

Direct detection of humoral marker correlates of COVID-19, glycated HSA and hyperglycosylated IgG3, by MALDI-ToF mass spectrometry.

June 2021

Ray K Iles^{1,2,3*}, Jason K Iles^{1,2}, Raminta Zmuidinaite^{1,2}, Anna Gardiner¹ Jonathan Lacey¹, Stephen Harding⁴ Jonathan Heeney^{2,5}, Helen Baxendale⁶

¹MAPSciences, The iLab, Stannard Way, Bedford, MK44 3RZ, UK

²Laboratory of Viral Zoonotics, Department of Veterinary Medicine, University of Cambridge, Madingley Road, CB3 0ES, Cambridge, UK.

³NISAD, Sundstorget 2, 252-21 HELSINGBORG, Sweden

⁴The Binding Site Group Ltd, 8 Calthorpe Road, Edgbaston, Birmingham B15 1QT, UK

⁵DIOSynVax, University of Cambridge, Madingley Road, CB3 0ES, Cambridge, UK

⁶Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK

*corresponding author

Short title: Direct and rapid MALDI-ToF MS screening for glycated/glycovariant HSA and IgG3 in COVID-19.

Keywords: COVID-19, MALDI-ToF MS, plasma, glycated Albumin, Glycovariant/glycated IgG3

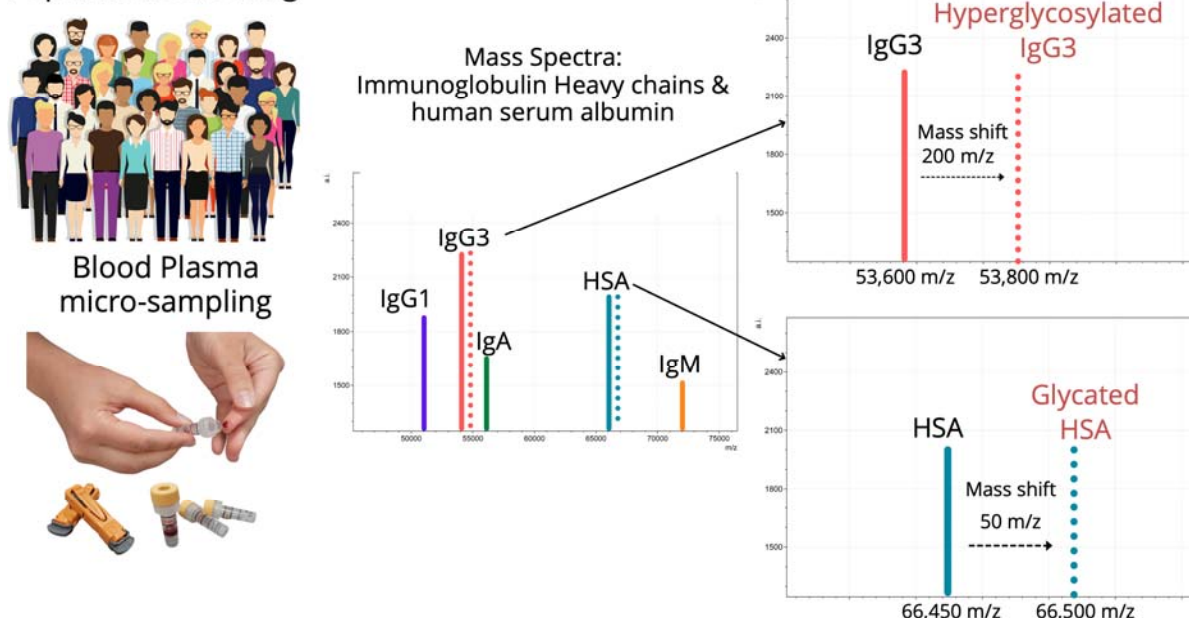
Abbreviations: HCW = Health care workers, ARDS= Acute respiratory distress syndrome, ITU= Intensive Therapy unit, AGE= advanced glycation end products, MALDI-ToF = Matrix assisted laser desorption ionization – time of flight. HSA= Human serum albumin, IgG3 = Immunoglobulin G subtype 3.

Abstract

The prefusion Spike protein of SARS-CoV2 binds advanced glycation end product (AGE) glycated human serum albumin (HSA) and a higher mass, hyperglycosylated/glycated, IgG3, as determined by matrix assisted laser desorption mass spectrometry (MALDI-ToF MS). We set out to investigate if the total blood plasma of patients who had recovered from acute respiratory distress as a result of COVID-19, contained more glycated HSA and higher mass (glycosylated/glycated) IgG3 than those with only clinically mild or asymptomatic infections. A direct dilution and disulphide bond reduction method was development and applied to plasma samples from SARS-CoV2 seronegative (N = 30) and seropositive (N = 31) healthcare workers and 38 convalescent plasma samples from patients who had been admitted with acute respiratory distress syndrome (ARDS) associated with COVID-19.

Patients recovering from COVID-19 ARDS had significantly higher mass, AGE-glycated HSA and higher mass IgG3 levels. This would indicate that increased levels and/or ratios of hyperglycosylation (probably terminal sialic acid) IgG3 and AGE glycated HSA may be predisposition markers for development of ARDS as a result of COVID-19 infection. Furthermore, rapid direct analysis of plasma samples by MALDI-ToF MS for such humoral immune correlates of COVID-19 presents a feasible screening technology for the most at risk; regardless of age or known health conditions.

Population screening



Introduction

As the COVID-19 pandemic has gripped the world, the medical technology strategy to deal with the dynamic situation has evolved to include mass screening and mass vaccination, additional to extensive investigations to identify individuals at risk of severe disease. Mass screening has proved challenging given the logistics involved and need for affordable, easy, repeat testing, track and trace [1]. Vaccination has made a huge impact on infection transmission and hospitalisations in countries where over 70% of the adult population has been immunised [2]. Booster vaccinations are likely needed to deal with emerging variant forms of SARS CoV-2 is a new long-term realisation [3]. A huge amount of money, time and national resources have been invested in medical technologies to meet the ongoing demand for diagnostics. Given that the infection with SARS-CoV2 virus was life threatening to minority of the population [4,5] and that vaccination may reduce this risk further; there are growing calls to open back up social interaction, business, trade and international travel [6]. How this can be justified remains an ethical and global-social/political dilemma. In response to this changing focus, clinicians and scientist are seeking to identify those who are at the greatest risk of severe disease to help inform and rationalise targeted testing and booster vaccinations [7].

