Comparative glycosylation profiles in Rheumatoid Arthritis

Principle results of the K-119459 NKFI Project

The main objective of the research project was the study of changes in protein glycosylation in the case of Rheumatoid Arthritis (RA). We have planned, and performed a comparative study of protein glycosylation, in the case of various proteins. We have also compared protein glycosylation in healthy individuals and in RA patients. We have found glycosylation markers, which clearly differentiate Rheumatoid Arthritis. In the course of research we have shifted focus towards glycosylation of immunoglobulin G (IgG), although glycosylation of other proteins (in accordance with the original Workplan) have also been studied. The main findings of the project are discussed in the subsequent paragraphs. Due to the Covid disease research has been slowed, and several major publications are still in preparation (these will include various key findings listed below).

1) We have adapted and **developed various techniques for sample preparation and analysis of protein glycosylation**. These include sample enrichment procedures, HPLC-MS/MS analysis, optimization of collision energy for glycopeptide MS/MS analysis and development of data evaluation processes. Some of these optimized processes have been published, some others are included among the research protocols of the ELKH Research Centre for Natural Sciences MS Proteomics research group. The method adapted for the present research, including the optimized procedures, consists of the following main steps:

- a1) Enrichment of glycoproteins from sera
- a2) Isolation of IgG form sera using column chromatography
- a3) Isolation of the antiACPA fraction of IgG using immunoaffinity chromatography
- b) Tryptic digestion of the samples (very small sample amounts, corresponding to ca. 20-50 mL) and sample preparation for LC-MS analysis
- c) nanoLC-MS and nanoLC-MS/MS analysis of the samples. This type of experiments is useful for identification and quantitation of protein- and site-specific N-glycosylation. In particular, IgG subclasses (IgG1, IgG2, and a mixture of IgG3 and IgG4) can also be separated and individually analyzed.
- d) Identification of glycosylation of various proteins using MS/MS analysis
- e) Quantitation of various glycopeptide isoforms using MS analysis
- f) Comparative, statistical analysis of glycosylation

Note, sample preparation steps a1, a2, and a3 are alternatives of each other. On many samples all three methods have been used, and the resulting samples were studied subsequently.

2) **Glycoproteins in the sera** of healthy individuals and RA patients have been identified, and their glycosylation features studied. These are listed in Table 1. Based on preliminary analysis and consultation with immunologists, we have decided to focus on glycosylation changes in immunoglobulins, and in particular, IgG.

Uniprot ID	Glycoprotein
P02763	AGP 1
P01011	Alpha-1-antichymotrypsin
P01023	Alpha-2-macroglobulin
P05090	Apolipoprotein D
P02749	Beta-2-glycoprotein 1
P01024	Complement C3
P0C0L4	Complement C4
P08603	Complement Factor H
P02765	Fetuin-A
P02751	Fibronectin
P00738	Haptoglobin
P02790	Hemopexin
P01876	IgA1
P0DOX2	IgA2
P0DOX5	IgG1
P01859	IgG2
P01861	IgG4
P01591	IgJ
P01871	IgM
P06681	Isoform 2 of Complement C2

Table 1. List of glycoproteins studied

3) **Glycosylation of Immunoglobulin G**.

Glycopeptides derived from IgG subclasses can be distinguished based on a combination of molecular mass and chromatographic retention times of the respective glycopeptides. However, those of IgG3 and 4 are isobars, and cannot be distinguished this way. This yields results on the glycosylation of IgG1, IgG2, and a mixture of IgG3 and 4. Altogether ca. 20 different glycoforms have been identified in each IgG subclass using MS/MS analysis. The identified glycoforms have been quantified in ca. 100 sera samples; some taken from healthy persons, some from RA patients. In some cases glycosylation of the total IgG, and glycosylation of antiACPA-specific fraction of IgG have also been compared. The average glycosylation profiles are shown in Table 2.

