Published in final edited form as:

Mol Immunol. 2016 August ; 76: 123–133. doi:10.1016/j.molimm.2016.07.004.

Uncommon Structural Motifs Dominate the Antigen Binding Site in Human Autoantibodies Reactive with Basement Membrane Collagen

Mary H. Foster^{a,b,c}, Elizabeth S. Buckley^a, Benny J. Chen^{a,c}, Kwan-Ki Hwang^{a,d}, and Amy G. Clark^{a,b}

Mary H. Foster: mary.h.foster@dm.duke.edu; Elizabeth S. Buckley: elizabeth.buckley@duke.edu; Benny J. Chen: benny.chen@dm.duke.edu; Kwan-Ki Hwang: kwan-ki.hwang@dm.duke.edu; Amy G. Clark: amy.clark@duke.edu

^aDepartment of Medicine, Duke University Medical Center, Durham, N.C., USA

^bDurham VA Medical Center, Durham, N.C., USA

^cDuke Cancer Institute, Duke University Medical Center, Durham, N.C., USA

^dDuke Human Vaccine Institute, Duke University Medical Center, Durham, N.C., USA

Abstract

Autoantibodies mediate organ destruction in multiple autoimmune diseases, yet their origins in patients remain poorly understood. To probe the genetic origins and structure of disease-associated autoantibodies, we engrafted immunodeficient mice with human CD34⁺ hematopoietic stem cells and immunized with the non-collagenous-1 (NC1) domain of the $\alpha 3$ chain of type IV collagen. This antigen is expressed in lungs and kidneys and is targeted by autoantibodies in anti-glomerular basement membrane (GBM) nephritis and Goodpasture syndrome (GPS), prototypic human organ-specific autoimmune diseases. Using Epstein Barr virus transformation and cell fusion, six human anti- $\alpha 3$ (IV)NC1 collagen monoclonal autoantibodies (mAb) were recovered, including subsets reactive with human kidney and with epitopes recognized by patients' IgG. Sequence analysis reveals a long to exceptionally long heavy chain complementarity determining region3 (HCDR3), the major site of antigen binding, in all six mAb. Mean HCDR3 length is 25.5 amino acids (range 20–36), generated from inherently long DH and JH genes and extended regions of non-templated N-nucleotides. Long HCDR3 are suited to forming noncontiguous antigen contacts and to binding recessed, immunologically silent epitopes hidden from conventional antibodies, as seen with self-antigen crossreactive broadly neutralizing anti-HIV Ig (bnAb). The anti- $\alpha 3$ (IV)NC1 collagen mAb also show preferential use of unmutated variable region genes that are enriched among human chronic lymphocytic leukemia antibodies that share features with natural polyreactive Ig. Our findings suggest unexpected relationships between pathogenic anti-collagen Ig, bnAb, and autoreactive Ig associated with malignancy, all of which arise from B cells

Address Correspondence: Mary H. Foster, Dept of Medicine, Duke University Medical Center, Box 103015, Durham, NC 27710, mary.h.foster@dm.duke.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

expressing unconventional structural elements that may require transient escape from tolerance for successful expansion.

Keywords

Anti-collagen autoimmunity; HCDR3 (heavy chain complementarity determining region 3); Human monoclonal autoantibody

1. Introduction

Autoimmune diseases affect an estimated 10–20% of the US population, a prevalence that equals that of cancer and heart disease. Autoimmunity often strikes young adults, destroying joints, kidneys, lungs, and other organs and requiring medications that can have devastating long-term effects. Kidney involvement is common and particularly disabling, as it often leads to renal failure requiring dialysis or transplantation. Up to 70–80% of patients with systemic lupus erythematosus or anti-neutrophil cytoplasmic antibody vasculitis develop glomerulonephritis requiring immunosuppression (Pagnoux et al., 2008). Humoral autoimmunity is often prominent: autoantibodies are key to diagnosis and organ destruction, and their elimination is a major goal of therapy. Adverse side effects of current interventions have spurred efforts to develop better targeted, less toxic treatments. These efforts would be aided by insight into the origin and structure of the pathogenic autoantibodies, and of their transmembrane counterparts on the surface of autoreactive B cells. Sequence information could refine our understanding of autoantibody-antigen (Ag) interactions, and provide clues to the origin, regulation, and evolution of the autoimmune response in patients.

Progress on this front has been slow. Antibodies and lymphocyte receptors are structurally complex and extremely diverse, generated by multiple gene rearrangements and enzymatic and somatic processes. This complexity renders isolated proteomic dissection of serum polyclonal antibody mixtures difficult. Proteomic approaches are further complicated by the uniqueness of individual Ig repertoires, particularly the critical heavy chain complementarity determining region 3 (HCDR3) that dominates Ag binding; subject-specific reference libraries are required for Ig proteomics data interpretation. Attempts to isolate Ag-specific Ig sequences from circulating B cells are stymied by the rarity of Ag-specific B cells in peripheral blood (Stevens et al., 1979). Immunization studies reveal a brief circulation half-life for Ag-specific B cells, which rapidly migrate between relatively inaccessible niches in the spleen or other peripheral lymphoid organs and the bone marrow, where they differentiate into resident plasma cells that are the source of the vast majority of serum Ig (Reddy et al., 2010). This may explain difficulties in recovering Ag-specific human monoclonal Ab (mAb) by viral transformation of autoimmune patients' blood cells, and in any case dissociates specificity of serum autoantibodies from that of contemporaneous circulating cells. Moreover, for many autoimmune diseases for which Ab and receptors recognize conformational (non-linear) epitopes, reagents or assays are not yet available that permit capture of Ag-specific B cells for study.

To circumvent these obstacles, we turned to a humanized model as a platform to probe origins of human pathogenic autoimmunity at the molecular level, and to sample and capture

human Ag-reactive cells and autoantibodies. We focus on anti-glomerular basement membrane (GBM) nephritis, and its systemic counterpart, Goodpasture Syndrome (GPS), because they have long been recognized as prototypic organ-specific autoimmune diseases. The autoantigen, alpha3(IV)NC1 collagen, is well characterized and autoantibodies play a prominent role in organ injury (reviewed in (Borza and Hudson, 2003). Patient IgG can transfer nephritis to susceptible animal hosts, such as squirrel monkeys, in whom pathogenic epitopes are accessible in vivo (Lerner et al., 1967; Luo et al., 2010), and glomerulonephritis rapidly recurs in transplant recipients with persistent circulating anti-GBM IgG (Wilson and Dixon, 1973). Linear Ig deposits along the glomerular and/or alveolar basement membranes are a hallmark of disease, and either this histopathological finding or detection of circulating anti-alpha3(IV)NC1 collagen Ig is required for diagnosis. Serum autoantibodies levels guide treatment decisions, including duration of plasmapheresis and timing of transplantation (Hellmark and Segelmark, 2014).

Herein, we report isolation and sequence analysis of six human mAb reactive with alpha3(IV)NC1 collagen, subsets of which bind human kidney GBM and alpha3(IV)NC1 collagen epitopes recognized by nephritis patients' IgG. These mAb derive from mice in which a human immune system was established by engraftment with cord blood-derived human CD34+ hematopoietic stem cells (HSC), followed by immunization with purified alpha3(IV)NC1 collagen. Sequence analysis reveals unusual features shared by the six mAbs, including strikingly long HCDR3s reminiscent of those reported among protective broadly neutralizing mAb (bnAb) reactive with human immunodeficiency virus (HIV), another Ab specificity rarely encountered in human serum, even among HIV-infected individuals. The results provide previously unattainable insight into structural motifs of human autoreactivity with basement membrane collagen, and clues into origins and regulation of enigmatic B cells autoreactive with basement membrane collagen.

2. Materials and Methods

2.1. Mouse studies and immunizations

Female NOD-scid-gamma (NSG) immunodeficient mice were purchased from the Duke Cancer Center Isolation Facility Breeding Core and housed under specific pathogen free conditions. Hu-HSC mice were generated essentially as described (Worni-Schudel et al., 2015), with the following modifications; 3–4 hours after irradiation (100 cGy, X-RAD 320, Precision X-ray, North Branford, CT) 9.5–10 week old NSG recipients were intravenously injected with purified human CD34+ HSC (AllCells LLC, Alameda, CA), using $2.5\text{--}10 \times 10^4$ CD34+ HSC per mouse (Pearson et al., 2008; Shultz et al., 2005). All studies were approved by the Duke University Institutional Review Board and Animal Care and Use Committee and conform to institutional standards and to the National Institutes of Health guide for the care and use of Laboratory animals. 14 weeks after HSC injection, Hu-HSC mice were immunized with 25 micrograms bovine alpha3(IV)NC1 collagen (Eurodiagnostica, Malmö, Sweden) emulsified in Freund's Complete Adjuvant, followed by boost 3 weeks later with antigen in Freund's Incomplete Adjuvant. Tissue was harvested at 1.5–3 weeks following this boost (19–20 weeks post-engraftment).

