. Solomon, Elizabeth W. Karlson, Eric B. Rimm, Carolyn C. Cannuscio, Lisa A. Mandl, JoAnn E. Manson, Meir J. Stampfer, Gary C. Curhan *"Cardiovascular Morbidity and Mortality in Women Diagnosed With Rheumatoid Arthritis"* Circulation. 2003;107:1303-1307 Originally published March 11, 2003

[2] Akinkuolie, A. O., Buring, J. E., Ridker, P. M., & Mora, S. *"A novel protein glycan biomarker and future cardiovascular disease events"* Journal of the American Heart Association, 3(5), e001221. 2014.

[3] R. Mallol, N. Amigó, M. A. Rodríguez, M. Heras, M. Vinaixa, N. Plana, E. Rock, J. Ribalta, O. Yanes, L. Masana, and X. Correig. *"Liposcale: a novel advanced lipoprotein test based on 2D diffusion-ordered 1H NMR spectroscopy.,"* J. Lipid Res., vol. 56, no. 3, pp. 737–46, Mar. 2015.

With thanks to the EAS for support in the form of a Young Investigator Fellowship.