A range of medical technologies are being explored and developed to help inform this targeted or precision medicine approach. Currently, the focus has been on artificial intelligence (AI) identifying the “at risk” from medical notes and pre-existing test results via algorithms trained through machine learning [8,9,10]. However, whether this will be sufficient to give the level of, sensitivity and specificity, “precision” required in identifying all that are at risk is questionable. The fault will lay not with the machine learning (ML) or AI, but that insufficient robust correlating test results are being provided to the algorithms, where the diagnostic accuracy of each of the correlate will be amplified by ML. New significant and non-subjective evaluations will be required such as biomarkers and immune correlates [11].

We have recently shown that the prefusion spike protein of SARS-Cov2 binds glycated albumin and a higher molecular mass, due to glycation or hyper glycosylation, IgG3. This we proposed benefits the virus immune evasion strategies and may contribute to COVID-19 pathologies as a result of a selective cloud of serum proteins triggering thrombolytic and detrimental inflammatory pathways [12,13]. Furthermore, an increased ratio of IgG3 compared to IgG1 in the humoral response to SARS-Cov-2 is associated with an adverse outcome of a COVID-19 infection [14]. The possibility is that elevated glycated albumin along with elevated levels of IgG3 characterised by an increased molecular mass (due to glycation/glycosylation) may act as biomarkers of susceptibility to develop ARDS as a result of SARS-CoV2 infection.

We examined the mass spectra of samples from convalescent patients and sero-negative health care workers (HCW), without immunocapture, to see if direct MALDI-ToF of diluted and disulphide reduced blood plasma would reveal gross characteristic differences in resolved immunoglobulin heavy chains and (glycated) albumin spectra. Given the logistic of simple pin prick blood home sampling, high throughput low cost MALDI-ToF mass spectrometry could provide one of the marker tests for the “at risk” population.

Material and Methods

Samples

Serum and plasma samples were obtained from Health care workers (HCWs) and patients referred to the Royal Papworth Hospital for critical care. COVID-19 patients hospitalised during the first wave and as well as NHS healthcare workers working at the Royal Papworth Hospital in Cambridge, UK served as the exposed HCW cohort (Study approved by Research Ethics Committee Wales, IRAS: 96194 12/WA/0148. Amendment 5). NHS HCW participants from the Royal Papworth Hospital were recruited through staff email over the course of 2 months (20th April 2020-10th June 2020) as part of a prospective study to establish seroprevalence and immune correlates of protective immunity to SARS-CoV-2. Patients were recruited in convalescence either pre-discharge or at the first post-discharge clinical review. All participants provided written, informed consent prior to enrolment in the study. Sera from NHS HCW and patients were collected between July and September 2020, approximately 3 months after they were enrolled in the study.

For cross-sectional comparison, representative convalescent serum and plasma samples from seronegative HCWs, seropositive HCW and convalescent PCR-positive COVID-19 patients were obtained. The serological screening to classify convalescent HCW as positive or negative was done according to the results provided by a CE-validated Luminex assay detecting N-, RBD- and S-specific IgG, a lateral flow diagnostic test (IgG/IgM) and an Electro-chemiluminescence assay (ECLIA) detecting N- and S-specific IgG. Any sample that produced a positive result by any of these assays was classified as positive. Thus, the panel of convalescent plasma samples (3 months post-infection) were grouped in three categories: A) Seronegative Staff (N = 30 samples) B) Seropositive Staff (N = 31 samples); C) Patients (N = 38 samples) [15].

Sample Analysis by MALDI-ToF Mass spectrometry

Mass spectra were generated using a 15mg/ml concentration of sinapinic acid (SA) matrix. One µl of plasma was diluted in 40µl of ultra-pure mass spec grade water (Romil Ltd, Water beach Cambridgeshire UK). After thorough vortex mixing, 10µl of the diluted plasma was mixed with mixed 1:1 with 10µl of 20mM tris(2-carboxyethyl)phosphine (TCEP) (Sigma-Aldrich, UK) in ultra-pure water. After incubation for 15mins at room temperature 1µl of the diluted samples were taken and plated on a 96 well stainless-steel target plate using a sandwich technique [16]. The MALDI-ToF mass spectrometer (microflex® LT/SH, Bruker, Coventry, UK) was calibrated

using a 2-point calibration of 2mg/ml bovine serum albumin (2+ ion 33,200 m/z and 1+ion 66,400 m/z) (PierceTM, ThermoFisher Scientific). Mass spectral data were generated in a positive linear mode. The laser power was set at 65% and the spectra was generated at a mass range between 10,000 to 200,000 m/z; pulsed extraction set to 1400ns.

A square raster pattern consisting of 15 shots and 500 positions per sample was used to give 7500 total profiles per sample. An average of these profiles was generated for each sample, giving a reliable and accurate representation of the sample across the well. The raw, averaged spectral data was then exported in a text file format to undergo further mathematical analysis.

Spectral Data processing

Mass spectral data generated by the MALDI-ToF instrument were uploaded to an open-source mass spectrometry analysis software mMassTM [17], where it was processed by using; a single cycle, Gaussian smoothing method with a window size of 300 m/z, and baseline correction with applicable precision and relative offset depending on the baseline of each individual spectra. In the software, an automated peak-picking was applied to produce peak list which was then tabulated and used in subsequent statistical analysis.

Statistical analysis

Peak mass and peak intensities were tabulated in excel and plotted in graphic comparisons of distributions for each antigen capture and patient sample group. Means and Medians were calculated and, given the asymmetric distributions found, non-parametric statistics were applied, such as Mann Whitney U test, when comparing differences in group distributions.

