



***N*-Glycosylation Site Analysis of Citrullinated Antigen-Specific B-Cell Receptors Indicates Alternative Selection Pathways During Autoreactive B-Cell Development**

Rochelle D. Vergroesen¹, Linda M. Slot¹, Barbera D. C. van Schaik², Marvyn T. Koning³, Theo Rispens⁴, Antoine H. C. van Kampen^{2,5}, René E. M. Toes¹ and Hans U. Scherer^{1*}

¹ Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands, ² Bioinformatics Laboratory, Amsterdam Public Health Research Institute, Amsterdam Infection & Immunity Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands, ³ Department of Hematology, Leiden University Medical Center, Leiden, Netherlands, ⁴ Department of Immunopathology, Sanquin Research and Landsteiner Laboratory, Academic Medical Center, Amsterdam, Netherlands, ⁵ Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, Netherlands

OPEN ACCESS

Edited by:

Stefano Casola,
IFOM - The FIRC Institute of Molecular
Oncology, Italy

Reviewed by:

Wenxia Song,
University of Maryland, College Park,
United States
Francesco Forconi,
University of Southampton,
United Kingdom

*Correspondence:

Hans U. Scherer
h.u.scherer@lumc.nl

Specialty section:

This article was submitted to
B Cell Biology,
a section of the journal
Frontiers in Immunology

Received: 22 December 2018

Accepted: 19 August 2019

Published: 04 September 2019

Citation:

Vergroesen RD, Slot LM, van
Schaik BDC, Koning MT, Rispens T,
van Kampen AHC, Toes REM and
Scherer HU (2019) *N*-Glycosylation
Site Analysis of Citrullinated
Antigen-Specific B-Cell Receptors
Indicates Alternative Selection
Pathways During Autoreactive B-Cell
Development.
Front. Immunol. 10:2092.
doi: 10.3389/fimmu.2019.02092

Many autoimmune diseases are hallmarked by autoreactive B and plasma cell responses that are directly or indirectly involved in disease pathogenesis. These B-cell responses show large variability between diseases, both in terms of the secreted autoantibody repertoire and the dynamics and characteristics of the underlying B-cell responses. Hence, different mechanisms have been proposed to explain the emergence of autoreactive B cells in an otherwise self-tolerant immune system. Notably, most mechanistic insights have been obtained from murine studies using models harboring genetic modifications of B and T cells. Given recent technological advances that have rendered autoreactive human B cells accessible for analysis, we here discuss the phenomenon of extensive *N*-glycosylation of the B-cell receptor (BCR) variable domain of a prototypic human autoreactive B-cell response and its potential role in the generation of autoimmunity. Anti-citrullinated protein antibodies (ACPA) hallmark the most disease-specific autoimmune response in Rheumatoid Arthritis (RA). Remarkably, ACPA-IgG are heavily *N*-glycosylated in the variable domain due to somatic mutations that generate abundant *N*-glycosylation consensus sequences. These sites, obtained from full-length BCR sequences of ACPA-expressing B cells from 12 ACPA-positive RA patients, were here analyzed in detail. Sites that required a single nucleotide mutation to be generated were defined as single somatic hypermutation (s-SHM) sites, whereas sites requiring multiple mutations were defined as m-SHM sites. IgG sequences of 12 healthy donors were used as control. Computational modeling of the germinal center reaction (CLONE algorithm) was used with the germline counterparts of ACPA-IgG heavy chain (HC) sequences to simulate the germinal center response. Our analyses revealed an abundance of *N*-glycosylation sites in ACPA-IgG HC that frequently required multiple mutations and predominated in specific positions. Based on these data, and taking into account recent insights into the dynamics of the ACPA-response during

disease development, we here discuss the hypothesis that *N*-glycosylation sites in ACPA-IgG variable domains could lead to alternative, possibly antibody affinity-independent selection forces. Presumably, this occurs during germinal center responses allowing these B cells to escape from putative tolerance checkpoints, thereby driving autoreactive B cell development in the pathogenesis of RA.

Keywords: anti-citrullinated protein antibodies, B cells, glycosylation, variable domain (Fab), rheumatoid arthritis, germinal center

INTRODUCTION

The mechanisms leading to the development of human B-cell mediated autoimmunity are incompletely understood. In particular, it is unclear how autoreactive B cells arise and escape tolerance checkpoints that normally control their development and expansion. While transient, short-lived autoreactive B-cell responses are frequently observed in the context of external, innate triggers such as infections, the development of long-lived, autoreactive B and plasma cells in the context of autoimmune disease likely requires T-cell help and the involvement of germinal centers (GCs) in lymph nodes or GC-like structures in inflamed tissues. As these specialized structures are equipped with various control mechanisms to prevent the development of autoreactivity (1, 2), it is important to understand why and how these protective mechanisms fail in the development of human autoimmunity. Using sequencing data of full-length variable regions of B-cell receptors (BCRs) specific for the most relevant autoantigen in rheumatoid arthritis, a prototypic autoimmune disease, we here discuss the hypothesis that human autoreactive B cells can circumvent classical mechanisms of negative selection during GC reactions by introducing *N*-glycosylation sites in the antibody variable (*V*-) domain. In fact, we provide evidence for the notion that the remarkable *V*-domain *N*-glycosylation observed for antibodies against citrullinated protein antigens (ACPA) in this disease allows for the outgrowth of autoreactive B cells that show extensive somatic hypermutation (SHM) with limited concurrent affinity maturation. Hence, the classical process of affinity maturation as a means to increase antigen specificity and to avoid negative selection is uncoupled from the process of SHM in this particular response. While the exact mechanisms by which *V*-domain *N*-glycans allow these B cells to escape tolerance control remain to be determined, our data provide an example of BCR diversification through abundant *N*-glycosylation sites in the antigen-specific BCR repertoire of a human autoreactive B-cell response and form the basis for further exploration of this intriguing phenomenon.