The core glycoform consists of 4 N-acetylglucosamine and 3 mannose units, its composition abbreviated as N4H3, and often indicated as the G0 glycoform. This can be galactosylated (G1, G2), and the galactose units might be sialylated (e.g. G2S1). These glycoforms are most often fucosylated (e.g. G1F), and might contain an N-acetylglucosamine at the 'bisecting' position (e.g. G2FB). Galactosylation, sialylation, fucosylation and process of adding a bisecting N-acetylglucosamine are often characterized by glycosylation indices (Table 3), which can be determined from the relative abundance of the various glycoforms.

abbreviation	composition	relative abundance

		IgG1	IgG2	IgG3/4
M1F	N3H4F1	0.6	0.4	
M1SF	N3H4S1F1	0.1	0.2	
G0	N4H3	1.7	0.8	
G0F	N4H3F1	22.0	28.7	29.8
G1	N4H4	3.1	1.0	
G1F	N4H4F1	31.5	30.7	28.3
G1S	N4H4S1		0.1	
G1SF	N4H4S1F1	1.9	5.6	4.5
G2	N4H5	1.7	0.5	
G2F	N4H5F1	13.8	9.9	10.6
G2S	N4H5S1	0.7	0.2	
G2SF	N4H5S1F1	7.5	6.3	8.7
G2S2F	N4H5S2F1	0.1		
G0B	N5H3	0.4	0.2	
G0BF	N5H3F1	5.2	7.8	9.5
G1B	N5H4	0.7		
G1BF	N5H4F1	7.1	5.8	6.5
G1BSF	N5H4S1F1	0.2	0.2	0.4
G2BF	N5H5F1	1.4	1.3	1.5
G2BSF	N5H5S1F1	0.2	0.2	

Table 2. Average glycosylation pattern of each of the IgG subclasses.

Index	IgG1	IgG2	IgG3/4
G0 (%)	29.3	37.5	39.3
G1 (%)	44.5	43.5	39.8
G2 (%)	25.5	18.3	20.9
G average (%)	95.5	80.2	81.6
defucosylated (%)	8.3	2.8	0.0
sialylated (%)	10.6	12.7	13.7
bisecting (%)	15.3	15.5	17.9

Table 3. Average glycosylation indices of IgG subclasses.

4) **Comparison of glycosylation between healthy persons and RA patients**

We have compared glycosylation features in the two groups, in schematic form this is shown in Figure 1. This suggests two main differences between the two groups, in the relative abundance of IgG subclasses and in fucosylation. In healthy persons IgG1 is the most abundant. RA patients show a decreased amount of IgG1 and a significantly increased amount of IgG3/4 subclasses. Even more interesting and important are differences in fucosylation. Compared to the healthy population, IgG1 has a smaller, while IgG2 has a larger amount of defucosylated glycoforms in RA patients. Note, the average amount of IgG fucosylation changes only very little. Most studies of IgG glycosylation do not study subclasses

separately, so this difference have not been observed before. These results suggest that glycosylation markers might be good markers for Rheumatoid Arthritis.