Results

The developed direct plasma sample MALDI-ToF mass spectra analysis method was rapid and required minimal reagents, optimised pulse extraction allowed for the IgG subtypes to be resolved in most samples. From previous spectra calibration using pure preparation IgG1 heavy chains averaged at peak apex ~51,000m/z, IgG3 at ~54,000m/z, IgA at ~56,000m/z and IgM at ~74,000m/z. Human Albumin was at 66,400m/z (1+) and Transferrin at 79,600m/z. All are broad heterogenous peaks reflecting glycosylation/glycation & sequence variation, and were

identifiable in the plasma sample spectra (see figure 1). A further as yet to be confirmed Ig heavy chain (IgX), presumed to be IgG4, was often seen resolving at $\sim 48,000m/z$ [12,13].

Direct plasma total IgG1 spectra

IgG1Hc was detected in all samples and was the most dominant and abundant immunoglobulin such that less abundant IgG3 and the related Ig heavy chain termed IgX Hc were often masked under the broad polyclonal IgG1 peak. Although intensity levels varied no significant association with the COVID-19 sample sets was found. Similarly, variance in IgG1 average molecular mass, although slightly lower for the ARDS COVID-19 patient set, was not statistically significant: Mean difference $\Delta 31m/z$, -0.06% change in mass, not significant $p=0.177$. (See figure 2)

Direct plasma total IgG3 spectra

IgG3 Hc mass spectral peaks were detected in 43/99 samples and showed an increasing intensity in those who had been infected with SARS-CoV2, with highest levels being found in the ARDS-COVID-19 patients' convalescent plasma: In the sero-negative HCW samples, peak detected in 37% of samples, mean intensity 30AU; sero-positive HCW samples, peak detected in 32% of samples, mean intensity 27AU and COVID-19 ARDS patient samples peak detected in 58% of samples, mean intensity 53AU. The molecular mass of the total IgG3 Hc in the plasma samples of the ARDS-COVID-19 patient convalescent plasma was significantly higher than the other groups (Mean difference: $\Delta 200m/z$, +0.37% change in mass, $p=0.00236$) (see figure 2)

Direct plasma HSA spectra

Human Albumin was detected in all samples and no statistically significant variance in albumin Intensity could be found in these post infection convalescence plasma samples. However, the molecular mass of the convalescent plasma HSA in the patients who were recovering from ARDS as a result of SARS-CoV2 was statistically significantly larger than total HSA found in HCWs who had recovered from SARS-COV2 infection with only mild symptoms: Mean difference $\Delta 50m/z$, +0.08% change in mass, $p=0.00758$).

Discussion

That the total HSA in patients who had developed ARDS COVID-19 had an average molecular weight higher than those developing only mild symptoms, is a significant clinical biomarker finding. But not entirely surprising, given that the elderly and particularly those with diabetes are most at risk [18]; and our previous study that the prefusion complete Spike protein binds AGE/glycated albumin [12]. Thus, it is entirely consistent with the finding of higher molecular weight glycated albumins are at such a high proportion in unextracted plasma that it skews the average mass of HSA measured, in those who had developed ARDS as a result of COVID-19.

More surprising was that, when detected, the MALDI-ToF mass spectra of total IgG3Hc was also significantly higher in the COVID-19 ARDS convalescent plasma (see figure 2). It had been reported that a dominance of anti-SARS-CoV2 IgG3 occurs in the immune responses to SARS-CoV2 in those who died from COVID-19 [14]. We had shown an increase in the mass of IgG3Hc bound to SARS-CoV2 prefusion Spike protein and eluted from magnetic bead to which it was conjugated [13]. However, to find a dominance of this higher molecular weight IgG3Hc in total plasma from patients who had recovered from COVID-19 ARDS indicates this is probably a characteristic of their inherent immunoglobulin synthesis, i.e., hyper glycosylation [19, 20]. Further to being a measurable risk biomarker, the magnitude of the mass difference at $\sim 200\text{m/z}$ is larger than that previously encountered when extracting SARS-CoV2 antigen bound antibodies (medium increase $\sim 150\text{m/z}$), and better resolved [13].

This increased resolution of a higher mass/ hyper-glycosylated total IgG3Hc, compared to when specific IgG3 capture and elution was analysed by MALDI-ToF mass spectra, could be explained by terminal glycan sugar hydrolysis and loss from the N- and O- linked moieties. Sialic acid is a common terminal structure on all glycans; however, several factors result in loss of detectable sialic acid residues including low pH and/or high temperature [21 – 25]. This is particularly common in MALDI-ToF mass spectrometry with acidic matrices (such as sinipinic acid - SA) [26,27]; and certainly, will have occurred to a large extent during our extraction of anti SARS-CoV2 Immunoglobulins from antigen bound magnetic beads with 5% acetic acid [12,13]. Indeed, we have noted that when recovering bound albumin 5% acetic acid was

sufficient to hydrolyse the N-terminal aspartic acid residue from the albumin, resulting in a mass decrease of ~115m/z compared to HSA not exposed to 5% acetic acid. Such prolonged acetic acid hydrolysis at elevated temperatures will selectively cleave at all aspartic acid residues within a protein [28].