RHEUMATOID ARTHRITIS

Rheumatoid arthritis is a chronic inflammatory disease that affects ~1% of the population (3). The majority of patients harbors autoantibodies, of which those most frequently observed target protein antigens in which arginine residues have undergone enzymatic, posttranslational modification into the non-classical amino acid citrulline (4). Citrulline is the common

determinant recognized by the antibodies generated. In principal, any arginine containing protein can be citrullinated. Because of this, and as a consequence of the structural determinants of citrulline recognition by anti-citrullinated protein antibodies (ACPA) at the molecular level (5), ACPA display extensive cross-reactivity and frequently recognize various citrullinated proteins (6, 7). So far, this has hampered the identification of single antigens as triggers of this response, and it is likely that no such single antigen exists. From a clinical perspective, ACPA, together with rheumatoid factors, represent the most specific and clinically most relevant biomarker in this disease (8, 9). Despite strong clinical associations, the question whether or not these autoantibodies and/or the underlying citrullinated-protein-related T- and B-cell responses are involved in disease initiation and chronicity remains a matter of research and debate. However, the autoreactive B-cell immune response that generates ACPA has lately received much attention, especially because therapeutic depletion of CD20⁺ B cells has proven to be effective in established disease, in particular in the autoantibody-positive subset of patients (10–12). In addition, technology has advanced such that ACPA-expressing B cells can now be identified in, and isolated from peripheral blood and synovial fluid of affected patients, despite their very low frequency in the circulation (13–17). This allows for an in depth analysis of the BCR repertoire and of the molecular characteristics of the BCR. Besides, it allows to gain insight into the development of this disease-specific autoreactive B-cell response.

EXTENSIVE V-DOMAIN N-GLYCOSYLATION OF ACPA-IgG

Previously, we could demonstrate that ACPA of the immunoglobulin (Ig)G isotype are extensively *N*-glycosylated in the *V*-domain (18). In fact, on average >90% of secreted ACPA-IgG molecules isolated from serum carry such glycans, which is in marked contrast to the much lower frequency of *V*-domain *N*-glycans observed in the total serum IgG repertoire. In-depth analysis of glycans found on ACPA-IgG *V*-domains revealed that these are expressed as fully-developed, biantennary *N*-glycans that frequently terminate with negatively charged sialic acid residues (19). This remarkable presence of glycans causes a shift in molecular size if affinity-purified ACPA-IgG or their corresponding F(ab)₂-fragments are analyzed by gel electrophoresis or size-exclusion chromatography. As this is not the case for ACPA-IgM (20), we hypothesized that ACPA-IgG acquire *N*-glycans by introducing somatic mutations at positions

that generate the *N*-glycosylation consensus sequence N-X-S/T (where X is any amino acid except proline), which is required for protein *N*-glycosylation. Such tripeptide sequons are only expressed by very few germline V-region genes. However, the repertoire of V-regions harbors multiple locations (termed progenitor sites) at which a single somatic mutation can lead to the generation of the sequon, i.e., of an *N*-glycosylation site (21). Indeed, recent sequence analysis of the V-regions of heavy chains (HC) of citrullinated antigen-specific BCRs showed a high frequency (>80%) of such *N*-glycosylation sites in the polyclonal ACPA-IgG repertoire, in contrast to a remarkable scarcity of such sites in a comparative, V-region matched repertoire of healthy donor IgG (22, 23). Notably, all of the *N*-glycosylation sites identified in the ACPA-IgG V-region repertoire were *de-novo* generated sites originating from somatic mutations, while no germline-encoded sites were observed (22).

LOW-AVIDITY AND CROSS-REACTIVITY OF THE ACPA RESPONSE

Interestingly, the B-cell response directed against citrullinated antigens generates a secreted polyclonal antibody repertoire that is of remarkably low-avidity (24). This low-avidity response can likely explain the concurrent presence of ACPA-IgM and ACPA-IgG, as IgM-expressing B cells will not be outcompeted by high-affinity IgGs normally seen during immune responses after e.g., tetanus vaccination. Nonetheless, the ACPA-response undergoes extensive isotype-switching (25) and somatic hypermutation (22), indicating that ACPA-expressing B cells are able to acquire T-cell help required for the expression of Activation Induced Cytidine Deaminase (AID). In addition, there is considerable cross-reactivity of the antibody pool toward different citrullinated proteins and toward other post-translational modifications (6, 26). On the clonal level, the response harbors clones that recognize citrullinated antigens only, while others are cross-reactive to antigens modified at lysine residues to express homocitrulline. However, all clones require the posttranslational protein modification as common determinant and do not bind the native protein. This cross-reactivity and the lack of high avidity are remarkable features, given that ACPA-IgG V-regions carry extensive somatic mutations. In fact, the number of nucleotide mutations in ACPA-IgG exceeds those of the total IgG repertoire and those of common, high-avidity vaccine responses directed against, for example, tetanus toxoid (22). Hence, ACPA-expressing B cells undergo extensive somatic mutation events, presumably during multiple cycles through GCs, yet their selection for survival does not seem to require an extensive increase of antibody affinity, as would be expected from a persistent immune response that is not clonally deleted (27).

NON-RANDOM ACCUMULATION OF *N*-GLYCOSYLATION SITES

Given the observations described, we hypothesized that ACPA-expressing B cells could be selected in GCs or GC-like structures

based on the presence of *N*-glycosylation sites in the V-domain rather than on the affinity by which their cognate antigens are recognized. If correct, one could postulate that this alternative selection mechanism depends on the accumulation of *N*-glycosylation sites in the V-region, either at specific positions or dispersed over the V-region at high frequency with the presence of progenitor sites and/or the site-specific activity of AID [at so-called hotspot motifs (28)] governing their distribution. In fact, such an accumulation of *N*-glycosylation sites should not be a simple (“random”) reflection of frequent somatic mutations, but rather occur in both highly mutated clones and in those harboring fewer mutations. Indeed, our first analysis indicated that *N*-glycosylation sites in ACPA BCR IGHV-regions do not accumulate randomly over the V-region sequence. In fact, although we observed that somatic mutations were required for the generation of *N*-glycosylation sites, it was striking that the number of nucleotide mutations in ACPA IGHV-regions did not correlate with the number of *N*-glycosylation sites that were generated. Hence, zero, one or multiple *N*-glycosylation sites were observed in clones with few and frequent mutations, and not only highly mutated clones showed a high number of *N*-glycosylation sites (22). This latter finding is relevant as it suggests that the introduction of *N*-glycosylation sites is not the mere consequence of a high mutation rate. Rather, the mutation rate is likely a consequence of continuous (re-)activation of the autoreactive B-cell population, while the *N*-glycans could be involved in the provision of additional signals that provide a selective advantage to these cells. To further substantiate this notion, we performed a subsequent, in-depth analysis of *N*-glycosylation sites and their characteristics in the ACPA-IgG IGHV-region repertoire. Full length V-region sequences of IgG HC were obtained from single cell isolation of ACPA-expressing B cells using cyclic citrullinated peptide (CCP)2 as a model antigen. This model peptide has been designed for clinical diagnostic use to capture a broad repertoire of polyclonal ACPA. Importantly, antibodies recognizing this peptide are bona-fide ACPA that also recognize various citrullinated proteins (7). Cell isolation was followed by Anchoring Reverse Transcription of Immunoglobulin Sequences and Amplification by Nested PCR [“ARTISAN PCR” (29)] and Sanger sequencing. In addition, ACPA-expressing B cells were isolated as pools followed by ARTISAN-PCR and next generation sequencing (22). IgG sequences were selected for analysis from single-sorted ACPA-expressing B cells, whereas sequencing data from pool-sorted ACPA-expressing B cells contained heavy chain transcripts from IgG (~83%), IgA (~11%), and IgM (~6%) expressing cells. Sequences were analyzed using IMGT-(High)V-quest (30), ARGalaxy (31), and Microsoft Access to define BCR sequences as unique productive clones based on VJ rearrangement + CDR3(AA) sequence identity. In addition, we retrieved information about the introduction of *N*-glycosylation sites across the variable region. As described, we observed acquired, non-germline encoded, *N*-glycosylation sites in >80% of ACPA-IgG V-region genes. Our subsequent analysis revealed that most clones harbored one site per HC, while two or three sites per HC were also observed but less frequently (**Figure 1A**). Importantly, the frequency of sites was highly comparable