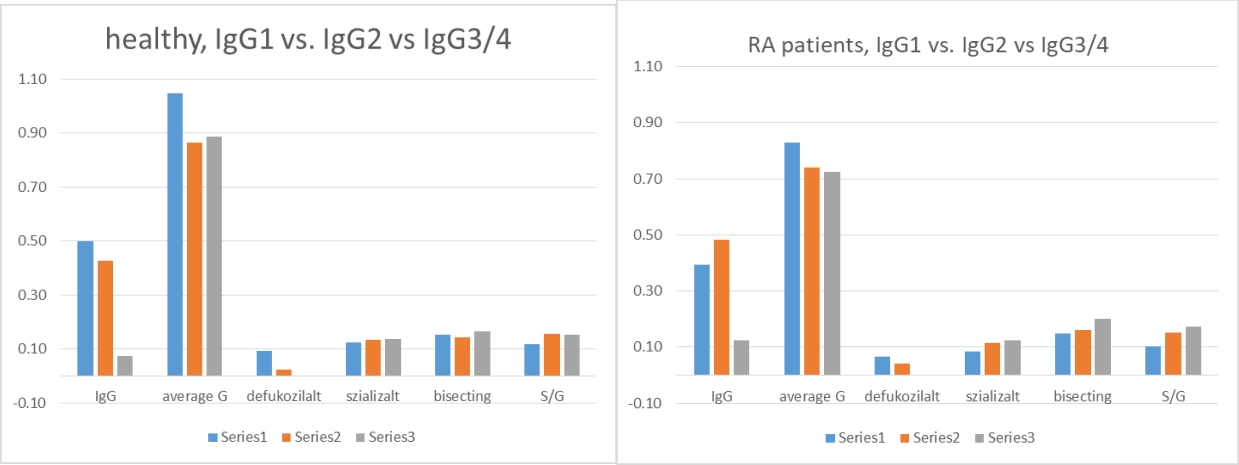

Figure 1: Comparison of the main glycosylation features between healthy persons and RA patients. IgG1 (blue), IgG2 (orange), IgG3/4 (gray).

5) **Detailed correlations among glycosylation features and other parameters**

We have compared correlations among various glycoform abundances, glycosylation indices and medical descriptors for all studied persons. Some of the most important correlations observed are the following:

a) ***Galactosylation decreases significantly with age***, both in the case of healthy persons and RA patients (Fig.2). This has a major impact on studies on RA, and makes it imperative that healthy persons and RA patients should be in the same age-group. In the comparisons between healthy and RA groups discussed below we have selected only women in the 50-70 age bracket and, adjusted the two groups so that the average age in the two studied groups should be within one year (58.1 and 58.6, respectively). Looking over prior studies on RA and glycosylation, we have found that in most studies the average age of the healthy volunteers and RA patients were widely different. We strongly suspect that such unbiased selection of volunteers invalidates many published results on the effect of RA on the glycosylation of IgG, which are currently widely accepted.

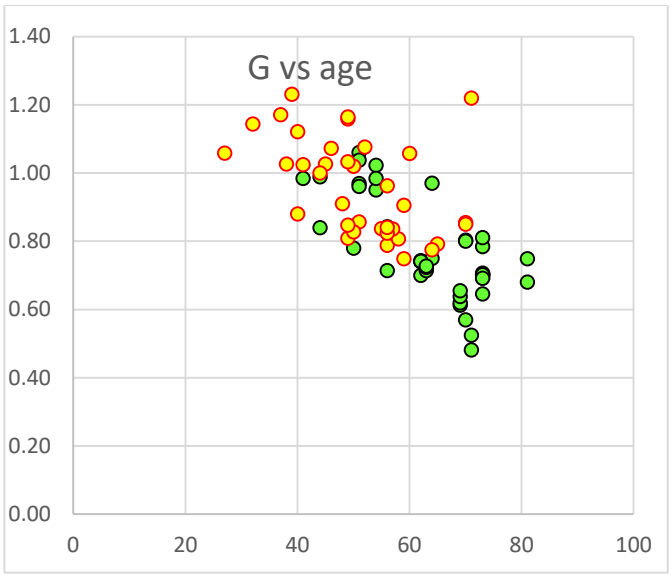

Fig. 2. Average number of galactose units in IgG in the case of various individuals as a function of age (orange: healthy, green: RA)

b) ***Galactosylation and sialylation shows a very good linear correlation*** (R^2=0.72). Data relating to healthy individuals and RA patients fall on the same trendline, and this is irrespective of age. The ratio of sialic acid and galactose units is a good measure of the degree of sialylation. Considering all results, between 10-20% of galactose units are sialylated (Fig.3). In the case when individuals show a high degree of galactosylation, these are slightly more probable to get sialylated. Low degree of galactosylation, on the other hand, corresponds to a slightly lower degree of sialylation. Note that galactosylation changes significantly among the studied population. The degree of sialylation (fraction of sialylated galactose units), on the other hand, changes less, and the difference between healthy and RA patients is smaller.