Terminal sialylation of glycan branches have been implicated in many regulatory processes of glycoprotein including circulatory half-life [29]. Terminal sialylation of O-linked glycan of the carboxy terminal region of the beta-subunit of human chorionic gonadotropin are responsible for this hormone's increased blood circulatory half-life to 24hrs compared to its 82% homologous hormone Luteinising hormone which has a half-life of only ~20mins [30-32]. IgG3 has a reported variable half-life and is the only IgG subtype to have serine O-linked glycosylation sites within the defining neck region of the common structure 9 [33] (see figure 3). The variability and extent to which these serine residues of the IgG3 neck domain are glycosylated and terminally sialylation may be decided by other factors including a genetic preference. Similarly, IgG3 has an additional potential N-linked glycosylation site within the neck domain which may or may not be so processed [33]. Furthermore, tri-antennary rather than bi-antennary branched glycans are a common feature of variant glycosylation [34,35]. Thus, the loss of terminal sialic acids and other chain saccharides by acidic hydrolysis may mean that only residual branching sugars in extracted and purified immunoglobulin peptide chains are left to, poorly, reflect the true extent of glycan variances [24]. The direct analysis of unextracted plasma by MALDI-ToF circumvents sialic acid and other glycan losses, due to acid hydrolysis, which are an unintended consequence of purification. Thus, explaining the much clearer size differences in IgG3Hc, revealed in direct plasma sample analysis by MALDI-ToF mass spectrometry.

The importance of sialic acid residues as a target for attachment of Influenza is well known [36,37]. However, human coronaviruses OC43 and HKU1 also bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A [38]. Equivalent to the S1 subunit of SARS-CoV2 this sialic acid residue binding domain may still be partially present in this corona virus. Indeed, sialic acid binding has been postulated to be the key feature of corona virus zoonotic transfer [39]. Furthermore, so strong and specific is the spike complexes binding to sialic acid that one group has used terminally sialylated glycans as a solid phase for SARS-

CoV2 capture in a lateral flow diagnostic device [40]. It is not a large binding domain and may be a charge specific attraction. Terminal sialic acids confer a strong net negative charge, whilst other terminal glycan saccharides do not. Similarly, some of the AGE glycation product deposit a structured negative charge to albumin [41]. This may be a major component of the biochemistry of glycated HSA and hyperglycosylated (and/or glycated) IgG3 binding by SARS-CoV2 prefusion Spike protein.

The human cost of the COVID-19 pandemic has been nearly four million lives lost worldwide. The financial cost to the USA is estimated at more than \$16 trillion, or approximately 90% of the annual gross domestic product (GDP) of the US [42]. In the UK, borrowing (deficit) in 2021 grew by £300 billion, directly attributable to COVID-19 and in 2020 GDP decreased by 11.3% [43]. With estimates of gross domestic product lagging behind pre-pandemic forecast by 3% annually for many years, the final cost for the UK maybe a similar 90% of GDP. If this is reflected globally then 90% of the global GDP will put the final cost at \$76 Trillion.

Clearly the investment in medical technologies to detect the virus, prevent its infection and identify those at most risk from COVID-19, along with future viral outbreaks, is both ethically and financially justified. In this respect clinical mass spectrometry has only just started to indicate its utility [44, 45].

The potential molecular mechanisms by which sialic acid glycosylation and alike glycation contribute to the pathogenesis of COVID-19 require further investigation. Nevertheless, the direct detection of AGE glycated HSA and high mass (hyperglycosylated and or glycated) IgG3Hc, as a rapid biomarker screening test for the identification of those most at risk of developing life-threatening ARDS as a result of SAR-CoV2 infection, is a compelling possibility.

References

[1] Iacobucci G. Covid-19: Universal screening is likely to miss infected people, review finds *BMJ* 2020; 370 :m3619 doi:10.1136/bmj.m3619

- [2] García-Montero, C.; Fraile-Martínez, O.; Bravo, C.; Torres-Carranza, D.; Sanchez-Trujillo, L.; Gómez-Lahoz, A.M.; Guijarro, L.G.; García-Honduvilla, N.; Asúnsolo, A.; Bujan, J.; Monserrat, J.; Serrano, E.; Álvarez-Mon, M.; De León-Luis, J.A.; Álvarez-Mon, M.A.; Ortega, M.A. An Updated Review of SARS-CoV-2 Vaccines and the Importance of Effective Vaccination Programs in Pandemic Times. *Vaccines* **2021**, *9*, 433. <https://doi.org/10.3390/vaccines9050433>
- [3] Rubin R. COVID-19 Vaccines vs Variants—Determining How Much Immunity Is Enough. *JAMA*.2021;325(13):1241–1243. doi:10.1001/jama.2021.3370
- [4] Tomáš Sobotka, Zuzanna Brzozowska, Raya Muttarak, Kryštof Zeman, Vanessa di Lego Age, gender and COVID-19 infections medRxiv 2020.05.24.20111765; doi:<https://doi.org/10.1101/2020.05.24.20111765>
- [5] Covid-19: death rate is 0.66% and increases with age, study estimates [Mahase, Elisabeth.BMJ : British Medical Journal \(Online\); London Vol. 369, \(Apr 1, 2020\).DOI:10.1136/bmj.m1327](#)
- [6] Han, Emeline, Tan, Melisa Mei Jin, Turk, Evaqet al. Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe The Lancet, Volume 396, Issue 10261, 1525 – 1534 [https://doi.org/10.1016/S0140-6736\(20\)32007-9](https://doi.org/10.1016/S0140-6736(20)32007-9)
- [7] Booth A, Reed AB, Ponzo S, Yassaee A, Aral M, Plans D, et al. (2021) Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. *PLoS ONE* 16(3): e0247461. <https://doi.org/10.1371/journal.pone.0247461>
- [8] Clift A K, Coupland C A C, Keogh R H, Diaz-Ordaz K, Williamson E, Harrison E M et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study *BMJ* 2020; 371 :m3731 doi:10.1136/bmj.m3731
- [9] Wise J. Covid-19: Extra 1.7 million people in England are asked to shield *BMJ* 2021; 372 :n467 doi:10.1136/bmj.n467
- [10] Cave S, Whittlestone J, Nyrup R, O hEigeartaigh S, Calvo R A. Using AI ethically to tackle covid-19 *BMJ* 2021; 372 :n364 doi:10.1136/bmj.n364
- [11] Assaf, Dan, Ya'ara Gutman, Yair Neuman, Gad Segal, Sharon Amit, Shiraz Gefen-Halevi, Noya Shilo et al. "Utilization of machine-learning models to accurately predict the risk for critical COVID-19." *Internal and emergency medicine* 15, no. 8 (2020): 1435-1443.