2.2. Flow cytometry

Single-cell suspensions of cells were analyzed by flow cytometry as described (Clark et al., 2011; Worni-Schudel et al., 2015), using fluorescent-conjugated antibodies (Becton Dickinson-Pharmingen, San Jose, CA). Cells were analyzed on a FACSCalibur or BD CANTO apparatus (Becton Dickinson-Pharmingen) and results analyzed using FlowJo (Treestar, Ashland, OR, USA), with live cells gated on forward and side scatter. Chimerism, or the level of human lymphocyte engraftment, in Hu-HSC mouse tissue was calculated as follows: % Chimerism = (%humanCD45+ cells ×100)/(%humanCD45+ cells + %mouseCD45+ cells).

2.3. Cell culture, Epstein Barr virus (EBV) transformation, and hybridoma generation

Harvested human B cells were transformed and fused essentially as described (Bonsignori et al., 2011; Worni-Schudel et al., 2015). Briefly, cells were cultured in bulk overnight in complete medium supplemented with Cyclosporin A, CpG oligonucleotides, checkpoint 2 kinase inhibitor, and EBV B95-8 virus suspension (ATCC, Manassas, VA, USA), and then seeded into 96-well plates and refed weekly, with cyclosporin A and CpG oligonucleotides withdrawal at two weeks. Transformed B cell lines for which supernatants tested consistently positive for reactivity with antigen were electrofused with HMMA2.5 mouse-human chimera myeloma cells (gift of Lisa Cavacini, Beth Israel Deaconess Medical Center, Boston) (Posner et al., 1987) using a PA-4000/PA-101S apparatus and hypoxanthine, aminopterin, thymidine and ouabain selection, as described (Yu et al., 2008). Antigen-reactive lines were subcloned using a standard limiting dilution method.

2.4. Enzyme Linked Immunosorbent Assay (ELISA)

Human immunoglobulin (Ig) class, light chain type, Ig concentration, and anti-alpha3(IV)NC1 collagen Ab in serum and cell supernatants were detected by ELISA, as described (Worni-Schudel et al., 2015; Zhang et al., 2008). Results for binding to antigen [bovine alpha3(IV)NC1 collagen (Eurodiagnostica, Malmö, Sweden)] were recorded as OD on antigen minus OD on wells coated with diluent only (6M guanidine-HCL), using as controls mouse anti-alpha3(IV)NC1 collagen IgG (Mab3) (Eurodiagnostica) and Goodpasture patient serum. For inhibition assays, IgG was purified from patient plasma (previously collected under an IRB approved protocol) (Worni-Schudel et al., 2015) using HiTrap Protein G HP columns (GE Healthcare, Pittsburgh, PA). Human IgM mAb was preincubated with varying concentrations of inhibitor (purified patient IgG), and competition for alpha3(IV)NC1 collagen binding assessed by ELISA.

Binding to single-stranded DNA (ssDNA) was detected by coating Immulon II HB 96-well plates (Thermo Scientific, Waltham, MA) at 2.25 microgram per well with ssDNA, produced by boiling deproteinized calf thymus DNA (Worthington Biochemical, Lakewood NJ). Positive control antibody was H241 IgG mAb (Ali et al., 1985). Binding to myeloperoxidase (MPO) was detected by coating Immulon II HB plates with human leukocyte MPO (Sigma Aldrich, St. Louis, MO) in 1X PBS at 0.001 units per well. MAb anti-MPO clone CLB-MPO-1/1 (Sigma Aldrich) served as positive control. Blocking, human mAb incubation, and detection steps were as described above. Binding to cardiolipin was detected using the

QUANTA Lite ACA IgM III kit (InovaDiagnostics, San Diego, CA), following manufacturer's instructions.

2.5. Immunohistochemistry

Hu-HSC mouse kidneys were immersed in Optimal Cutting Temperature (OCT) embedding compound and snap frozen. Two de-identified normal human kidneys, frozen in OCT blocks, were obtained from the Duke Biospecimen Repository and Processing Core. Sections were cut by the Duke Surgical Sciences Histopathology Lab Core Facility, fixed in acetone, and blocked in PBS containing 3% bovine serum albumin and 3% normal goat serum. Human kidney sections were incubated overnight with mAb-containing cell culture supernatants or human IgM control (Thermo Fisher Scientific, Waltham, MA) at 4°C in a humid chamber. Human Ig in Hu-HSC mouse kidneys (direct IF) and on human kidney (indirect IF) was detected using goat-anti-human IgM-Alexa Fluor488 antibodies (Thermo Fisher Scientific), and images acquired using NIS-Elements software (Nikon USA). Anti-nuclear Abs (ANA) were determined by indirect immunofluorescence of fixed HEp-2 cells (ANA Hep-2 Test System, Zeus Scientific, Branchburg, NJ, USA) following the manufacturer's instructions and using mAb at a concentration of 1 µg/mL. Stained slides were imaged as described above.

2.6. Analysis of human anti-alpha3(IV)NC1 mAb variable region genes

RNA was isolated from subcloned EBV-transformed HMMA-fused B cells producing anti-alpha3(IV)NC1 collagen Ig and converted to cDNA as described (Worni-Schudel et al., 2015). Productively rearranged Ig heavy chain (HC) and light chain (LC) variable (V) region genes were amplified by PCR in chain-specific reactions, using published panels of 5' primers for human VH (VH1-Ext – VH6-Ext), Vlambda (Vlambda1-Ext – Vlambda10-Ext), and (Vk) (Vk1/2-Ext - Vk6/7-Ext) subgroups paired with a 3' primer for IgM (IgM-ext) and the lambda (Cλ–new-ext) and kappa (Cκ-ext) light chain constant regions, respectively (Liao et al., 2009), with the exception of mAb C HC, for which a different set of VH primers (Ippolito et al., 2012) was used to obtain an appropriate size PCR band for sequencing. PCR products were purified using the DNA Clean and Concentrator kit (Genesee Scientific, San Diego, CA, USA) and sequence determined by the Duke University DNA Analysis facility. Sequences were analyzed and compared using the BLAST software and databases (blast.ncbi.nlm.nih.gov/Blast.cgi) and the International ImMunoGeneTics Information System (IMGT) public Ig gene database (www.imgt.org) with the IMGT/VQUEST tool (Giudicelli et al., 2004) integrated with IMGT/JunctionAnalysis (Yousfi Monod et al., 2004).

2.7. Statistical analyses

Data are shown as mean ± standard deviation (SD), unless otherwise indicated. Comparisons between two groups were performed using a nonparametric Wilcoxon rank-sum test and JMP (SAS Institute, Cary, NC) software, with $p < 0.05$ considered significant.

3. Results

3.1. Recovery of human B cells and anti- $\alpha 3(\text{IV})\text{NC1}$ collagen Ab from immunized Hu-HSC mice

Human CD34⁺ HSC were injected into 11 conditioned adult NSG mice to establish a human lymphoid system in vivo. 5 received HSC derived from a single donor and 6 received HSC pooled from 3 donors. At 3 months post HSC injection, 8 recipients (4 each having received single or mixed donor HSC) demonstrated evidence of human cell engraftment in peripheral blood (mean blood chimerism 15.0%) and were immunized with $\alpha 3(\text{IV})\text{NC1}$ collagen at 14 and 17 weeks post-engraftment. At tissue harvest at 19–20.5 weeks post-engraftment, chimerism in spleen was higher in recipients of single donor HSC than for recipients of mixed HSC (mean 46.0% vs 10.8%, $n=4/\text{group}$, with range 1.9% to 80.6%) (Fig. 1A–D). Of human CD45⁺ cells in spleen, B cells and T cells constituted mean 85.8% and 3.5%, respectively (Fig. 1E and not shown). The total number of human B cells in spleen was mean $1,497 \pm 3,020 \times 10^3$ (range $7 \times 10^3 - 8,620 \times 10^3$) (Fig. 1F). Plasma levels of human IgM paralleled splenic population by human B cells ($R^2=0.64$, $p=0.016$), with higher concentrations in recipients of single donor cells (mean $31.1 \pm 19.8 \mu\text{g/ml}$ human IgM vs 4.9 ± 4.8 for recipients of mixed donor cells). Plasma levels of human IgG were low in all mice, mean $0.08 \pm 0.05 \mu\text{g/ml}$, range 0.02–0.14 $\mu\text{g/ml}$). Direct immunofluorescence examination of kidneys of immunized Hu-HSC mice revealed scattered deposits of human IgM, with faint linear capillary wall staining in some glomeruli, whereas no human Ig was detected in kidneys of unimmunized Hu-HSC mice (Fig. 2A). Ag-reactive Ig were not detected in serum.