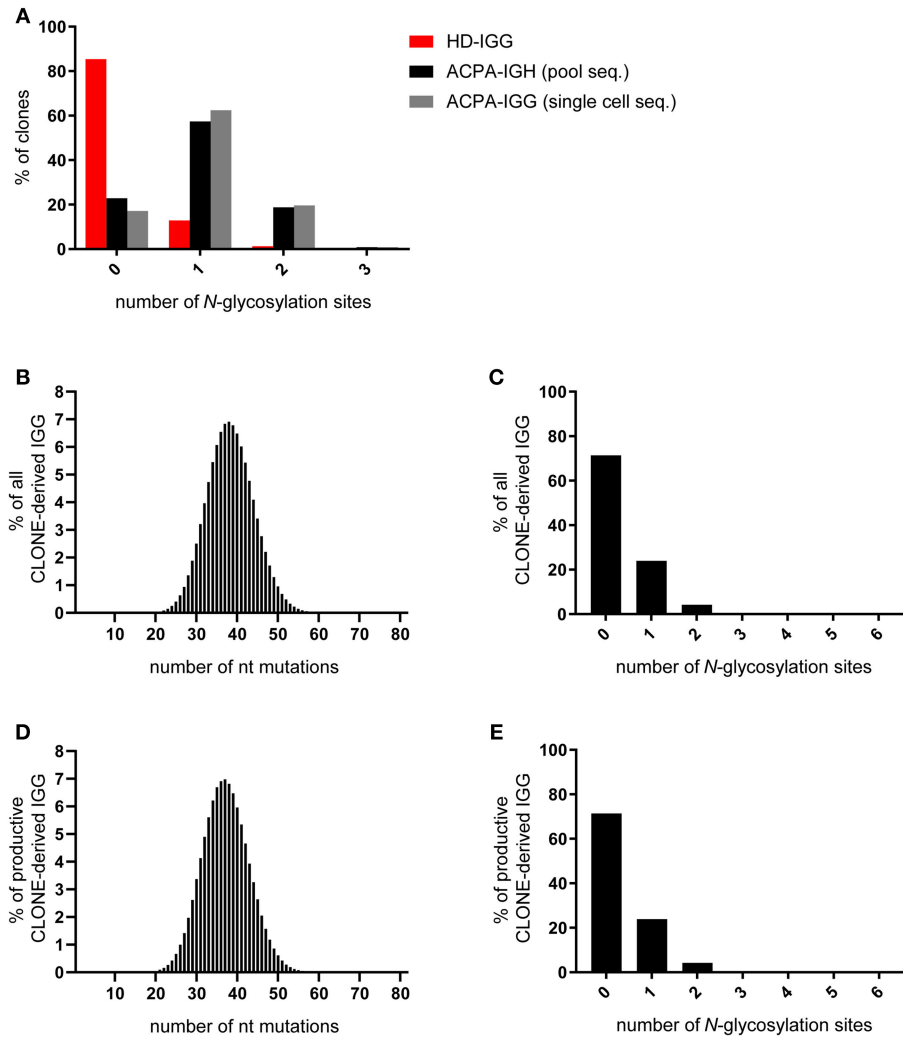


FIGURE 1 | Frequency of *N*-glycosylation sites in the immunoglobulin heavy chain variable domain. **(A)** Percentage of ACPA- and healthy donor-derived IgG heavy chain clones containing zero, one, or more *N*-glycosylation sites in their BCR. Unique productive ACPA-derived clones were obtained using pool-sequencing (black bar) and single cell-sequencing (gray bar) and compared to unique productive pool-sequenced HD-IgG (red bar). **(B)** Frequency of nucleotide mutations in all sequences derived from the CLONE-algorithm with germline-reverted ACPA-IgG sequences as input data. **(C)** Percentage of all CLONE algorithm-derived clones containing zero, one, or more *N*-glycosylation sites. **(D,E)** Subset analysis exclusively on CLONE-algorithm derived productive sequences, without stop codons in IGHV, as analyzed in **(B,C)**.

between sequences of both isolation and sequencing techniques, excluding sequencing errors of either technique as a potential source of bias in the data. More specifically, 77% of pool-sorted ACPA B cells (223 clones) and 83% of single cell-sorted ACPA B cells (117 clones) contained one or more acquired sites (Figure 1A). In contrast, sites were absent in >80% of IgG-clones isolated from a healthy donor IgG HC repertoire. As expected, the number of mutations found in the healthy donor sequences was substantially lower compared to the mutation frequency observed in ACPA-IgG. In fact, a mean of almost 50 nucleotide mutations was observed in ACPA-IgG sequences compared to ~25 nucleotide mutations in the healthy donor repertoire (22). While this comparison demonstrates the remarkable degree of V-domain *N*-glycosylation of ACPA-IgG, we felt that caution

should be taken as to a direct comparison of *N*-glycosylation site frequencies because of the substantially lower number of nucleotide mutations present in the healthy donor repertoire.