- [12] Iles JK, Zmuidinaite, R Saddee, C et al. SARS-CoV-2 Spike protein binding of glycosylated serum albumin - its potential role in the pathogenesis of the COVID-19 clinical syndromes and bias towards individuals with pre-diabetes/type 2 diabetes & metabolic diseases. medRxiv 2021.06.14.21258871; doi:<https://doi.org/10.1101/2021.06.14.21258871>
- [13] Iles JK, Zmuidinaite, R Saddee, C et al. Determination of IgG1 and IgG3 SARS-CoV-2 spike protein and nucleocapsid binding. Who is binding who and why? medRxiv 2021.06.17.21259077; doi:<https://doi.org/10.1101/2021.06.17.21259077>
- [14] Jennifer L. Yates, Dylan J. Ehrbar, Danielle T. Hunt, *et al.* Serological analysis reveals an imbalanced IgG subclass composition associated with COVID-19 disease severity, Cell Reports Medicine, 2021, 100329, <https://doi.org/10.1016/j.xcrm.2021.100329>.
- [15] Castillo-Olivares, J., Wells, D.A., Ferrari, et al . Towards Internationally standardised humoral Immune Correlates of Protection from SARS-CoV-2 infection and COVID-19 disease. medRxiv 2021.05.21.21257572; doi:<https://doi.org/10.1101/2021.05.21.21257572>
- [16] Kussmann M, Roepstorff P. Sample preparation techniques for peptides and proteins analyzed by MALDI-MS. *Methods Mol Biol.* 2000;146:405–424.
- [17] Strohm, M., Kavan, D., Novak, P., Volny, M. and Havlicek, V., mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Analytical chemistry*, 2010 82, 4648-4651.
- [18] Editorial, COVID-19 and diabetes: a co-conspiracy? The Lancet Diabetes & Endocrinology, 2020, 8, 801. [https://doi.org/10.1016/S2213-8587\(20\)30315-6](https://doi.org/10.1016/S2213-8587(20)30315-6)
- [19] Lauc G, Rudan I, Campbell H, Rudd PM. Complex genetic regulation of protein glycosylation. *Mol Biosyst.* 2010 Feb;6(2):329-35. doi: 10.1039/b910377e.
- [20] Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell.* 2006 Sep 8;126(5):855-67. doi: 10.1016/j.cell.2006.08.019. PMID: 16959566.
- [21] Ruhaak, L.R., Zauner, G., Huhn, C. *et al.* Glycan labeling strategies and their use in identification and quantification. *Anal Bioanal Chem* 2010 **397**, 3457–3481. <https://doi.org/10.1007/s00216-010-3532-z>
- [22] Guttman, A., Chen, F.-T. A., Evangelista, R. A. & Cooke, N. High-resolution capillary gel electrophoresis of reducing oligosaccharides labeled with 1-aminopyrene-3, 6, 8-trisulfonate. *Anal. Biochem.* 1996, **233**, 234–242, doi:[10.1006/abio.1996.0034](https://doi.org/10.1006/abio.1996.0034)

- [23] Chen, F.-T. A., Dobashi, T. S. & Evangelista, R. A. Quantitative analysis of sugar constituents of glycoproteins by capillary electrophoresis. *Glycobiology* 1998, **8**, 1045–1052, doi:[10.1093/glycob/8.11.1045](https://doi.org/10.1093/glycob/8.11.1045).
- [24] Yang, S, Zhang, L, Thomas, S et al. Modification of sialic acids on solid-phase: accurate characterization of protein sialylation. *Anal. Chem.* 2017 **89**, 6330–6335, doi:[10.1021/acs.analchem.7b01048](https://doi.org/10.1021/acs.analchem.7b01048)(2017).
- [25] Yang, S., Clark, D., Liu, Y. *et al.* High-throughput analysis of N-glycans using AutoTip via glycoprotein immobilization. *Sci Rep* **7**, 10216 (2017). <https://doi.org/10.1038/s41598-017-10487-8>
- [26] Nie H, Li Y, Sun XL. Recent advances in sialic acid-focused glycomics. *J Proteomics*. 2012;75(11):3098-3112. doi:10.1016/j.jprot.2012.03.050
- [27] Lattova, E, Perreault H, Krokhin O, Matrix-assisted laser desorption/ionization tandem mass spectrometry and post-source decay fragmentation study of phenylhydrazones of N-linked oligosaccharides from ovalbumin, *Journal of the American Society for Mass Spectrometry*, 2004, **15**, 725-735, <https://doi.org/10.1016/j.jasms.2004.01.012>.
- [28] Remily-Wood, Dirscherl, EH & Koomen, JM. Acid Hydrolysis of Proteins in Matrix Assisted Laser Desorption Ionization Matrices, *Journal of the American Society for Mass Spectrometry*, 2009, **20**: 2106-2115, <https://doi.org/10.1016/j.jasms.2009.07.007>.
- [29] Wide L, Eriksson K, Sluss PM, Hall JE. Serum half-life of pituitary gonadotropins is decreased by sulfonation and increased by sialylation in women. *J Clin Endocrinol Metab*. 2009 Mar;94(3):958-64. doi: 10.1210/jc.2008-2070.
- [30] Iles RK, Chard T. Molecular insights into the structure and function of human chorionic gonadotrophin. *J Mol Endocrinol*. 1993 Jun;10(3):217-34. doi: 10.1677/jme.0.0100217.
- [31] Furuhashi M, Shikone T, Fares FA, Sugahara T, Hsueh AJ, Boime I. Fusing the carboxy-terminal peptide of the chorionic gonadotropin (CG) beta-subunit to the common alpha-subunit: retention of O-linked glycosylation and enhanced in vivo bioactivity of chimeric human CG. *Mol Endocrinol*. 1995 Jan;9(1):54-63. doi: 10.1210/mend.9.1.7539107.
- [32] Iles RK, Cole LA, Butler SA. Direct analysis of hCGβcf glycosylation in normal and aberrant pregnancy by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Int J Mol Sci*. 2014 Jun 5;15(6):10067-82. doi: 10.3390/ijms150610067.