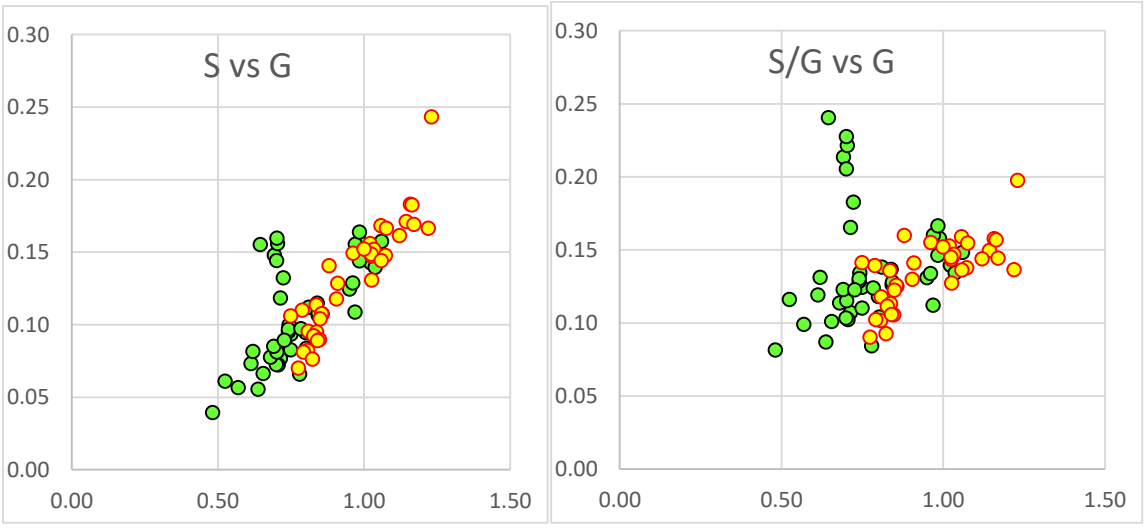

Fig. 3: Correlation between galactosylation and sialylation, and the degree of sialylation (S/G) plotted against galactosylation (orange: healthy, green: RA). Blue and purple dots correspond to

c) ***Correlations among glycosylation features***: We have compared galactosylation and fucosylation; galactosylation and bisecting glycans; and fucosylation and bisecting glycans (Fig.4). These features do not show any correlation; i.e. are independent features.

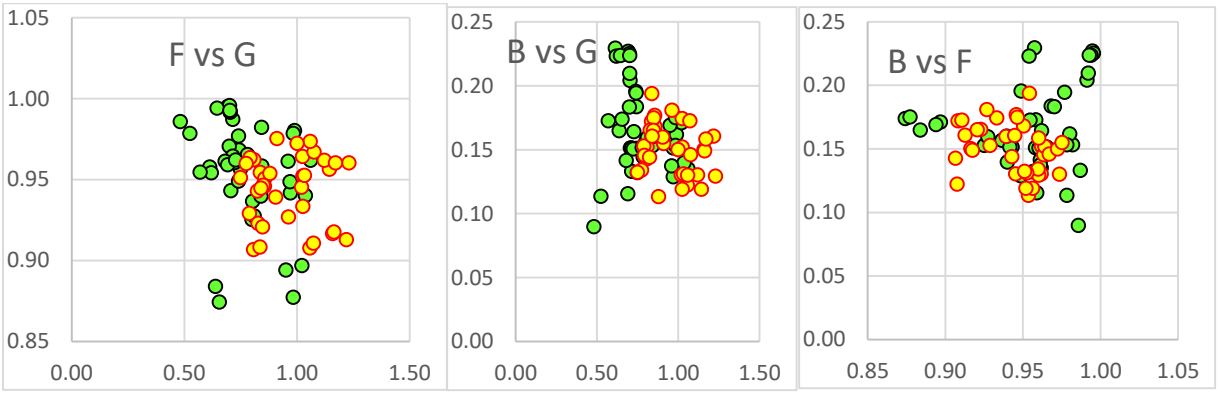

Fig. 4: Galactosylation and fucosylation; galactosylation and bisecting glycans; and fucosylation and bisecting glycans (orange: healthy, green: RA)