COMPUTATIONAL GERMINAL CENTER SIMULATION SUPPORTS THE CONCEPT OF SELECTIVE INTRODUCTION OF *N*-GLYCOSYLATION SITES

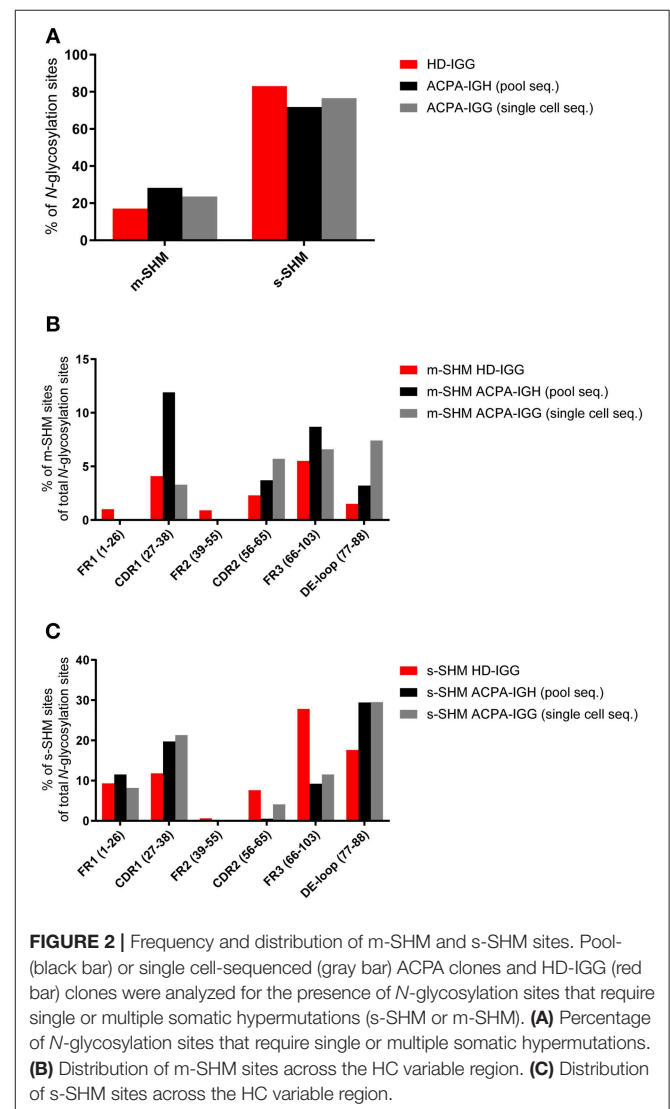
To account for this difference in mutation frequency, we simulated the generation of *N*-glycosylation sites during the process of SHM in the GC response using an adapted, pre-existing computational model [CLONE (32, 33)] that provides

a stochastic simulation of affinity maturation. The simulation is initiated with a single B cell associated to a selected germline sequence obtained from reverting IgG sequences ($n = 163$) obtained from ACPA-expressing B cells. This process necessarily conserved the CDR3 region. In the model, all cells divide once and acquire a Poisson distributed number of somatic mutations with an average that matches the number of mutations observed in the ACPA sequences as a proxy for repetitive cycles of B cell circulation through GCs. According to pre-defined probabilities, a mutation may have no effect, causes cell death (negative selection) or provides a selective advantage (positive selection) by decreasing its apoptosis rate. The simulation was terminated after 40 cell divisions. The simulation was repeated for each of the 163 input sequences and for each sequence we ran the simulation at least 10 times. The simulation yielded a Gaussian distribution of nucleotide mutations based on 26.799.339 clones (median of 37 nucleotide mutations per clone; **Figure 1B**). Despite this enhanced number of somatic mutations, the vast majority of clones did not harbor any *N*-glycosylation sites (**Figure 1C**). This was the case if all generated sequences were taken into account, but also if the large number of non-productive sequences that the algorithm generated with stop codons (81%) were removed from the analysis (**Figures 1D,E**). Together, these data indicate that the process of somatic hypermutation during GC-responses in the healthy situation, even if artificially “pushed” to mutation rates exceeding those of protective, antigen-specific responses, does not lead to the accumulation of *N*-glycosylation sites comparable to the degree observed for ACPA-IgG. This also indicates that *N*-glycans in the IGHV-domain do not contribute to the selection of (presumably) higher affinity clones, which is in line with the frequency of *N*-glycosylation sites in the healthy donor data set (**Figure 1A**), our observations from tetanus toxoid-specific clones (22), the low avidity of the ACPA response (24), and work from others (34). Hence, it is conceivable that additional, glycan-dependent selective mechanisms influence the accumulation of sites observed in the citrullinated antigen-directed, autoreactive B-cell response.

***N*-GLYCOSYLATION SITES RESULT FROM SINGLE AND MULTIPLE SOMATIC HYPERMUTATION EVENTS**

While only few germline-encoded immunoglobulin heavy chain variable (IGHV) region genes, such as IGHV1-8, IGHV4-34, and IGHV5-1, harbor an *N*-glycosylation site consensus sequence, there is an abundance of sequence positions which require only a single somatic mutation to generate a *de-novo* *N*-glycosylation site (so called progenitor glycosylation sites) (21). Additional sites can be formed on other positions but these require multiple mutations. Given the observations described above indicating the preferred outgrowth of ACPA-expressing B cells that have introduced *N*-glycosylation sites and a glycosylation-dependent selective mechanism, we hypothesized that single somatic hypermutation sites (s-SHM), i.e., sites generated from progenitor glycosylation sites, would be abundant in both the

healthy donor and the ACPA-IgG repertoire. In contrast, ACPA-IgG V-regions would be expected not only to harbor a higher number of sites originating from multiple mutations (m-SHM) in absolute terms, but also in relative terms when corrected for the total number of *N*-glycosylation sites in both data sets. Indeed, while s-SHM sites were present in both sets of data, ACPA-IgG harbored ~9% more m-SHM sites. Again, this finding was observed irrespective of the sequencing approach (**Figure 2A**). Of note, within the ACPA repertoire, certain regions within the IGHV-domain, such as complementarity determining region (CDR) 2 and framework region 2 (FR2) more frequently harbored m-SHM sites, whereas s-SHM sites were most abundant in FR1 (**Figures 2B,C**). Besides, *N*-glycosylation sites were more frequent in the DE-loop [the loop between the second and third CDR, position 77–88 of FR3 (21)] than in the remaining part of FR3. The *N*-glycosylation sites in the remaining part of FR3 required more often multiple mutations compared to sites in the DE-loop. Possibly, the distribution of *N*-glycosylation sites



is also influenced by the skewed activity of AID toward hotspot motifs. Indeed, a random analysis of a set of *N*-glycosylation sites and progenitor sites indicated that some but not all acquired sites were situated on AID hotspot motifs (WRCY/WA, data not shown). In addition, Koers et al. (35) analyzed the relative occurrence of “other” mutations on the progenitor sites than the *N*-glycosylation site inducing mutations. Overall, this analysis indicated that there was no intrinsic bias concerning the introduction of *N*-glycosylation sites. In summary, these data show that ACPA-IgG contain not only an absolute and relative increase in *N*-glycosylation-sites, but also indicate that these *N*-glycosylation sites are more often acquired by multiple mutations. Again, these data are in line with the notion that the introduction of *N*-linked glycans in the V-domain of ACPA confers an advantage to citrullinated-protein directed B cells.