- [33] Plomp R, Dekkers G, Rombouts Y, et al. Hinge-Region O-Glycosylation of Human Immunoglobulin G3 (IgG3). *Mol Cell Proteomics*. 2015;14(5):1373-1384. doi:10.1074/mcp.M114.047381
- [34] Medzihradszky KF, Kaasik K, Chalkley RJ. Tissue-Specific Glycosylation at the Glycopeptide Level. *Mol Cell Proteomics*. 2015 Aug;14(8):2103-10. doi: 10.1074/mcp.M115.050393.
- [35] Reily, C., Stewart, T.J., Renfrow, M.B. et al. Glycosylation in health and disease. *Nat Rev Nephrol* 2019. 15, 346–366 <https://doi.org/10.1038/s41581-019-0129-4>
- [36] Stencel-Baerenwald JE, Reiss K, Reiter DM, Stehle T, Dermody TS. The sweet spot: defining virus-sialic acid interactions. *Nat Rev Microbiol*. 2014;12(11):739-749. doi:10.1038/nrmicro3346
- [37] Ayora-Talavera G. Sialic acid receptors: focus on their role in influenza infection. *Journal of Receptor, Ligand and Channel Research*. 2018;10:1-11 <https://doi.org/10.2147/JRLCR.S140624>
- [38] Hulswit RJG., Lang, Y., Bakkers, MJG., et al. Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proceedings of the National Academy of Sciences* Feb 2019, 116 (7) 2681-2690; DOI:10.1073/pnas.1809667116
- [39] Kuchipudi, S.V; Nelli, R.K; Gontu, A. et al. Sialic Acid Receptors: The Key to Solving the Enigma of Zoonotic Virus Spillover. *Viruses* 2021, 13, 262. <https://doi.org/10.3390/v13020262>
- [40] Baker, A.N., Richards, S-J, Guy, C.S. et al The SARS-COV-2 Spike Protein Binds Sialic Acids and Enables Rapid Detection in a Lateral Flow Point of Care Diagnostic Device *ACS Central Science* 2020 6 (11), 2046-2052 DOI: 10.1021/acscentsci.0c00855
- [41] Wu JT, Tu MC, Zhung P. Advanced glycation end product (AGE): characterization of the products from the reaction between D-glucose and serum albumin. *J Clin Lab Anal*. 1996;10(1):21-34. doi: 10.1002/(SICI)1098-2825(1996)10:1<21::AID-JCLA5>3.0.CO;2-O. PMID: 8926563.
- [42] Cutler DM, Summers LH. The COVID-19 Pandemic and the \$16 Trillion Virus. *JAMA*. 2020;324(15):1495–1496. doi:10.1001/jama.2020.19759
- [43] <https://www.instituteforgovernment.org.uk/explainers/cost-coronavirus>

- [44] Iles, R.K.; Zmuidinaite, R.; Iles, J.K.; Carnell, G.; Sampson, A.; Heeney, J.L. Development of a Clinical MALDI-ToF Mass Spectrometry Assay for SARS-CoV-2: Rational Design and Multi-Disciplinary Team Work. *Diagnostics* **2020**, *10*, 746. <https://doi.org/10.3390/diagnostics10100746>
- [45] Justin H. Griffin, Kevin M. Downard, Mass spectrometry analytical responses to the SARS-CoV2 coronavirus in review, *TrAC Trends in Analytical Chemistry*, 2021 142, 116328, <https://doi.org/10.1016/j.trac.2021.116328>.

Acknowledgements

This study was undertaken by the Humoral Immune Correlates to COVID-19 (HICC) consortium, funded by the UKRI and NIHR; grant number G107217 (COV0170 - HICC: Humoral Immune Correlates for COVID-19). RKI is also funded by NISAD Ideell Förening (charitable association) Organisationsnummer 802528-6157

We grateful thank Bruker UK Ltd, Coventry for the loan of Microflex MALDI-ToF mass spectrometer and Dr Erika Tranfield and Dr Julie Green for technical support with Bruker MALDI-ToF.

Legend to figures

Figure 1 - Mass spectra profile, 40,000m/z to 85,000m/z, of human reduced plasma samples (treated with TCEP) in order to reveal Immunoglobulin heavy chains (IgM Hc ~75,000m/z, IgA ~56,000 m/z not found, IgG3 54,000 and unconfirmed IgG(x), thought to be IgG4, at 48,000 m/z. Human serum albumin resolved at 66,400+ and transferrin at 79,000m/z.

Figure 2 – Box and whisker plots of relative intensities and variance in peak apex molecular mass of IgG1 heavy chains (IgG1 Hc), IgG3 heavy Chains (IgG3 Hc) and human Serum albumin (HAS) for the different sample groups: **Blue** represents data from SARS-CoV2 sero-negative HCW, **Orange** from SARS-CoV2 HCW sero-positive having recovered from COVID-19 with mild symptoms and **Red** sample data from convalescent patients recovering from COVID-19 ARDS. The table to the right is numeric data from the plots detailing the mean and medium values of peak intensity (AU) and molecular mass (m/z) of IgG1Hc, IgG3Hc and HSA for the respective groups.

Figure 3 – Diagrammatic representation of IgG1 and IgG3 illustrating the differences in heavy chain structure with special reference to the larger neck domain and the O-linked glycosylation sites found there.