At 4–7 weeks after splenocyte culture with CpG oligonucleotides and EBV (mean 66.8×10^3 splenocytes per well, range $37.5 - 155.6 \times 10^3$), 134 (44.7%) of 300 wells had cell clusters consistent with transformed human B cells, and 26 (19.4%) of 134 wells that showed cell growth contained Ag-reactive human Ig. Six transformed human B cell lines that consistently produced Ag-reactive Ig with the high ODs and maintained favorable growth characteristics were derived from 3 immunized mice, and were expanded and successfully electrofused. Five lines were derived from single donor HSC, and one line from pooled HSC. All lines produced human IgM. Resulting hybridoma lines were subsequently subcloned by limiting dilution, and binding of human mAb to $\alpha 3(\text{IV})\text{NC1}$ collagen confirmed in ELISA (Fig 2B.).

Reactivity of the six mAb with human GBM was assayed using indirect immunohistochemistry by incubation of mAb on normal human kidney sections, obtained from two donors. Linear GBM staining was detected for a subset of mAb (Fig. 2C), variably accompanied by patchy linear staining of tubular BM (not shown). Competition ELISA was used to assess reactivity of the human mAb with pathogenic GP epitopes, defined as those epitopes recognized by patients' IgG. Inhibition of mAb binding to $\alpha 3(\text{IV})\text{NC1}$ collagen by patients' circulating IgG was demonstrated for a subset of mAb (Fig. 2D).

3.2. Sequence analysis of human anti- α 3(IV)NC1 collagen mAb

The productively rearranged Ig HC and LC V region sequences for each human mAb were determined by sequence analysis after RT-PCR of cDNA. The identified genes, putative germline gene segments, and HC and LC deduced amino acid (a.a.) sequences were determined using IMGT search tools and key features are shown in Table 1 and Fig. 3. Full Ab sequences are available in the DDBJ/EMBL/GenBank nt databases under the accession numbers KU212251-KU212260.

The mAb A HC is encoded by an IGHV1–8 gene with 100% identity to an unmutated IGHV1–8*01 allele. The V region sequence has a single nucleotide (nt) difference from IGHV1–8*02 gene in CDR1 that is silent. Review of the databases reveals that IGHV1–8 genes are enriched among human mantle cell lymphomas. The mAb A HC CDR3 is encoded by 3 P-nt, 3 N1-region nontemplated (somatically generated) nt, 22 nt contributed by an unmutated IGHD3–3 gene, allele IGHD3–3*01, translated in reading frame 1, and 15 N2-region nt (Fig. 3), in addition to unmutated IGHV1–8*01 allele and an unmutated IGHJ6 gene, allele IGHJ6*02, that also encodes framework (FR) 4.

The mAb A kappa LC uses an unmutated IGKV2–28 gene, with 100% identity to both the IGKV2–28*01 and IGKV2D-28*01 alleles, which have identical nt sequences in the expressed region. MAb A also uses an unmutated IGKJ1 gene, with 100% match to the IGKJ1*01 allele.

The mAb B HC is encoded by an IGHV3–48 gene with 100% identity to an unmutated IGHV3–48*04 allele. The V region sequence has a single nt difference from IGHV3–48*01 and IGHV3–48*02 genes in FR3 that is silent. Review of the databases reveals that IGHV3–48 genes are enriched among human chronic lymphocytic leukemias (CLL), and have prognostic value in this disease. The HC CDR3 is encoded by 22 N1-region and 12 N2-region nt, 37 nt contributed by an unmutated IGHD3–16 gene, allele IGHD3–16*02, translated in reading frame 2, and 3 P-nts bookending the D-gene contributions (Fig. 3). The 3' end of HCDR3 is encoded by an unmutated IGHJ6 gene, allele IGHJ6*02.

The mAb B lambda LC uses an IGLV1–40 gene, with 100% identity to both the IGLV1–40*01 and IGLV1–40*02 alleles, which have identical nt sequences in the expressed region. MAb B also uses an unmutated IGLJ1 gene, with 100% match to the IGLJ1*01 allele.

The mAb C HC is encoded by an IGHV3–21 gene. Available sequence (3 codons of FR1 through FR3) has 100% identity to unmutated IGHV3–21*01 and IGHV3–21*02 alleles. Review of the databases reveals that IGHV3–21 genes are also enriched among CLL, in which they are associated with poor prognosis (Dal-Bo et al., 2011). The mAb C HCDR3 is encoded by 10 N1-region and 1 N2-region nts, 30 nts contributed by an unmutated IGHD3–16 gene, allele IGHD3–16*02, translated in reading frame 2, and 3 P-nts, as well as an unmutated IGHJ4 gene, allele IGHJ4*02. The current panel of LC primers does not yield a detectable PCR product for the mAb C LC, suggesting use of a rare V-gene segment or polymorphic variant not captured by the primer panel.

The mAb D HC is encoded by an IGHV3–23 gene with 100% identity to three unmutated alleles (IGHV3–23*01, IGHV3–23*04, and IGHV3–23D*01) that have identical nt sequences in the expressed region. Review of the databases reveals that IGHV3–23 genes are enriched among mantle cell lymphomas (a feature shared with IGHV1–8 genes), and are among the most common IGHV genes expressed in CLL. The mAb D HCDR3 is encoded by 1 P-nt, 2 N1-region and 11 N2-region nts, 24 nts contributed by an unmutated IGHD3–3 gene, allele IGHD3–3*01, translated in reading frame 2. Additional nts at either end of HC CDR3 can be accounted for by the unmutated IGHV3–23 gene and an unmutated IGHJ6 gene, allele IGHJ6*02.

The mAb D kappa LC uses a minimally mutated IGKV1–8 gene, with a single nt difference from the IGKV1–8*01 allele. This results in a conservative hydrophobic Phe for Leu a.a. replacement in codon 11 of FR1. MAb D also uses an unmutated IGKJ1 gene, with 100% match to the IGKJ1*01 allele.

The mAb E HC is encoded by an IGHV1–69 gene with 99% identity in the coding region to alleles IGHV1–69*01 (previously VH1/DP-10) and IGHV1–69D*01. The V region sequence has a single nt difference in FR3 that leads to a conservative mutation (valine for alanine) from germline. IGHV1–69 genes have been associated with rheumatoid factor activity and with several B cell proliferative diseases, including the unmutated subset of CLL associated with a poor prognosis, unusually long HCDR3, and polyreactivity for infectious agents including HIV (Forconi et al., 2010; Hwang et al., 2014; Johnson et al., 1997). MAb E uses the D3–3*01 allele, also known as DXP4, preferentially used by V1–69 expressing CLL B cells. The mAb E HCDR3 is encoded by 6 N1-region and 12 N2-region nts, and 26 nts contributed by an unmutated IGHD3–3 gene, allele IGHD3–3*01, translated in reading frame 2 (Fig. 3), as well as a short segment of an unmutated IGHJ4 gene, allele IGHJ4*02.

The mAb E lambda LC uses an IGLV1–40 gene, with 100% identity to both the IGLV1–40*01 and IGLV1–40*02 alleles, which have identical nt sequences within the sequenced region. MAb E also uses an unmutated IGLJ3 gene, with 100% match to the IGLJ3*02 allele. There is a single nt difference in the J-gene-encoded region from unmutated IGLJ2*01 and IGLJ3*01 alleles that is silent.

The mAb G HC is encoded by an IGHV4–61 gene with 100% identity to an unmutated IGHV4–61*02 allele. IGHV4–61 is overrepresented among cerebrospinal fluid B cells isolated from patients with multiple sclerosis (von Budingen et al., 2012). The HCDR3 is encoded by 18 N1-region and 19 N2-region nts, 10 nts contributed by an unmutated IGHD2–15 gene, allele IGHD2–15*01, translated in reading frame 2, and a single P-nt (Fig. 3), as well as an unmutated IGHJ6 gene, allele IGHJ6*02. The mAb G uses a lambda LC based on ELISA isotyping. Similar to mAb C, the current panel of universal Vlambda primers does not yield a detectable PCR product (LC band) for this mAb.