PHYSICOCHEMICAL PROPERTIES OF *N*-GLYCOSYLATION SITES IN ACPA-IgG HC

As the mechanism by which V-domain *N*-glycans influence the selection and survival of ACPA-expressing B cells is unclear, we determined the molecular characteristics of the *N*-glycosylation tripeptide sequons with the aim of defining features relevant to such sites in ACPA-IgG. Specifically, we determined the physicochemical properties of the amino acid “X” in “N-X-S/T” and analyzed its charge and hydrophobicity (Figure 3). These analyses revealed that most amino acids at position “X” (except, by definition, proline) could participate in the formation of *N*-glycosylation sites (Figure 3A), with a certain preference for hydrophobic, non-charged amino acids (Figure 3B). In general, no specific pattern or choice of particular amino acids was observed either at position “X” or at the flanking residues (Figure 3C), indicating that the presence of the *N*-glycan itself rather than the structural properties of the amino acid backbone are relevant in the aberrant selection of ACPA-expressing B cells. This observation is important, as the presence of an *N*-glycosylation site in a given protein does not necessarily imply the presence of a glycan. In fact, localization of the site and the conformational folding of the protein might preclude the use of the site for glycosylation, while it could still have functional relevance for the antibody as such. Nonetheless, analysis of 27 recombinantly produced monoclonal antibodies with, in total, 38 sSHM *N*-glycosylation sites in different positions suggests that a high frequency of these sites, and probably other sSHM sites, is occupied by *N*-glycans (35).

CLINICAL RELEVANCE OF ACPA-IgG V-DOMAIN *N*-GLYCOSYLATION

As described, V-domain *N*-glycosylation of ACPA-IgG provides insight in factors that are relevant to the development and selection of citrullinated antigen-directed B cells. However, the presence of this feature presumably also reflects the help from T cells to ACPA-expressing B cells, and as such has implications for its use as a clinical biomarker. Serum ACPA-IgG can persist

for many years in the absence of disease-specific symptoms (36, 37). In fact, not all ACPA-positive individuals will develop RA (38–41). Hence, it is of clinical relevance to identify those individuals who transition from symptom-free autoimmunity to clinically overt autoimmune disease. In this context, it is intriguing that the strongest genetic risk factor for ACPA-positive RA, the HLA shared epitope alleles, confer risk only to ACPA-positive disease but not to ACPA-positivity, i.e., to autoimmunity without clinical symptoms (42, 43). As HLA-molecules are crucial for the communication between T and B cells, it is possible that the transition from autoimmunity to overt, ACPA-positive disease is triggered by T cells that provide help to ACPA-expressing B cells. In doing so, this T cell help is likely to induce or enhance somatic mutations and, hence, the generation of *N*-glycosylation sites. Indeed, first evidence obtained from the study of ACPA-IgG in asymptomatic ACPA-positive individuals indicates that the degree of serum ACPA-IgG V-domain *N*-glycosylation is markedly lower than in RA patients (44). Moreover, although analyzed so far in only a small cohort of patients, the presence of extensive V-domain *N*-glycosylation in serum ACPA-IgG predicted RA-development. Together, these observations demonstrate that the generation of V-domain glycans and the biological consequences might be directly associated to disease-relevant processes, either caused by the activation of disease-relevant T cells, the outgrowth of the autoreactive B-cell population itself or by the secreted autoantibody repertoire. The exact biological mechanism by which the presence of V-domain glycans on ACPA-IgG influence the selection and outgrowth of autoreactive B cells is, however, still unknown. Potential effects mediated by the glycans include the modulation of antigen binding and/or of BCR signaling, or the direct binding to lectins expressed on either the same or neighboring cells (45). Although we did not observe an overrepresentation of *N*-glycosylation sites in the CDR regions, which are supposed to be the most important regions for antigen binding, subtle changes in antigen binding between glycosylated and non-glycosylated monoclonal ACPAs have already been described (18). Hence, additional, in-depth investigations are required to further dissect the biological effector functions associated with these ACPA-IgG V-domain glycans.

SUMMARY AND CONCLUSION

Taken together, we here describe the detailed analysis of the occurrence of *N*-glycosylation sites in the V-domain of a human autoreactive B-cell response and discuss its potential relevance for autoreactive B-cell development. Our data indicate that the extensive *N*-glycosylation found in ACPA-IgG is not the result of the random accumulation of *N*-glycosylation sites due to extensive SHM. In fact, we propose the concept that the *N*-glycosylation is involved, either directly or indirectly, in the (positive) selection of autoreactive B-cell clones, even in the absence of a high affinity recognition of the self-antigen. However, as discussed above, the functional mechanisms by which V-domain *N*-glycans promote positive selection (or inhibit