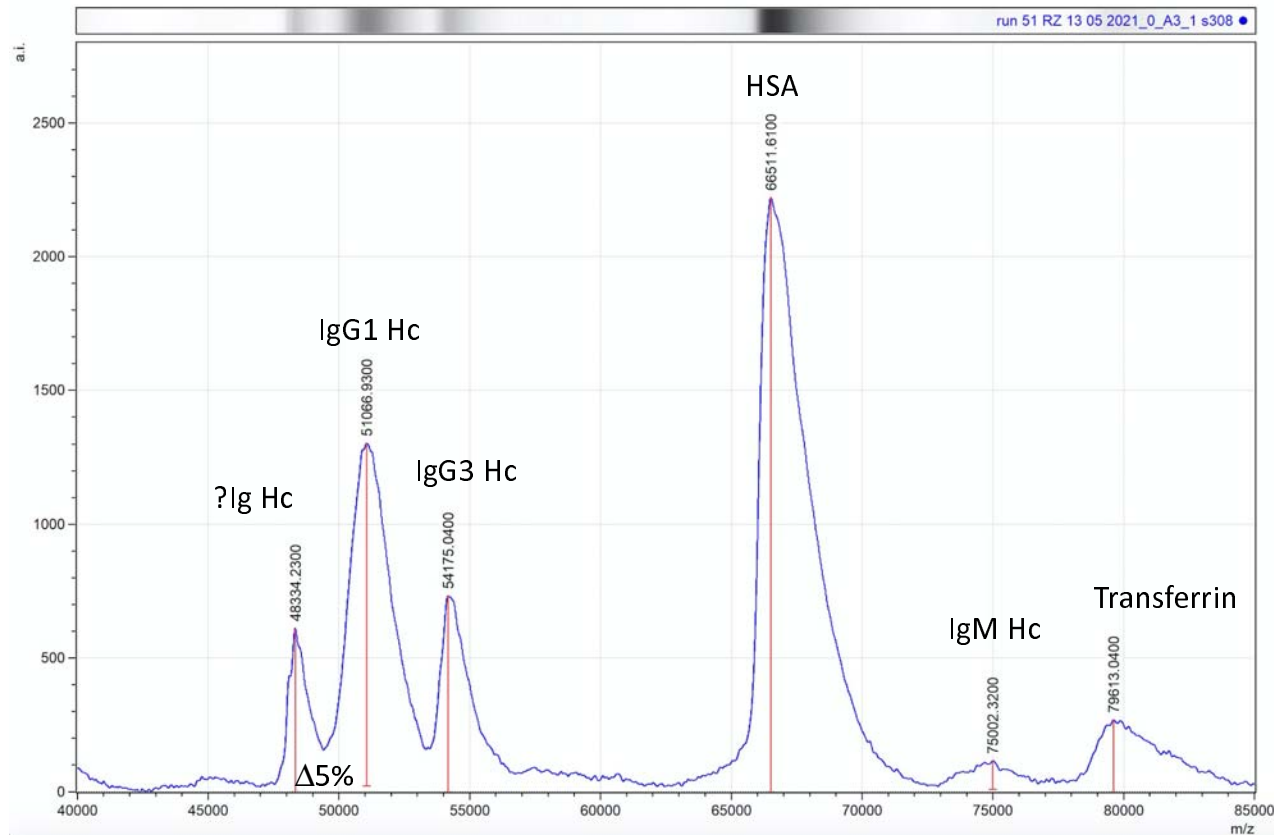


Figure 1

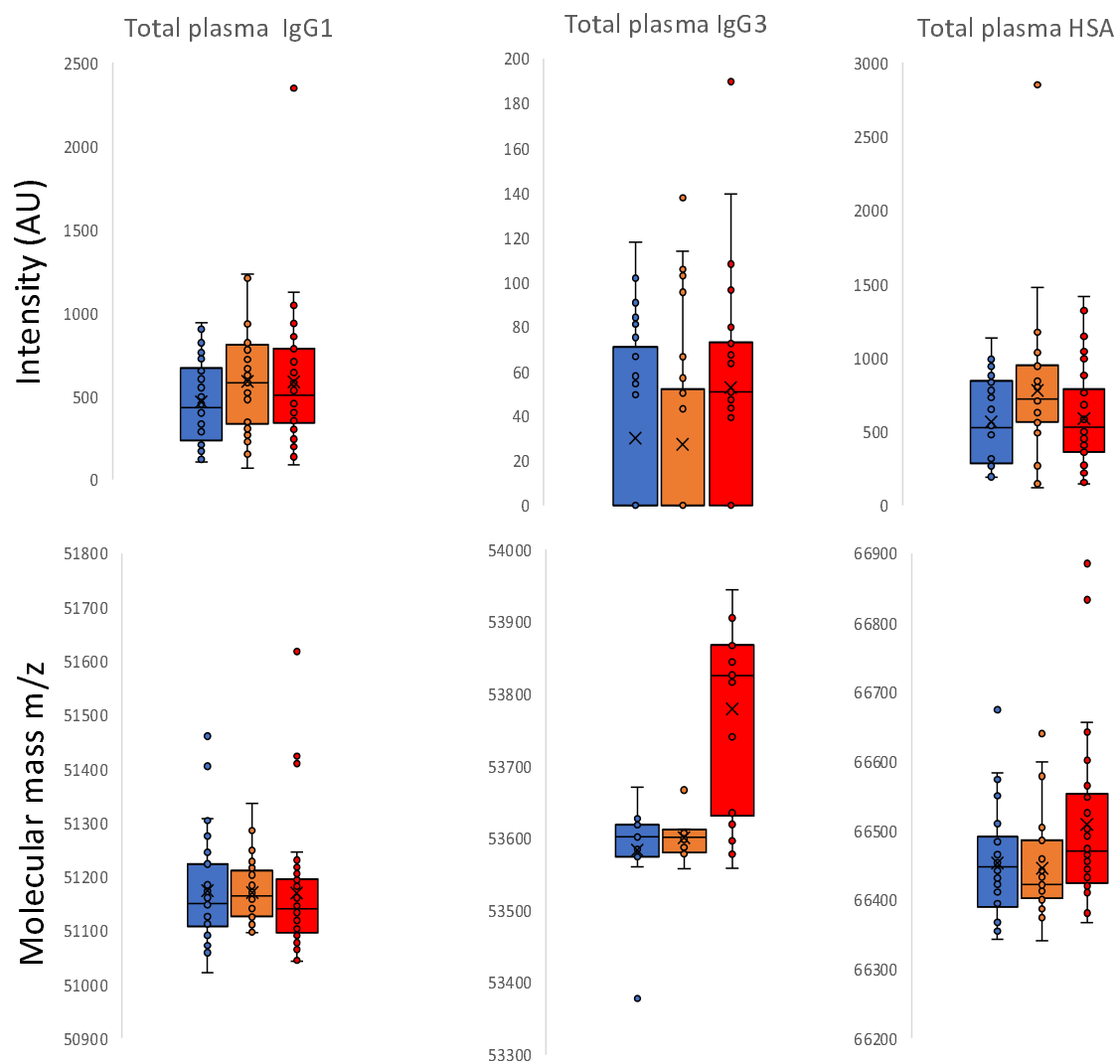


Figure 2

Spectral peak	Study group	N° samples peak detected	Average Mass (m/z)	Average Intensity (AU)
IgG1	HCW Sero negative	30/30 (100%)	Medium 51,151 Mean 51,174	Median 431 Mean 464
	HCW Sero positive	31/31 (100%)	Medium 51,164 Mean 51,170	Median 580 Mean 586
	Patient-COVID-19 ARD	38/38 (100%)	Medium 51,140 Mean 51,169	Median 506 Mean 585
IgG3	HCW Sero negative	11/30 (37%)	Medium 53,601 Mean 53,584	Median 0 Mean 30