d) ***Influence of age and disease on the proportion of IgG variants***. These are shown in Fig. 5, indicating the proportion of the IgG variant against age, for the healthy and RA group. The results show a significant increase of the relative amount of IgG3/4 both with age and with RA; although it is no clear correlation, like in the case of galactosylation. Note that RA patients show an elevated proportion of IgG3/4 in each age bracket (e.g. 40-50; 50-60, or 60-70). This may suggest that the proportion of increased IgG3/4 might be an indication of poor health.

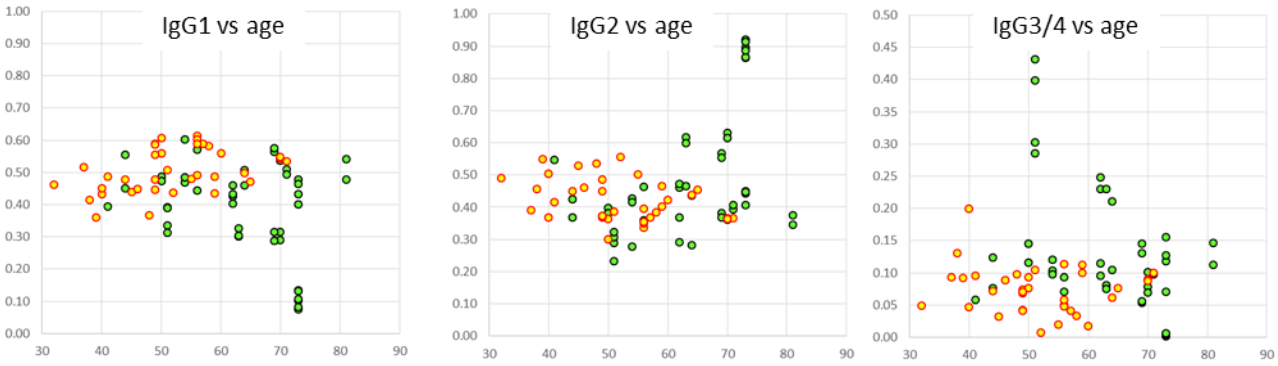

Fig. 5 Proportion of IgG subclasses correlated with age and RA (orange: healthy, green: RA)

e) ***Differential fucosylation of IgG1 and IgG2 subclasses***. It is one of the most unexpected, and potentially very important finding of the present study (Fig. 6). Fucosylation of IgG1 and IgG2 do not

depend on age. In the case of IgG1, the degree of fucosylation is slightly increased in RA. IgG2, on the other hand, shows a significant decrease of fucosylation in RA, although the inter-individual differences are large. The Figure also shows the cross-correlation between the degree of fucosylation of IgG1 and IgG2. In both the healthy and the RA group the results show a good linear correlation between the degree of fucosylation on IgG1 and IgG2. However, the slope of this correlation is markedly different in the healthy and RA group of subjects. The linear correlation suggests that if a person is characterized by low degree of fucosylation, both IgG1 and IgG2 will be under-fucosylated (compared to the average). Note that differences between healthy and RA subjects can be grouped together fairly well based on fucosylation only: all healthy persons fall inside the ellipsis in Fig. 6, while data of nearly all RA persons fall outside it. Note furthermore, that this is a feature which does not depend on the age of the person. The most important conclusion is, that RA influences fucosylation of IgG1 and IgG2 in a different manner.