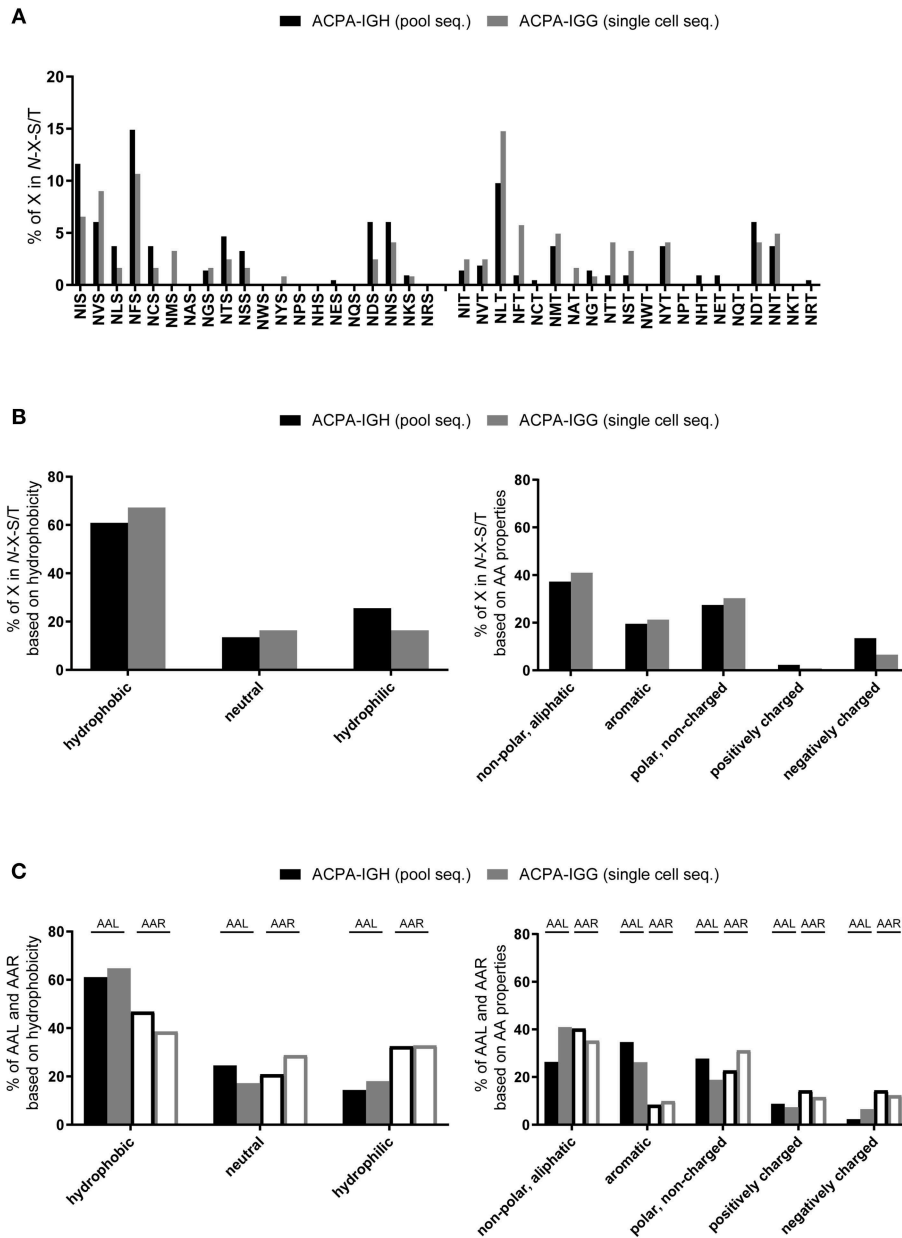


FIGURE 3 | In depth analysis of amino acid “X” in N-glycosylation sites. Pool- (black bar) and single cell-sequenced (gray bar) ACPA clones were analyzed for the properties of amino acid “X” in the consensus sequence N-X-S/T. **(A)** Frequency of all amino acids used at position “X” in N-glycosylation sites. **(B)** Analysis of the hydrophobicity of the amino acids on position “X” and of their non-polar (aliphatic), aromatic, polar (non-charged), positively-, and negatively-charged characteristics. **(C)** Amino acids left (AAL, closed bar) and right (AAR, open bar) flanking N-glycosylation sites ordered by hydrophobicity and by their non-polar (aliphatic), aromatic, polar (non-charged), positively-, and negatively-charged characteristics.

negative selection) are incompletely understood. As different B-cell responses have been described that demonstrate enhanced levels of V-domain N-glycans and/or glycosylation sites, such as those observed in Sjögren’s syndrome (46), those found in follicular lymphoma (47–49), those generating anti-hinge region antibodies, and those in immunoglobulin subclass G4-related disease and primary sclerosing cholangitis (50), it is likely that the mechanisms operational in the development of ACPA-expressing B cells are not restricted to this B-cell response

only. So far, however, we have not identified another antigen-specific B-cell response that generates a degree of V-domain N-glycosylation comparable to that of ACPA-IgG. In conclusion, future studies will need to profit from this extreme phenotype and focus on the consequences of V-domain N-glycans for the recognition and binding strength of citrullinated antigens, on modulation of B-cell receptor-mediated signaling events, and on potential binding interactions between V-domain N-glycans and lectins.

ETHICS STATEMENT

This study was approved by the Institutional Review Board of Leiden University Medical Center. All subjects involved in the study gave written informed consent prior to any study-related procedure.

AUTHOR CONTRIBUTIONS

RV, LS, BS, and MK designed the study, performed the experiments, and analyzed the data. TR contributed to the design of the study, provided algorithms for data analysis, interpreted the data, and critically revised the manuscript. AK, RT, and HS designed the study, analyzed data, and wrote the manuscript.

REFERENCES

- Melchers F. Checkpoints that control B cell development. *J Clin Invest.* (2015) 125:2203–10. doi: 10.1172/JCI78083
- Goodnow CC, Vinuesa CG, Randall KL, Mackay F, Brink R. Control systems and decision making for antibody production. *Nat Immunol.* (2010) 11:681–8. doi: 10.1038/ni.1900
- Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet.* (2010) 376:1094–108. doi: 10.1016/S0140-6736(10)60826-4
- Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J Clin Invest.* (1998) 101:273–81. doi: 10.1172/JCI1316
- Ge C, Xu B, Liang B, Lonnblom E, Lundstrom SL, Zubarev RA, et al. Structural basis of cross-reactivity of anti-citrullinated protein antibodies. *Arthritis Rheumatol.* (2019) 71:210–21. doi: 10.1002/art.40698
- Toes RE, Huizinga TJ. Update on autoantibodies to modified proteins. *Curr Opin Rheumatol.* (2015) 27:262–7. doi: 10.1097/BOR.000000000000166
- Ioan-Facsinay A, el-Bannoudi H, Scherer HU, van der Woude D, Menard HA, Lora M, et al. Anti-cyclic citrullinated peptide antibodies are a collection of anti-citrullinated protein antibodies and contain overlapping and non-overlapping reactivities. *Ann Rheum Dis.* (2011) 70:188–93. doi: 10.1136/ard.2010.131102
- Dorner T, Egerer K, Feist E, Burmester GR. Rheumatoid factor revisited. *Curr Opin Rheumatol.* (2004) 16:246–53. doi: 10.1097/00002281-200405000-00013
- Willemze A, Trouw LA, Toes REM, Huizinga TWJ. The influence of ACPA status and characteristics on the course of RA. *Nat Rev Rheumatol.* (2012) 8:144–52. doi: 10.1038/nrrheum.2011.204
- Isaacs JD, Cohen SB, Emery P, Tak PP, Wang J, Lei G, et al. Effect of baseline rheumatoid factor and anticitrullinated peptide antibody serotype on rituximab clinical response: a meta-analysis. *Ann Rheum Dis.* (2013) 72:329–36. doi: 10.1136/annrheumdis-2011-201117
- Chatzidionysiou K, Lie E, Nasonov E, Lukina G, Hetland ML, Tarp U, et al. Highest clinical effectiveness of rituximab in autoantibody-positive patients with rheumatoid arthritis and in those for whom no more than one previous TNF antagonist has failed: pooled data from 10 European registries. *Ann Rheum Dis.* (2011) 70:1575–80. doi: 10.1136/ard.2010.148759
- Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR, et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. *N Engl J Med.* (2004) 350:2572–81. doi: 10.1056/NEJMoa032534
- Kerkman PF, Fabre E, van der Voort EI, Zaldumbide A, Rombouts Y, Rispens T, et al. Identification and characterisation of citrullinated antigen-specific B cells in peripheral blood of patients with rheumatoid arthritis. *Ann Rheum Dis.* (2016) 75:1170–6. doi: 10.1136/annrheumdis-2014-207182
- Titcombe PJ, Wigerblad G, Sippl N, Zhang N, Shmagel AK, Sahlstrom P, et al. Pathogenic citrulline-multispecific B cell receptor clades in rheumatoid arthritis. *Arthritis Rheumatol.* (2018) 70:1933–45. doi: 10.1002/art.40590