Analysis of HCDR3 of the 6 human anti-alpha3(IV)NC1 collagen mAb reveals several striking similarities. All mAb express exceptionally long HCDR3, with a mean length of 25.5 a.a. (range 20–36) (Table 1 and Figure 3). This compares to an average human Ig

HCDR3 length of 15.2 a.a. (IMGT designation) (Zemlin et al., 2003). The mean length of the D-gene contribution to HCDR3 in our mAb is 24.8 nt (range 10–37 nt), compared to 14.3 nt among productive IgH rearrangements in adult peripheral blood naïve B cells (Larimore et al., 2012). Human full-length genomic IGHD genes range from 9 to 36 nt (Larimore et al., 2012). All of our mAb express IGHD genes from the intrinsically longest human D-gene families (D2 and D3), including the longest known D-gene, IGHD3–16, in mAb B and C. HCDR3 length is also extended by limited “trimming” of gene segments at the V-D and D-J junctions, and somatic insertion of long strings of non-DNA-templated (N-region) nt or palindromic (P) nt on one or both sides of the D-gene joints (Fig. 3). The mAb HCDR3 include mean 21.8 total N-region nt, including mean 10.2 nt (range 2–22) at the V-D junction and mean 11.7 nt (range 1–19) at the D-J junction. This compares to a mean 13.4 total N-region nt in productive IgH rearrangements in the large dataset of Larimore and colleagues (mean 7.3 nt at V-D and mean 6.1 nt at D-J junctions) (Larimore et al., 2012). Several mAb HCDR3 (A, B, and G) also retain the extended 5' region of the IGHJ6 gene that encodes a polyTyr motif. A partial contribution from unconventional VDJ rearrangements cannot be ruled out, as several short nt sequences within N-regions could conceivably represent a D-D rearrangement or VH replacement footprint (Huang et al., 2014).

There is evidence of preferential use of D and J gene segments. 5 of 6 mAb are encoded by an IGHD-gene from the D3 subgroup (Table 1), and all 5 mAb have a centrally positioned Trp (W) residue (Fig. 3), suggesting an important role for this residue in ligand binding. For 4 of these mAb (B, C, D and E), the IGHD3 allele (either D3-3*01 or D3-16*02) is read in the second reading frame (RF2), yielding similar central D-gene-derived peptides - (Y)-Y-D-Y/F-(V)-W-G/S-S/G-Y – that are highly enriched in aromatic amino acids: 46.2% to 62.5% of D-gene-encoded a.a. are Tyr (Y), Phe (F), or Trp (W). For mAb A, the D3-3*01 allele is read in RF1, yielding a rare D-gene-encoded heptapeptide (L-R-F-L-E-W-L) enriched in highly hydrophobic leucine (L) residues. Four of the 6 mAb are encoded by the J-gene J6*02 allele, the extended 5' end of which can contribute up to 5 contiguous tyrosine (Tyr) residues to the 3' half of HCDR3 (Fig. 3). The HCDR3s of all 6 mAb are enriched for aromatic (Y/F/W) residues (mean 32.6% of a.a., range 21.4%-47.6%). The HCDR3 of mAb G is unique in the group in that it is encoded by an IGHD2, not IGHD3 gene, and it contains glycine (Gly), a small residue of unique flexibility, at an unusually high frequency (9 of 28, or 32.1% of residues). MAb A HCDR3 contains fewer Gly residues, though still a relatively high percentage (19.2%), and MAb A and mAb G HCDR3s have similar long Gly- and Tyr-enriched 3' sequences generated by N-region nt and extended IGHJ6 genes. Four of 6 mAb (B, C, A, G) contain a rare germline IGHD-gene encoded arginine (R) (Corbett et al., 1997); all 4 of these mAb HCDR3 also contain additional cationic a.a. (R or lysine, K) encoded by N1-region nts.

Light chain sequences were obtained for 4 of the 6 mAb, including 2 kappa and 2 lambda chains, derived from 3 different LC V region genes (Table 1). LCDR3 lengths range from 9 to 12 a.a. Both mAb B and E express a serine-enriched long LCDR3 (12 and 11 a.a., respectively) encoded by the IGLV1–40 gene. Three of the 4 LCDR3 include a central V-gene-encoded highly hydrophobic leucine residue.

3.3. Polyreactivity of human anti-alpha 3(IV)NC1 collagen mAb

Because several of the mAb use Ig gene segments shared with polyreactive and autoreactive CLL B cells, some of which use long HCDR3, we assessed binding of five mAb to several self-Ags associated with polyreactivity (ANA, ssDNA, cardiolipin). Results revealed that none of the mAb bound to nucleus or cytoplasm of HEp-2 cells, as shown in Figure 4 for mAb A, C and D assayed at 1 ug/ml, the concentration that yields strong binding on alpha3(IV)NC1 collagen (Figure 2B), and no mAb bound to cardiolipin (not shown). MAb A displayed modest binding to ssDNA in ELISA (Figure 4). Because approximately 25% of patients with anti-GBM GN co-express antineutrophil cytoplasmic autoantibodies (ANCA), particularly ANCA reactive with myeloperoxidase (MPO) (Rutgers et al., 2005), we also assessed anti-alpha3(IV)NC1 mAb binding to human MPO; results show that no mAb bound to MPO in ELISA (figure 4).

4. Discussion

We recovered 6 non-clonally related human mAb reactive with alpha3(IV)NC1 collagen by fusion of viral-transformed human B cells from immunized Hu-HSC mice. Functional assessment using immunohistochemistry indicates that a subset of the mAb produce linear GBM staining when incubated on normal human kidney sections, indicating specificity for human GBM epitopes. By determining the nucleotide and deduced amino acid sequences of the antibody heavy and light chains, we gained insight into their genetic origins and structural requirements for antigen binding. Whereas the Ab use several different V and D gene alleles, suggesting several solutions for Ag binding, the presence of unmutated genes and preference for IGHV3 subgroup alleles and IGHD3 and IGHJ6 genes suggests that these elements create a particularly suitable scaffold within the human antigen-naïve (preimmune) repertoire for binding NC1 collagen epitopes. This includes use by all mAb of IGHD genes and/or extended IGHJ genes highly enriched for aromatic residues, and use by 4 mAb (A, B, C, and G) of rare IGHD genes (and reading frames) for which the germline sequence encodes cationic arginine, a residue selected against in the human naïve repertoire (Larimore et al., 2012). The most striking and unusual feature shared by all mAb is a long, to exceptionally long, HCDR3, an uncommon structural motif that is associated with Ab capable of accessing recessed epitopes that are generally “hidden” from conventional Abs.

The HCDR3 is of particular interest because it is the dominant site of direct Ag interaction and is typically the major determinant of antigen specificity (Kabat and Wu, 1991; Padlan, 1994; Wilson and Stanfield, 1994). In contrast to VH-encoded HCDR1 and HCDR2, which are constrained by canonical length and structure requirements, HCDR3 is generated during Ig gene rearrangement by unique mechanisms that permit enormous sequence diversity. In humans, the IMGT-designated HCDR3 is comprised of a.a. between the conserved FR3 cysteine (Cys, a.a. 104) and FR4 tryptophan (a.a. 118), and has an average length of 15.2 a.a., based on the extensive database of Zemlin and colleagues (Zemlin et al., 2003). (Using the alternative Kabat numbering and HCDR3 designation, human HCDR3 average length is approximately 13.2 a.a.). In comparison, the HCDR3 in our sequenced mAb are extremely long, with mean length 25.5 a.a. (range 20–36 a.a., Table 1, IMGT designation). These findings confirm and expand our initial report of the first Hu-HSC-derived human anti-

alpha3(IV)NC1 collagen mAb, termed 2D6, which had a longer than average HCDR3, as well as unusual hydrophobic and dual cysteine motifs (Worni-Schudel et al., 2015).