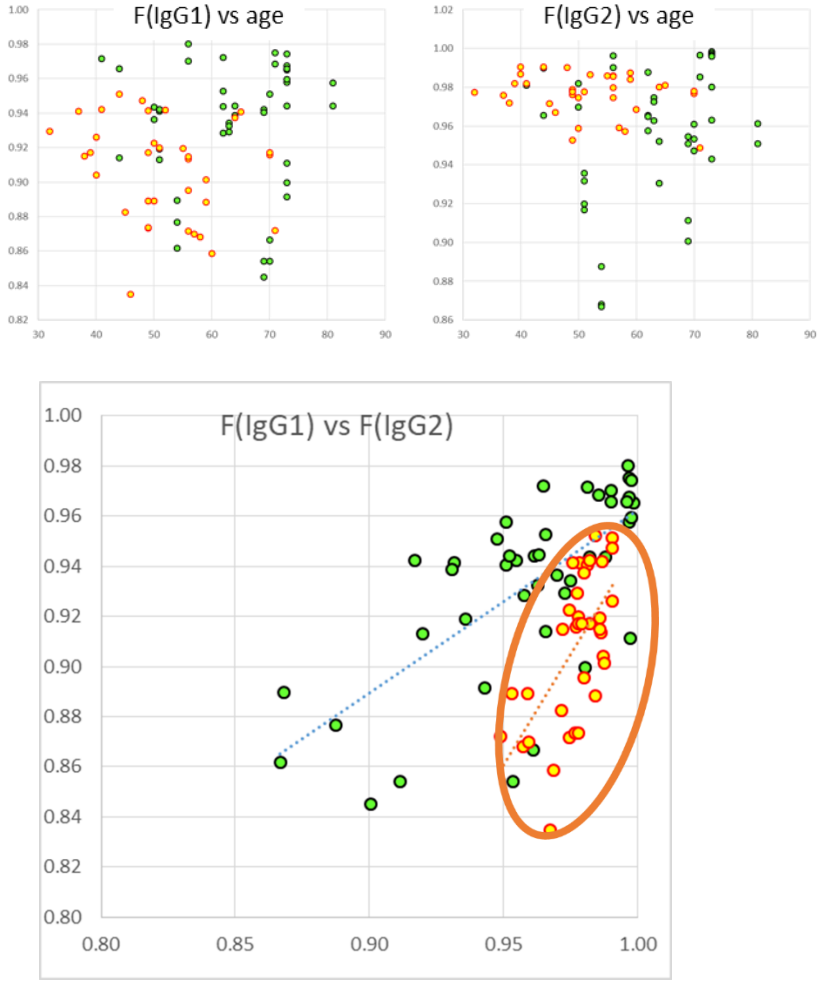

Fig. 6 Degree of fucosylation of IgG1 and IgG2 subclasses, and cross-correlation between IgG1 and IgG2 fucosylation (orange: healthy, green: RA)

f) ***Correspondence between glycosylation features and medical (laboratory) results of RA patients.*** We have compared IgG glycosylation features with disease activity score (DAS), C-reactive protein (CRP), rheumatoid factor (RF), cyclic citrullinated peptide (CCP) and Erythrocyte sedimentation rate (ESR). Reasonable linear correlations have been found between galactosylation and CRP (R^2=0.39), galactosylation and RF (R^2=0.42), and also between galactosylation and ESR (R^2=0.31) (Fig. 7.). Note that low galactosylation corresponds to high values of CRP, RF and ESR. There is one further case showing reasonable correlation between glycosylation and a disease marker, and that is the amount of bisecting glycoforms and RF (R^2=0.29, Fig. 7): A high degree of bisecting glycoforms correlate with high RF values.