FUNDING

The authors acknowledge grant support from the Netherlands Organization for Scientific Research (NWO) (projects 435000033 and 91214031), the IMI funded project BeTheCure (contract 1151422), and the IMI funded project RTCure (contract 777357). HS is the recipient of an NWO-ZonMW clinical fellowship (project 90714509), an NWO-ZonMW VENI grant (project 91617107), a ZonMW Enabling Technology Hotels grant (project 435002030), and received support from the Dutch Arthritis Foundation (project 15-2-402). MK received support from the Prof. Steenhuis Fund. This work was carried out on the Dutch national e-infrastructure with support of the SURF Cooperative.

- Elliott SE, Kongpachith S, Lingampalli N, Adamska JZ, Cannon BJ, Mao R, et al. Affinity maturation drives epitope spreading and generation of proinflammatory anti-citrullinated protein antibodies in rheumatoid arthritis. *Arthritis Rheumatol.* (2018) 70:1946–58. doi: 10.1002/art.40587
- Lu DR, McDavid AN, Kongpachith S, Lingampalli N, Glanville J, Ju CH, et al. T cell-dependent affinity maturation and innate immune pathways differentially drive autoreactive B cell responses in rheumatoid arthritis. *Arthritis Rheumatol.* (2018) 70:1732–44. doi: 10.1002/art.40578
- Corsiero E, Bombardieri M, Carloti E, Pratesi F, Robinson W, Migliorini P, et al. Single cell cloning and recombinant monoclonal antibodies generation from RA synovial B cells reveal frequent targeting of citrullinated histones of NETs. *Ann Rheum Dis.* (2016) 75:1866–75. doi: 10.1136/annrheumdis-2015-208356
- Rombouts Y, Willemze A, van Beers JJ, Shi J, Kerkman PF, van Toorn L, et al. Extensive glycosylation of ACPA-IgG variable domains modulates binding to citrullinated antigens in rheumatoid arthritis. *Ann Rheum Dis.* (2016) 75:578–85. doi: 10.1136/annrheumdis-2014-206598
- Hafkenschied L, Bondt A, Scherer HU, Huizinga TW, Wuhler M, Toes RE, et al. Structural analysis of variable domain glycosylation of anti-citrullinated protein antibodies in rheumatoid arthritis reveals the presence of highly sialylated glycans. *Mol Cell Proteomics.* (2017) 16:278–87. doi: 10.1074/mcp.M116.062919
- Kempers AC, Hafkenschied L, Dorjee AL, Moutousidou E, van de Bovenkamp FS, Rispens T, et al. The extensive glycosylation of the ACPA variable domain observed for ACPA-IgG is absent from ACPA-IgM. *Ann Rheum Dis.* (2018) 77:1087–8. doi: 10.1136/annrheumdis-2017-211533
- van de Bovenkamp FS, Derksen NIL, Ooijevaar-de Heer P, van Schie KA, Kruijthof S, Berkowska MA, et al. Adaptive antibody diversification through N-linked glycosylation of the immunoglobulin variable region. *Proc Natl Acad Sci USA.* (2018) 115:1901–6. doi: 10.1073/pnas.1711720115
- Vergoosen RD, Slot LM, Hafkenschied L, Koning MT, van der Voort EIH, Grooff CA, et al. B-cell receptor sequencing of anti-citrullinated protein antibody (ACPA) IgG-expressing B cells indicates a selective advantage for the introduction of N-glycosylation sites during somatic hypermutation. *Ann Rheum Dis.* (2018) 77:956–8. doi: 10.1136/annrheumdis-2017-212052
- Lloyd KA, Steen J, Amara K, Titcombe PJ, Israelsson L, Lundstrom SL, et al. Variable domain N-linked glycosylation and negative surface charge are key features of monoclonal ACPA: implications for B-cell selection. *Eur J Immunol.* (2018) 48:1030–45. doi: 10.1002/eji.201747446
- Suwanalai P, Scherer HU, van der Woude D, Ioan-Facsinay A, Jol-van der Zijde CM, van Tol MJ, et al. Anti-citrullinated protein antibodies have a low avidity compared with antibodies against recall antigens. *Ann Rheum Dis.* (2011) 70:373–9. doi: 10.1136/ard.2010.135509
- Verpoort KN, Jol-van der Zijde CM, Papendrecht-van der Voort EA, Ioan-Facsinay A, Drijfhout JW, van Tol MJ, et al. Isotype distribution of anti-cyclic citrullinated peptide antibodies in undifferentiated arthritis and rheumatoid arthritis reflects an ongoing immune response. *Arthritis Rheum.* (2006) 54:3799–808. doi: 10.1002/art.22279