The finding of long HCDR3 in all of the anti-alpha3(IV)NC1 collagen mAb is striking in part because HCDR3 longer than 20 a.a. are relatively rare in the normal healthy human adult (Gokmen et al., 1998). Briney and colleagues found that HCDR3 24 a.a. and 28 a.a. (>2 and >3 SD above mean HCDR3 length, respectively, in their studies) constituted only 3.5% and 0.43% of the naïve human B cell repertoire, and were even less frequent in memory B cell populations (Briney et al., 2012). Whereas some constraints on HCDR3 length are evident in early progenitor cells (before antigen selection) (Ivanov et al., 2005), considerable evidence indicates that negative selection plays a major role in removing B cells bearing Ig receptors with long HCDR3. During B cell development in human bone marrow, average HCDR3 length (by Kabat designation) drops significantly from 15.6 and 16.1 a.a. in pre-B and early immature B cells, respectively, to 13.5 a.a. in mature naïve cells (Meffre et al., 2001; Wardemann et al., 2003). HCDR3 are significantly shorter in productive (expressed) versus non-productive (out-of-frame) rearranged human IgH alleles (Larimore et al., 2012; Souto-Carneiro et al., 2004), consistent with negative selection. It is estimated that 69% or more of early B cells are subsequently removed on the basis of surface expression of receptors bearing a long HCDR3 (Larimore et al., 2012).

Selective bias against lengthy HCDR3 is attributed in large part to their propensity for autoreactivity and polyreactivity (Aguilera et al., 2001; Crouzier et al., 1995; Ichiyoshi and Casali, 1994). Wardemann et al found that the significant drop in HCDR3 length observed during human B cell maturation is paralleled by a dramatic decrease in frequency of self-reactivity, from 75.9% of Ab produced by early immature cells to 20.4% of those from mature naïve B cells (Wardemann et al., 2003). Nonetheless, some B cells with long HCDR3 and autoreactive receptors do reach the periphery: long HCDR3 (> 20 a.a., Kabat numbering) or autoreactivity were identified in approximately 5% and 20%, respectively, of human mature naïve B cells (Wardemann et al., 2003). Understanding the advantage of maintaining these B cells in the repertoire, and identifying their self and foreign ligand targets, will provide new insight into immune functions and disease origins, as well as suggest novel therapeutic targets. In this respect, it is intriguing that several of the anti-alpha3(IV)NC1 collagen mAb express variable region genes enriched among CLL B cells that share features with the natural or innate polyreactive B cell repertoire (Hwang et al., 2014). In our survey of mAb binding to selected self-Ags that have been associated with polyreactivity, we detected modest binding to ssDNA by mAb A, whereas no mAb bound to HEp-2 cells or cardiolipin. The mAb also did not bind MPO, a self-Ag target in a quarter of patients with anti-GBM GN (Rutgers et al., 2005).

Long HCDR3 can form tertiary structures particularly suitable for binding partially hidden conformational epitopes in complex protein antigens. Crystal structure analyses reveal protruding loops or convex Ag binding surfaces on long HCDR3 that penetrate deep cavities on viruses or other large protein targets. Long HCDR3 are prominent in neutralizing Ab responses to malaria and trypanosomes and among potent anti-HIV IgG bnAb (Cardoso et al., 2005; Zhou et al., 2015; Zwick et al., 2004). Long HCDR3 can extend between HIV surface glycans (notorious for rapid antigenic shift to evade immune detection) to contact

deep, conserved epitopes, such as the recessed CD4 “supersite” (Zhou et al., 2015) or the membrane-proximal Trp-rich region of HIV gp41 (Zwick et al., 2004). It is noteworthy that these same features also promote Ag crossreactivity and autoreactivity, prominent features of multiple anti-HIV bnAb with long HCDR3 (Haynes et al., 2005).

Long HCDR3 are also common among the unique small heavy-chain-only antibodies found in camelids (Harmsen et al., 2000; Lawson, 2012; Muyldermans et al., 2001). Camelid immunization has been exploited to generate domains that bind preferentially within enzyme active sites (De Genst et al., 2006; Li et al., 2012), a specificity rarely found among conventional Ig. This has enabled development of potent therapeutic enzyme inhibitors that neutralize bacterial toxins and snake venoms (Wesolowski et al., 2009). This suggests that Ig with long HCDR3 a critical void in the B cell and antibody repertoire, and may explain their persistence in the preimmune repertoire despite autoreactivity.

It is of interest that camelid VHH express clan III VH genes with a high degree of homology to the human IGHV3 subgroup (Harmsen et al., 2000), a subgroup used by half of our anti- α 3(IV)NC1 collagen mAb (Table 1). This suggests particular structural suitability of this pairing (IGHV3 and long HCDR3) in supporting antibody interactions with buried epitopes. This may have particular relevance to anti-GBM nephritis, in which the pathogenic epitopes are nonlinear and include key hydrophobic residues hidden within the NC1 hexamer (Borza and Hudson, 2003). Long HCDR3 enriched in Gly residues, as seen with mAb A and G, may provide additional conformational flexibility for binding such noncontiguous epitopes. Although α 345NC1 hexamer dissociation is required for binding by patients' IgG, it has also been demonstrated that patients' IgG can facilitate hexamer dissociation and epitope exposure, in part by displacing the α 5(IV)NC1 subunit to unmask cryptic epitopes and hydrophobic side chains on α 3(IV)NC1 (Borza et al., 2005; Pedchenko et al., 2010). It is thus tempting to speculate that long HCDR3 are particularly suited for this purpose, able to extend into the cleft between NC1 monomers to disrupt the partially hydrophobic interface.

Importantly, our recovery of long HCDR3 does not simply reflect alterations in B cell development and selection within the Hu-HSC model. Ippolito and colleagues used high-throughput sequencing to comprehensively characterize the humanized B cell repertoire in unimmunized NSG Hu-HSC mice and found that the engrafted immune system was globally diverse and almost indistinguishable from that of the healthy human adult (Ippolito et al., 2012). The HCDR3s in particular were similar in length (mean 15.6 a.a.), junctional diversity mechanisms, average neutral hydrophathy, and charged a.a. frequency, and reminiscent of independent results obtained using single cell PCR (Kolar et al., 2004). The results mirror those in adult human bone marrow or stem cell transplant recipients, in which engrafted repertoires resemble those of healthy adults (Gokmen et al., 1998). Moreover, B cell selection and tolerance mechanisms appear intact in Hu-HSC mice, judged by patterns of preferential Ig gene use and selection against “inherently autoreactive” V-region genes (Ippolito et al., 2012). Selection bias from other sources, such as expansion of rare retained neonatal B cells, also seems unlikely; the mean HCDR3 length of the human cord blood IgM repertoire (14.03 ± 0.02 a.a.) is similar to that of the adult human (Prabakaran et al., 2012).

The anti- $\alpha 3(\text{IV})\text{NC1}$ collagen mAb HCDR3 contain multiple features (long HCDR3, as well as highly cationic and/or highly hydrophobic a.a residues) that are associated with autoreactivity and normally selected against in the naïve B cell repertoire. Evidence from autoantibody transgenic models suggests that B cells expressing these receptors are indeed regulated in vivo. Mice expressing the human polyreactive anti-HIV bnAb 2F5 or 4E10 antibodies as targeted knockins demonstrate profound B cell regulation by deletion, receptor editing, and anergy (Chen et al., 2013; Verkoczy et al., 2011; Verkoczy et al., 2010). These regulatory mechanisms mirror the phenotype we previously reported in mice transgenic for the murine anti- $\alpha 3(\text{IV})\text{NC1}$ collagen mAb 1G6 (Clark et al., 2011; Zhang et al., 2008).