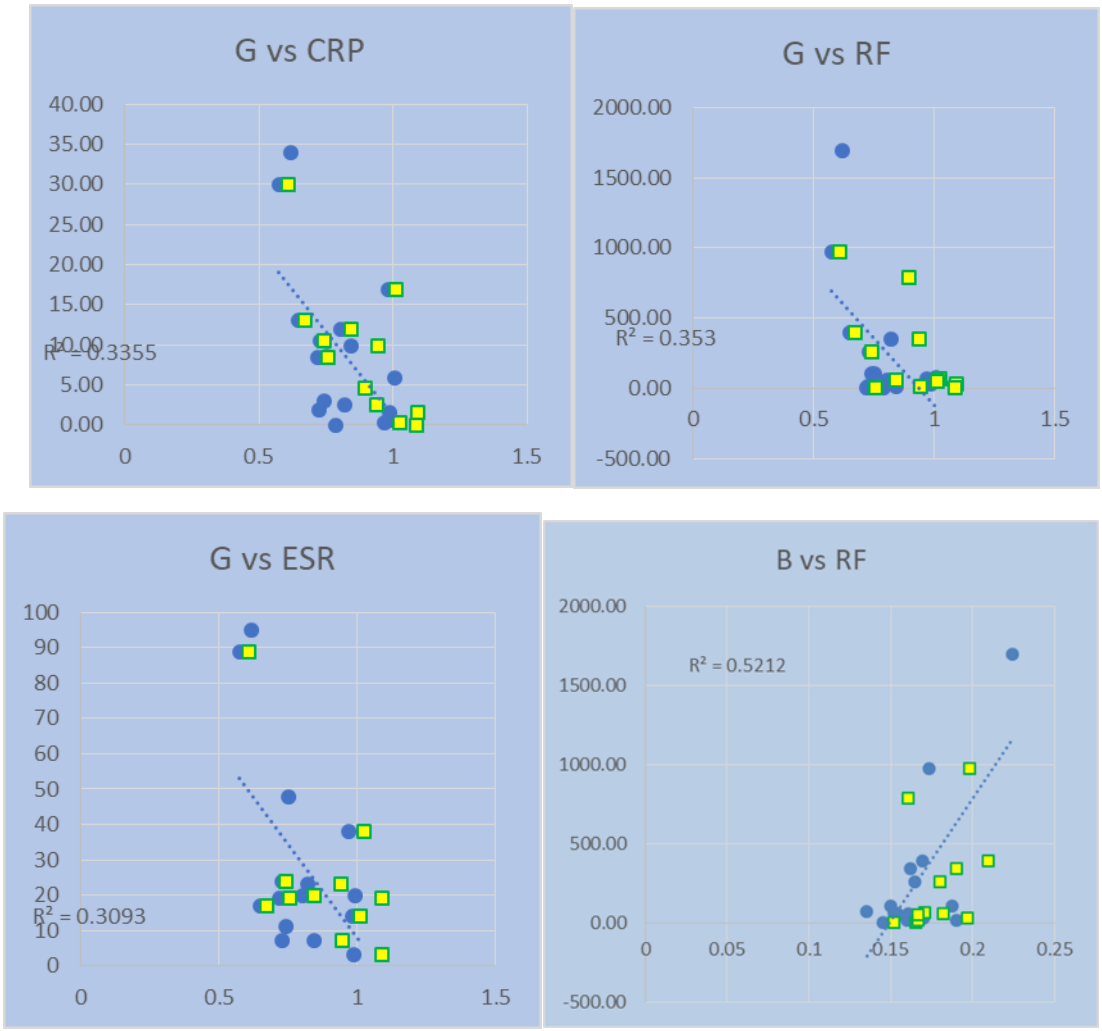

Fig. 7: Correlations between glycosylation features and medical laboratory results

5) Glycosylation-based biomarkers for diagnosis of RA patients.