	HCW Sero positive	10/31 (32%)	Medium 53,601 Mean 53,600	Median 0 Mean 27
	Patient-COVID-19 ARD	22/38 (58%)	Medium 53,828 Mean 53,779	Median 51 Mean 53
HSA	HCW Sero negative	30/30 (100%)	Medium 66,448 Mean 66,473	Median 528 Mean 567
	HCW Sero positive	31/31 (100%)	Medium 66,422 Mean 66,455	Median 720 Mean 778
	Patient-COVID-19 ARD	38/38 (100%)	Medium 66,471 Mean 66,509	Median 557 Mean 592

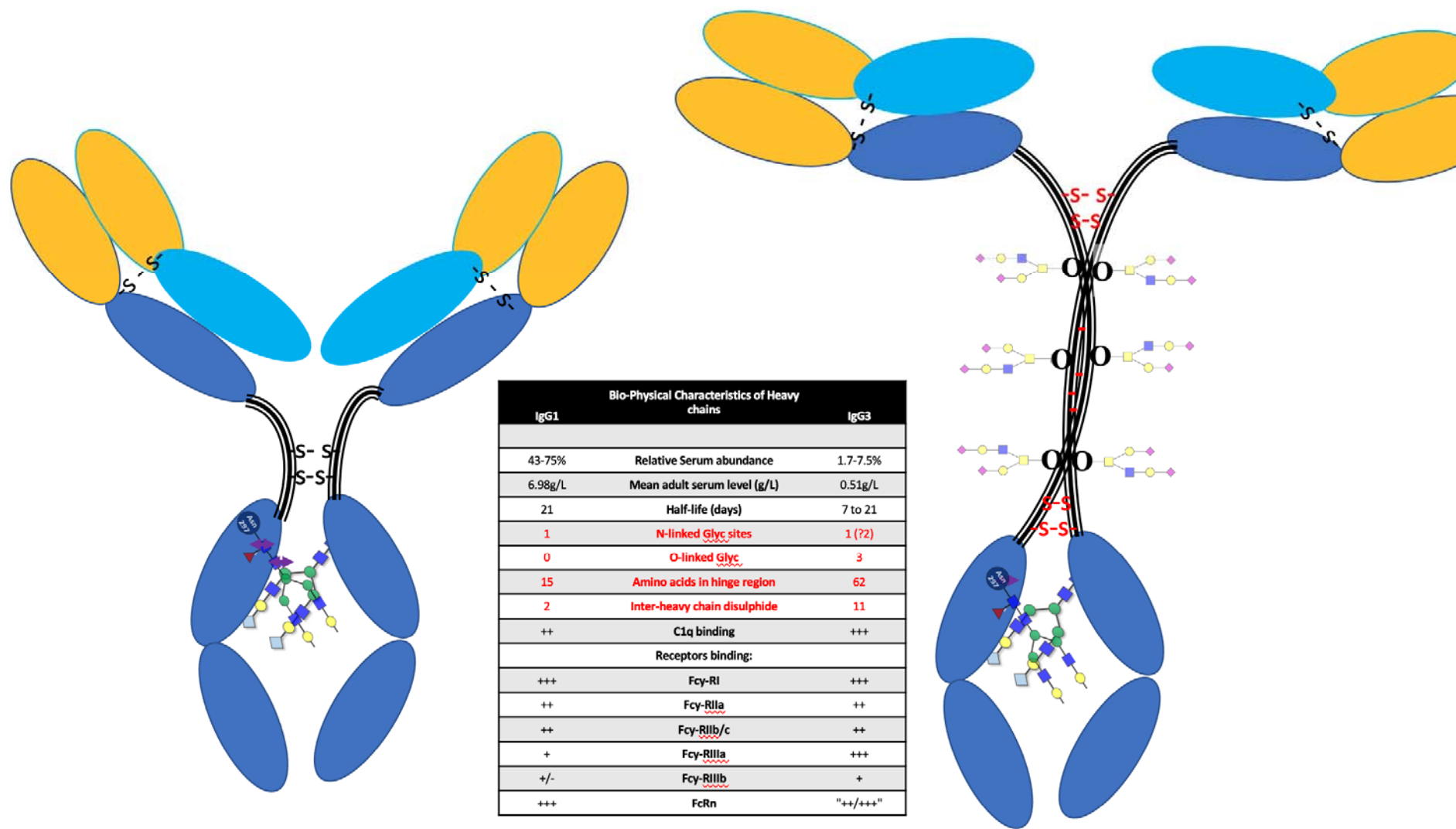


Figure 3