The relationship of the anti- $\alpha 3(\text{IV})\text{NC1}$ collagen IgM mAb and B cells to the differentiated and isotype-switched B cells producing IgG that cause disease in patients will ultimately require sequence analysis of patients' pathogenic IgG. To date this has not been possible using mAb technology and sequence analysis of hybridoma genomic DNA or cDNA, in part because, to our knowledge, no one has recovered a pathogenic IgG mAb from a patient with anti-GBM GN or Goodpasture disease. However, the possibility of direct sequence analysis may be within reach by combining powerful proteomics and next generation sequencing technologies and generating a patient donor-specific Ig DNA reference library, as described for Ag-reactive IgG in vaccinated volunteers (Lavinder et al., 2014). A direct precursor relationship between B cells producing the IgM mAb and patient IgG is suggested by competitive inhibition studies (Fig. 2D). This relationship can also be explored in a modified Hu-HSC model that expresses a relevant HLA class II molecule. Recovery of human IgG is poor in existing Hu-HSC models, in large part because human CD4⁺ T cells select on mouse MHC in the Hu-HSC mouse thymus, such that the mature human CD4⁺ T cells are ill equipped to engage and provide help to the HLA-expressing human B cells. This problem could be circumvented and IgG Ag-specific responses induced by transgenic expression of HLA class II molecules (and deletion of mouse Class II) and engraftment with HLA-matched HSC in novel Hu-HSC models. In this regard, lack of IgG switch is one reason that we do not observe significant kidney disease in our immunized Hu-HSC mice. In this regard, it is notable that whereas these rodent models are excellent for dissecting origins of human pathogenic autoimmunity, it is unlikely that the Hu-HSC will develop full-blown disease under most circumstances due to defects in downstream effector mechanisms. NSG mice lack C5 complement and multiple cytokines. Moreover, the target antigen, $\alpha 3(\text{IV})\text{NC1}$ collagen, is less accessible in rodent GBM compared to that in humans and other primates; rodent collagen IV NC1 hexamers are extensively crosslinked, effectively blocking in vivo binding by autoantibodies, including patients' anti-GBM IgG (Luo et al., 2010).

Of the 7 human anti- $\alpha 3(\text{IV})\text{NC1}$ collagen mAb recovered from immunized Hu-HSC mice (this report and Worni-Schudel et al., 2015), five were derived from the same HSC donor. It is thus notable that HCDR3 regions are not only highly variable between different Ig, but they can be highly unique to an individual. This raises the possibility that one factor that predisposes to anti-GBM GN is the capacity to generate a suitable Ig repertoire that includes long HCDR3. The relative rarity of anti-GBM GN suggests that this predisposition is uncommon. Our finding of long HCDR3 in unmutated human IgM further suggests that the predisposition is genetic, although it is also plausible that in patients there is a

contribution from the lifetime exposure to exogenous Ags that shapes the peripheral B cell repertoire.

Conclusions

Our findings suggest that genetic susceptibility that favors development of long HCDR3 and environmental exposure(s) that cause transient loss of immune tolerance may be critical for an immune response to successfully recruit anti- $\alpha 3(\text{IV})\text{NC1}$ collagen B cells. It has been hypothesized that transient disruption of central tolerance is necessary to release normally “forbidden” polyreactive B cell precursors bearing critical long HCDR3 into the periphery for development of anti-HIV bnAb in HIV infection or vaccination (Haynes, 2015). We propose that a similar phenomenon may be necessary in development of a pathogenic anti-GBM response, and may explain the rarity of anti-GBM Ig and associated diseases like GPS and anti-GBM nephritis. Ultimately, comparison of patients’ pathogenic IgG with IgM precursors will provide critical insight into the origins and regulation of the pathogenic anti- $\alpha 3(\text{IV})\text{NC1}$ collagen B cells, including the roles of Ig mutation, antigen crossreactivity, tolerance, and identity of factors that initiate activation.

Acknowledgments

This work was funded by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award number R01DK088904 (MHF), a Duke University Undergraduate Research grant (ESB), and a grant from the Institute of Medical Research, Durham VAMC (MHF). We thank the following Shared Resources of the Duke Cancer Institute: Cancer Center Isolation Facility Breeding Core, Flow Cytometry Shared Resource, DNA Analysis Facility, and Biospecimen Repository and Processing Core. We thank Clay Rouse, DVM for help with animal procedures, J. Michael Cook, PhD for help with flow cytometry, Julie Fuller of the Surgical Sciences Histopathology Core/Shared Resource Lab for help with histopathology, and Alexis Sharp, PA, Shannon McCall, MD, and David Howell, MD, PhD, for access to normal human kidney specimens.

Abbreviations

Ag	antigen
bnAb	broadly neutralizing anti-HIV Ig
CLL	chronic lymphocytic leukemia
EBV	Epstein Barr virus
GBM	glomerular basement membrane
GP	Goodpasture
GPS	GP syndrome
HC	heavy chain
HCDR3	heavy chain complementarity determining region3
HSC	hematopoietic stem cells
IMGT	ImMunoGeneTics Information System
LC	light chain

mAb	monoclonal antibody
NC1	non-collagenous-1 (domain)
NSG	NOD-scid-gamma
V	variable

REFERENCES

- Aguilera I, Melero J, Nunez-Roldan A, Sanchez B. Molecular structure of eight human autoreactive monoclonal antibodies. *Immunology*. 2001; 102:273–280. [PubMed: 11298825]
- Ali R, Dersimonian H, Stollar BD. Binding of monoclonal anti-native DNA autoantibodies to DNA of varying size and conformation. *Mol Immunol*. 1985; 22:1415–1422. [PubMed: 3879530]
- Bonsignori M, Hwang KK, Chen X, Tsao CY, Morris L, Gray E, Marshall DJ, Crump JA, Kapiga SH, Sam NE, Sinangil F, Pancera M, Yongping Y, Zhang B, Zhu J, Kwong PD, O'Dell S, Mascola JR, Wu L, Nabel GJ, Phogat S, Seaman MS, Whitesides JF, Moody MA, Kelsoe G, Yang X, Sodroski J, Shaw GM, Montefiori DC, Kepler TB, Tomaras GD, Alam SM, Liao HX, Haynes BF. Analysis of a clonal lineage of HIV-1 envelope v2/v3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors. *J Virol*. 2011; 85:9998–10009. [PubMed: 21795340]
- Borza DB, Hudson BG. Molecular characterization of the target antigens of anti-glomerular basement membrane antibody disease. *Springer Semin Immunopathol*. 2003; 24:345–361. [PubMed: 12778332]
- Borza DB, Bondar O, Colon S, Todd P, Sado Y, Neilson EG. Goodpasture autoantibodies unmask cryptic epitopes by selectively dissociating autoantigen complexes lacking structural reinforcement: Novel mechanisms for immune privilege and autoimmune pathogenesis. *J Biol Chem*. 2005; 280:27147–54. [PubMed: 15917228]
- Briney BS, Willis JR, Crowe JE Jr. Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes. *PloS One*. 2012; 7:e36750. [PubMed: 22590602]
- Cardoso RM, Zwick MB, Stanfield RL, Kunert R, Binley JM, Katinger H, Burton DR, Wilson IA. Broadly neutralizing anti-HIV antibody 4e10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity*. 2005; 22:163–173. [PubMed: 15723805]
- Chen Y, Zhang J, Hwang KK, Bouton-Verville H, Xia SM, Newman A, Ouyang YB, Haynes BF, Verkoczy L. Common tolerance mechanisms, but distinct cross-reactivities associated with gp41 and lipids, limit production of HIV-1 broad neutralizing antibodies 2f5 and 4e10. *J Immunol*. 2013; 191:1260–1275. [PubMed: 23825311]
- Clark AG, Mackin KM, Foster MH. Genetic elimination of alpha3(iv) collagen fails to rescue anti-collagen B cells. *Immunology Letters*. 2011; 141:134–139. [PubMed: 21963654]
- Corbett SJ, Tomlinson IM, Sonnhammer EL, Buck D, Winter G. Sequence of the human immunoglobulin diversity (D) segment locus: A systematic analysis provides no evidence for the use of DIR segments, inverted D segments, "minor" D segments or D-D recombination. *J Mol Biol*. 1997; 270:587–597. [PubMed: 9245589]
- Crouzier R, Martin T, Pasquali JL. Heavy chain variable region, light chain variable region, and heavy chain cdr3 influences on the mono- and polyreactivity and on the affinity of human monoclonal rheumatoid factors. *J Immunol*. 1995; 154:4526–4535. [PubMed: 7722307]
- Dal-Bo M, Del Giudice I, Bomben R, Capello D, Bertoni F, Forconi F, Laurenti L, Rossi D, Zucchetto A, Pozzato G, Marasca R, Efremov DG, Guarini A, Del Poeta G, Foa R, Gaidano G, Gattei V. B-cell receptor, clinical course and prognosis in chronic lymphocytic leukaemia: The growing saga of the IGHV3 subgroup gene usage. *Br J Haematol*. 2011; 153:3–14. [PubMed: 21303354]
- De Genst E, Silence K, Decanniere K, Conrath K, Loris R, Kinne J, Muyldermans S, Wyns L. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc Natl Acad Sci U S A*. 2006; 103:4586–4591. [PubMed: 16537393]