The results discussed above indicate various differences between the glycosylation of IgG subclasses and rheumatoid arthritis. We have done statistical analysis, among them principle component analysis (PCA)

and linear discriminant analysis (LDA) to test the capability of glycosylation markers to separate healthy persons and RA patients. Fig. 8 shows a typical LDA plot, indicating that glycosylation might become a good diagnostic marker for RA. In this study 4 glycan abundances were selected, N3H4SF of IgG1; N4H5 of IgG1, N4H5F of IgG1 and N4H5 of IgG2. We have done several checks to cross validate the results, and have shown that the model is independent of the selection of the training set. In this trial IgG glycosylation patterns of 28 RA patients and 34 healthy individuals were studied. The two classes were age matched, and all participants were women. The cross-validated results show 100% probability of identifying healthy persons, while 93% of correctly identifying RA patients.

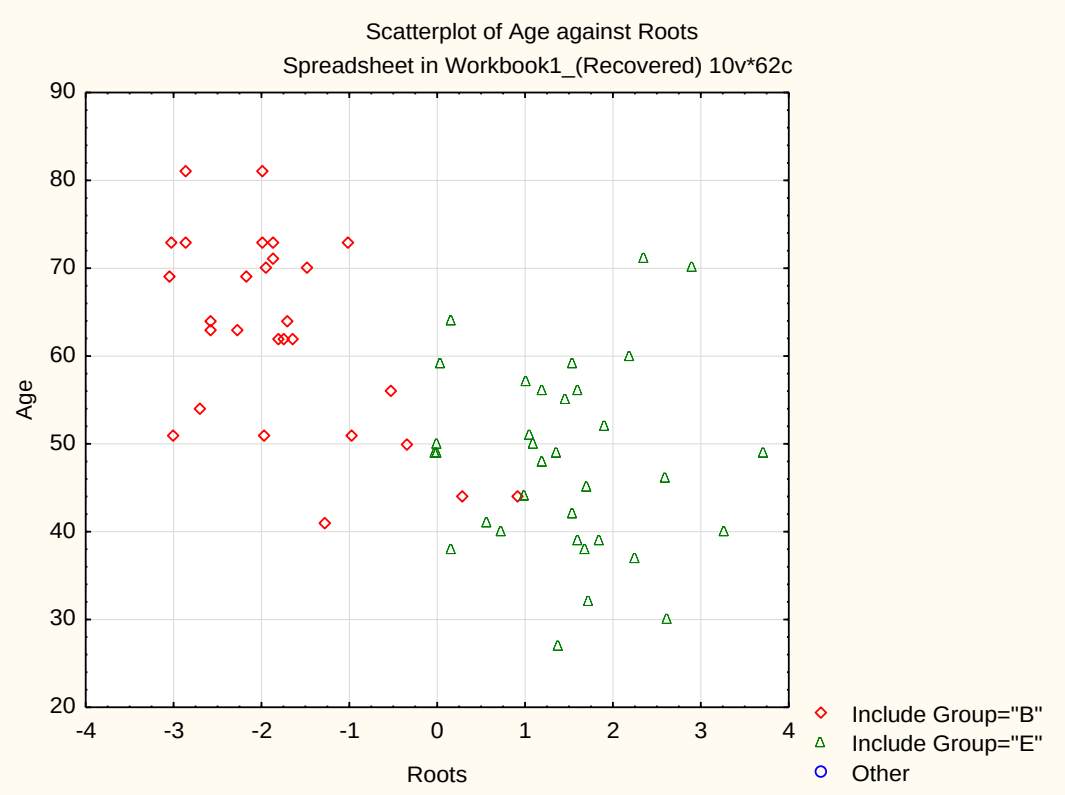

Figure 8. LDA plot of healthy persons (green) and RA patients (red)

6) Role of glycosylation on the pathomechanism of RA

We have performed functional assays, like capacity of TNFα production, in order to study pro-inflammatory activities. These show that ACPA-active IgG (present in RA patients) significantly enhance TNFα release on an FcγRI dependent manner, while healthy IgG does not. TNFα production inversely correlate with the relative abundance of the (agalactosylated) G0 glycoform. These suggest a novel mechanism in RA, contributing to disease progression.