26. Reed E, Jiang X, Kharlamova N, Ytterberg AJ, Catrina AI, Israelsson L, et al. Antibodies to carbamylated alpha-enolase epitopes in rheumatoid arthritis also bind citrullinated epitopes and are largely indistinct from anti-citrullinated protein antibodies. *Arthritis Res Ther.* (2016) 18:96. doi: 10.1186/s13075-016-1001-6
27. Victora GD, Schwickert TA, Fooksman DR, Kamphorst AO, Meyer-Hermann M, Dustin ML, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell.* (2010) 143:592–605. doi: 10.1016/j.cell.2010.10.032
28. Yeap LS, Hwang JK, Du Z, Meyers RM, Meng FL, Jakubauskaite A, et al. Sequence-intrinsic mechanisms that target AID mutational outcomes on antibody genes. *Cell.* (2015) 163:1124–37. doi: 10.1016/j.cell.2015.10.042
29. Koning MT, Kielbasa SM, Boersma V, Buermans HP, van der Zeeuw SA, van Bergen CA, et al. ARTISAN PCR: rapid identification of full-length immunoglobulin rearrangements without primer binding bias. *Br J Haematol.* (2017) 178:983–6. doi: 10.1111/bjh.14180
30. Giudicelli V, Brochet X, Lefranc MP. IGMT/V-QUEST: IGMT standardized analysis of the immunoglobulin (IG) and T cell receptor (TR) nucleotide sequences. *Cold Spring Harb Protoc.* (2011) 2011:695–715. doi: 10.1101/pdb.prot5633
31. IJspeert H, van Schouwenburg PA, van Zessen D, Pico-Knijnenburg I, Stubbs AP, van der Burg M. Antigen receptor galaxy: a user-friendly, web-based tool for analysis and visualization of T and B cell receptor repertoire data. *J Immunol.* (2017) 198:4156–65. doi: 10.4049/jimmunol.1601921
32. Hershberg U, Uduman M, Shlomchik MJ, Kleinstein SH. Improved methods for detecting selection by mutation analysis of Ig V region sequences. *Int Immunol.* (2008) 20:683–94. doi: 10.1093/intimm/dxn026
33. Shlomchik MJ, Watts P, Weigert MG, Litwin S. Clone: a Monte-Carlo computer simulation of B cell clonal expansion, somatic mutation, and antigen-driven selection. *Curr Top Microbiol Immunol.* (1998) 229:173–97. doi: 10.1007/978-3-642-71984-4_13
34. Sabouri Z, Schofield P, Horikawa K, Spierings E, Kipling D, Randall KL, et al. Redemption of autoantibodies on anergic B cells by variable-region glycosylation and mutation away from self-reactivity. *Proc Natl Acad Sci USA.* (2014) 111:E2567–75. doi: 10.1073/pnas.1406974111
35. Koers J, Derksen NIL, Ooijevaar-de Heer P, Nota B, van de Bovenkamp FS, Vidarsson G, et al. Biased N-glycosylation site distribution and acquisition across the antibody V region during B cell maturation. *J Immunol.* (2019) 202:2220–8. doi: 10.4049/jimmunol.1801622
36. Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum.* (2004) 50:380–6. doi: 10.1002/art.20018
37. Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum.* (2003) 48:2741–9. doi: 10.1002/art.11223
38. van Zanten A, Arends S, Roozendaal C, Limburg PC, Maas F, Trouw LA, et al. Presence of anticitrullinated protein antibodies in a large population-based cohort from the Netherlands. *Ann Rheum Dis.* (2017) 76:1184–90. doi: 10.1136/annrheumdis-2016-209991
39. Tasliyurt T, Kisacik B, Kaya SU, Yildirim B, Pehlivan Y, Kutluturk F, et al. The frequency of antibodies against cyclic citrullinated peptides and rheumatoid factor in healthy population: a field study of rheumatoid arthritis from northern turkey. *Rheumatol Int.* (2013) 33:939–42. doi: 10.1007/s00296-012-2458-5
40. Ten Brinck RM, van Steenberg HW, van Delft MAM, Verheul MK, Toes REM, Trouw LA, et al. The risk of individual autoantibodies, autoantibody combinations and levels for arthritis development in clinically suspect arthralgia. *Rheumatology (Oxford).* (2017) 56:2145–53. doi: 10.1093/rheumatology/kex340
41. van Steenberg HW, Mangnus L, Reijnierse M, Huizinga TW, van der Helm-van Mil AH. Clinical factors, anticitrullinated peptide antibodies and MRI-detected subclinical inflammation in relation to progression from clinically suspect arthralgia to arthritis. *Ann Rheum Dis.* (2016) 75:1824–30. doi: 10.1136/annrheumdis-2015-208138
42. Terao C, Ohmura K, Ikari K, Kawaguchi T, Takahashi M, Setoh K, et al. Effects of smoking and shared epitope on the production of anti-citrullinated peptide antibody in a Japanese adult population. *Arthritis Care Res.* (2014) 66:1818–27. doi: 10.1002/acr.22385
43. Hensvold AH, Magnusson PK, Joshua V, Hansson M, Israelsson L, Ferreira R, et al. Environmental and genetic factors in the development of anticitrullinated protein antibodies (ACPAs) and ACPA-positive rheumatoid arthritis: an epidemiological investigation in twins. *Ann Rheum Dis.* (2015) 74:375–80. doi: 10.1136/annrheumdis-2013-203947
44. Hafkenscheid L, de Moel E, Smolik I, Tanner S, Meng X, Jansen BC, et al. N-linked glycans in the variable domain of ACPA-IgG predict the development of rheumatoid arthritis. *Arthritis Rheumatol.* (2019). doi: 10.1002/art.40920. [Epub ahead of print].
45. Kempers AC, Hafkenscheid L, Scherer HU, Toes REM. Variable domain glycosylation of ACPA-IgG: a missing link in the maturation of the ACPA response? *Clin Immunol.* (2018) 186:34–7. doi: 10.1016/j.clim.2017.09.001
46. Hamza N, Hershberg U, Kallenberg CG, Vissink A, Spijkervet FK, Bootsma H, et al. Ig gene analysis reveals altered selective pressures on Ig-producing cells in parotid glands of primary Sjogren's syndrome patients. *J Immunol.* (2015) 194:514–21. doi: 10.4049/jimmunol.1302644
47. Zhu D, Ottensmeier CH, Du M-Q, McCarthy H, Stevenson FK. Incidence of potential glycosylation sites in immunoglobulin variable regions distinguishes between subsets of Burkitt's lymphoma and mucosa-associated lymphoid tissue lymphoma. *Br J Haematol.* (2003) 120:217–22. doi: 10.1046/j.1365-2141.2003.04064.x
48. Radcliffe CM, Arnold JN, Suter DM, Wormald MR, Harvey DJ, Royle L, et al. Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. *J Biol Chem.* (2007) 282:7405–15. doi: 10.1074/jbc.M602690200
49. McCann KJ, Ottensmeier CH, Callard A, Radcliffe CM, Harvey DJ, Dwek RA, et al. Remarkable selective glycosylation of the immunoglobulin variable region in follicular lymphoma. *Mol Immunol.* (2008) 45:1567–72. doi: 10.1016/j.molimm.2007.10.009
50. Culver EL, van de Bovenkamp FS, Derksen NIL, Koers J, Cargill T, Barnes E, et al. Unique patterns of glycosylation in immunoglobulin subclass G4-related disease and primary sclerosing cholangitis. *Ital J Gastroenterol Hepatol.* (2018). doi: 10.1111/jgh.14512. [Epub ahead of print].

Conflict of Interest Statement: RT and HS are co-authors on a provisional patent application based on research related to topics discussed in this study.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Vergroesen, Slot, van Schaik, Koning, Rispen, van Kampen, Toes and Scherer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.