- Forconi F, Potter KN, Wheatley I, Darzentas N, Sozzi E, Stamatopoulos K, Mockridge CI, Packham G, Stevenson FK. The normal IGHV1-69-derived B-cell repertoire contains stereotypic patterns characteristic of unmutated CLL. *Blood*. 2010; 115:71–77. [PubMed: 19887677]
- Giudicelli V, Chaume D, Lefranc MP. IMGT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis. *Nucleic Acids Res*. 2004; 32:W435–W440. [PubMed: 15215425]
- Gokmen E, Raaphorst FM, Boldt DH, Teale JM. Ig heavy chain third complementarity determining regions (H CDR3s) after stem cell transplantation do not resemble the developing human fetal H CDR3s in size distribution and Ig gene utilization. *Blood*. 1998; 92:2802–2814. [PubMed: 9763565]
- Harmsen MM, Ruuls RC, Nijman IJ, Niewold TA, Frenken LG, de Geus B. Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Mol Immunol*. 2000; 37:579–590. [PubMed: 11163394]
- Haynes BF. New approaches to HIV vaccine development. *Curr Opin Immunol*. 2015; 35:39–47. [PubMed: 26056742]
- Haynes BF, Fleming J, St Clair EW, Katinger H, Stiegler G, Kunert R, Robinson J, Searce RM, Plonk K, Staats HF, Ortel TL, Liao HX, Alam SM. Cardiophilin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science*. 2005; 308:1906–1908. [PubMed: 15860590]
- Hellmark T, Segelmark M. Diagnosis and classification of Goodpasture's disease (anti-GBM). *J Autoimmunity*. 2014; 48–49:108–112.
- Huang L, Lange MD, Zhang Z. VH replacement footprint analyzer-I, a java-based computer program for analyses of immunoglobulin heavy chain genes and potential VH replacement products in human and mouse. *Front Immunol*. 2014; 5:40. [PubMed: 24575092]
- Hwang KK, Trama AM, Kozink DM, Chen X, Wiehe K, Cooper AJ, Xia SM, Wang M, Marshall DJ, Whitesides J, Alam M, Tomaras GD, Allen SL, Rai KR, McKeating J, Cattera R, Yan XJ, Chu CC, Kelsoe G, Liao HX, Chiorazzi N, Haynes BF. IGHV1-69 B cell chronic lymphocytic leukemia antibodies cross-react with HIV-1 and hepatitis C virus antigens as well as intestinal commensal bacteria. *PloS one*. 2014; 9:e90725. [PubMed: 24614505]
- Ichiyoshi Y, Casali P. Analysis of the structural correlates for antibody polyreactivity by multiple reassortments of chimeric human immunoglobulin heavy and light chain V segment. *J Exp Med*. 1994; 180:885–895. [PubMed: 8064239]
- Ippolito GC, Hoi KH, Reddy ST, Carroll SM, Ge X, Rogosch T, Zemlin M, Shultz LD, Ellington AD, Vandenberg CL, Georgiou G. Antibody repertoires in humanized NOD-scid-IL2Rgamma(null) mice and human B cells reveals human-like diversification and tolerance checkpoints in the mouse. *PLoS One*. 2012; 7:e35497. [PubMed: 22558161]
- Ivanov II, Schelonka RL, Zhuang Y, Gartland GL, Zemlin M, Schroeder HW Jr. Development of the expressed Ig CDR-H3 repertoire is marked by focusing of constraints in length, amino acid use, and charge that are first established in early B cell progenitors. *J Immunol*. 2005; 174:7773–7780. [PubMed: 15944280]
- Johnson TA, Rassenti LZ, Kipps TJ. Ig VH1 genes expressed in B cell chronic lymphocytic leukemia exhibit distinctive molecular features. *J Immunol*. 1997; 158:235–246. [PubMed: 8977195]
- Kabat EA, Wu TT. Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VI genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *J Immunol*. 1991; 147:1709–1719. [PubMed: 1908882]
- Kolar GR, Yokota T, Rossi MI, Nath SK, Capra JD. Human fetal, cord blood, and adult lymphocyte progenitors have similar potential for generating B cells with a diverse immunoglobulin repertoire. *Blood*. 2004; 104:2981–2987. [PubMed: 15256421]
- Larimore K, McCormick MW, Robins HS, Greenberg PD. Shaping of human germline IGH repertoires revealed by deep sequencing. *J Immunol*. 2012; 189:3221–3230. [PubMed: 22865917]
- Lavinder JJ, Wine Y, Giesecke C, Ippolito GC, Horton AP, Lungu OI, Hoi KH, DeKosky BJ, Murrin EM, Wirth MM, Ellington AD, Dorner T, Marcotte EM, Boutz DR, Georgiou G. Identification and characterization of the constituent human serum antibodies elicited by vaccination. *Proc Natl Acad Sci U S A*. 2014; 111:2259–2264. [PubMed: 24469811]

- Lawson AD. Antibody-enabled small-molecule drug discovery. *Nat Rev Drug Discov.* 2012; 11:519–525. [PubMed: 22743979]
- Lerner RA, Glasscock RJ, Dixon FJ. The role of anti-glomerular basement membrane antibody in the pathogenesis of human glomerulonephritis. *J Exp Med.* 1967; 126:989–1004. [PubMed: 4964566]
- Li JW, Xia L, Su Y, Liu H, Xia X, Lu Q, Yang C, Reheman K. Molecular imprint of enzyme active site by camel nanobodies: Rapid and efficient approach to produce abzymes with alliinase activity. *J Biol Chem.* 2012; 287:13713–13721. [PubMed: 22374998]
- Liao HX, Levesque MC, Nagel A, Dixon A, Zhang R, Walter E, Parks R, Whitesides J, Marshall DJ, Hwang KK, Yang Y, Chen X, Gao F, Munshaw S, Kepler TB, Denny T, Moody MA, Haynes BF. High-throughput isolation of immunoglobulin genes from single human B cells and expression as monoclonal antibodies. *J Virol Methods.* 2009; 158:171–179. [PubMed: 19428587]
- Luo W, Wang XP, Kashtan CE, Borza DB. Alport alloantibodies but not Goodpasture autoantibodies induce murine glomerulonephritis: Protection by quinary crosslinks locking cryptic alpha3(IV) collagen autoepitopes in vivo. *J Immunol.* 2010; 185:3520–3528. [PubMed: 20709951]
- Meffre E, Milili M, Blanco-Betancourt C, Antunes H, Nussenzweig MC, Schiff C. Immunoglobulin heavy chain expression shapes the B cell receptor repertoire in human B cell development. *J Clin Invest.* 2001; 108:879–886. [PubMed: 11560957]
- Muyldermans S, Cambillau C, Wyns L. Recognition of antigens by single-domain antibody fragments: The superfluous luxury of paired domains. *Trends Biochem Sci.* 2001; 26:230–235. [PubMed: 11295555]
- Padlan E. Anatomy of the antibody molecule [review]. *Mol Immunol.* 1994; 31:169–217. [PubMed: 8114766]
- Pagnoux C, Hogan SL, Chin H, Jennette JC, Falk RJ, Guillevin L, Nachman PH. Predictors of treatment resistance and relapse in antineutrophil cytoplasmic antibody-associated small-vessel vasculitis: Comparison of two independent cohorts. *Arthritis Rheum.* 2008; 58:2908–2918. [PubMed: 18759282]
- Pearson T, Greiner DL, Shultz LD. Creation of "humanized" mice to study human immunity, *Curr Protoc Immunol.* Chapter. 2008; 15 Unit 15 21.
- Pedchenko V, Bondar O, Fogo AB, Vanacore R, Vozizyan P, Kitching AR, Wieslander J, Kashtan C, Borza DB, Neilson EG, Wilson CB, Hudson BG. Molecular architecture of the Goodpasture autoantigen in anti-GBM nephritis. *N Engl J Med.* 2010; 363:343–354. [PubMed: 20660402]
- Posner MR, Elboim H, Santos D. The construction and use of a human-mouse myeloma analogue suitable for the routine production of hybridomas secreting human monoclonal antibodies. *Hybridoma.* 1987; 6:611–625. [PubMed: 3436625]
- Prabakaran P, Chen W, Singarayan MG, Stewart CC, Streaker E, Feng Y, Dimitrov DS. Expressed antibody repertoires in human cord blood cells: 454 sequencing and IMGT/HIGHV-QUEST analysis of germline gene usage, junctional diversity, and somatic mutations. *Immunogenetics.* 2012; 64:337–350. [PubMed: 22200891]
- Reddy ST, Ge X, Miklos AE, Hughes RA, Kang SH, Hoi KH, Chrysostomou C, Hunicke-Smith SP, Iverson BL, Tucker PW, Ellington AD, Georgiou G. Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nature Biotech.* 2010; 28:965–969.
- Rutgers A, Slot M, van Paassen P, van Breda Vriesman P, Heeringa P, Tervaert JW. Coexistence of anti-glomerular basement membrane antibodies and myeloperoxidase-ANCAs in crescentic glomerulonephritis. *Am J Kidney Dis.* 2005; 46:253–62. [PubMed: 16112043]
- Shultz LD, Lyons BL, Burzenski LM, Gott B, Chen X, Chaleff S, Kotb M, Gillies SD, King M, Mangada J, Greiner DL, Handgretinger R. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol.* 2005; 174:6477–6489. [PubMed: 15879151]
- Souto-Carneiro MM, Longo NS, Russ DE, Sun HW, Lipsky PE. Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. *J Immunol.* 2004; 172:6790–6802. [PubMed: 15153497]

- Stevens RH, Macy E, Morrow C, Saxon A. Characterization of a circulating subpopulation of spontaneous antitetanus toxoid antibody producing B cells following in vivo booster immunization. *J Immunol.* 1979; 122:2498–2504. [PubMed: 312871]
- Verkoczy L, Chen Y, Bouton-Verville H, Zhang J, Diaz M, Hutchinson J, Ouyang YB, Alam SM, Holl TM, Hwang KK, Kelsoe G, Haynes BF. Rescue of HIV-1 broad neutralizing antibody-expressing B cells in 2F5 VH × VL knockin mice reveals multiple tolerance controls. *J Immunol.* 2011; 187:3785–3797. [PubMed: 21908739]
- Verkoczy L, Diaz M, Holl TM, Ouyang YB, Bouton-Verville H, Alam SM, Liao HX, Kelsoe G, Haynes BF. Autoreactivity in an HIV-1 broadly reactive neutralizing antibody variable region heavy chain induces immunologic tolerance. *Proc Natl Acad Sci U S A.* 2010; 107:181–186. [PubMed: 20018688]
- von Budingen HC, Kuo TC, Sirota M, van Belle CJ, Apeltsin L, Glanville J, Cree BA, Gourraud PA, Schwartzburg A, Huerta G, Telman D, Sundar PD, Casey T, Cox DR, Hauser SL. B cell exchange across the blood-brain barrier in multiple sclerosis. *J Clin Invest.* 2012; 122:4533–4543. [PubMed: 23160197]
- Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. Predominant autoantibody production by early human B cell precursors. *Science.* 2003; 301:1374–1377. [PubMed: 12920303]
- Wesolowski J, Alzogaray V, Reyelt J, Unger M, Juarez K, Urrutia M, Cauerhff A, Danquah W, Rissiek B, Scheuplein F, Schwarz N, Adriouch S, Boyer O, Seman M, Licea A, Serreze DV, Goldbaum FA, Haag F, Koch-Nolte F. Single domain antibodies: Promising experimental and therapeutic tools in infection and immunity. *Med Microbiol Immunol.* 2009; 198:157–174. [PubMed: 19529959]
- Wilson C, Dixon F. Anti-glomerular basement membrane antibody-induced glomerulonephritis. *Kidney Int.* 1973; 3:74–89. [PubMed: 4571918]
- Wilson IA, Stanfield RL. Antibody-antigen interactions: New structures and new conformational changes. *Curr Opin Struct Biol.* 1994; 4:857–867. [PubMed: 7536111]
- Worni-Schudel IM, Clark AG, Chien T, Hwang KK, Chen BJ, Foster MH. Recovery of a human natural antibody against the noncollagenous-1 domain of type IV collagen using humanized models. *J Transl Med.* 2015; 13:185–196. [PubMed: 26048777]
- Yousfi Monod M, Giudicelli V, Chaume D, Lefranc MP. IMGT/JunctionAnalysis: The first tool for the analysis of the immunoglobulin and t cell receptor complex V-J and V-D-J junctions. *Bioinformatics.* 2004; 20(Suppl 1):I379–I385. [PubMed: 15262823]
- Yu X, McGraw PA, House FS, Crowe JE Jr. An optimized electrofusion-based protocol for generating virus-specific human monoclonal antibodies. *J Immunol Methods.* 2008; 336:142–151. [PubMed: 18514220]
- Zemlin M, Klinger M, Link J, Zemlin C, Bauer K, Engler JA, Schroeder HW Jr, Kirkham PM. Expressed murine and human CDR-hH3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures. *J Mol Biol.* 2003; 334:733–749. [PubMed: 14636599]
- Zhang Y, Su SC, Hecox DB, Brady GF, Mackin KM, Clark AG, Foster MH. Central tolerance regulates B cells reactive with Goodpasture antigen alpha3(IV)NC1 collagen. *J Immunol.* 2008; 181:6092–6100. [PubMed: 18941198]
- Zhou T, Lynch RM, Chen L, Acharya P, Wu X, Doria-Rose NA, Joyce MG, Lingwood D, Soto C, Bailer RT, Ernandes MJ, Kong R, Longo NS, Louder MK, McKee K, O'Dell S, Schmidt SD, Tran L, Yang Z, Druz A, Luongo TS, Moquin S, Srivatsan S, Yang Y, Zhang B, Zheng A, Pancera M, Kirys T, Georgiev IS, Gindin T, Peng HP, Yang AS, Mullikin JC, Gray MD, Stamatatos L, Burton DR, Koff WC, Cohen MS, Haynes BF, Casazza JP, Connors M, Corti D, Lanzavecchia A, Sattentau QJ, Weiss RA, West AP Jr, Bjorkman PJ, Scheid JF, Nussenzweig MC, Shapiro L, Mascola JR, Kwong PD. Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors. *Cell.* 2015; 161:1280–1292. [PubMed: 26004070]
- Zwick MB, Komori HK, Stanfield RL, Church S, Wang M, Parren PW, Kunert R, Katinger H, Wilson IA, Burton DR. The long third complementarity-determining region of the heavy chain is important in the activity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5. *J Virol.* 2004; 78:3155–3161. [PubMed: 14990736]

Highlights

- We report isolation of fully human monoclonal antibodies (mAb) reactive with alpha3(IV)NC1 collagen and human kidney from immunized humanized Hu-HSC mice.
- The antigen binding sites of human anti-alpha3(IV)NC1 collagen mAb are dominated by exceptionally long heavy chain complementarity determining region3 (HCDR3).
- Long HCDR3 are suited to binding recessed, partially hidden epitopes that are often immunologically silent to conventional antibodies with shorter HCDR3, and are found among polyreactive anti-microbial antibodies including broadly neutralizing self-antigen cross-reactive anti-HIV Ig.
- B cells expressing long HCDR3 are selected against in the healthy immune repertoire, suggesting that transient escape from tolerance may be necessary for expansion of anti-alpha3(IV)NC1 collagen B cells.
- Human anti-alpha3(IV)NC1 collagen monoclonal autoantibodies use unmutated variable region genes that are enriched among chronic lymphocytic leukemia (CLL) antibodies that share features with natural polyreactive Ig.

Figure 1.

Human immune cell engraftment in Hu-HSC mice immunized with alpha3(IV)NC1 collagen. A and B. Flow cytometry analysis shows human leukocytes (hCD45+) and human B (CD19+) and T (CD3+) cells in spleens of Hu-HSC mice. Cells were gated on live cells based on forward and side scatter. Mice were injected with HSC from single (A) or mixed (B) donors. Percentage of cells within each marker is indicated. C. Staining controls include human peripheral blood leukocytes (PBL) (left) and mouse splenocytes (right). Shown are human chimerism (D), human B cell percent (E) and human B cell count (F) in spleen, and

human IgM concentration in serum (G). Each symbol represents an individual Hu-HSC mouse. Chimerism is calculated as described in Methods.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

A. Hu-HSC mouse kidney **C. Binding to human kidney** **B. Alpha3(IV)NC1 collagen binding** **D. Inhibition of Ag binding** **Figure 2.**

Glomerular basement membrane antigen binding by human mAb reactive with anti-alpha3(IV)NC1 collagen. A. Direct immunofluorescence of kidney sections from Hu-HSC mice with and without (control) immunization, visualized with labeled anti-human Ig (40X). B. Shown is antigen binding as measured in ELISA using concentration-matched (1 ug/ml) mAb (OD405 on antigen after subtraction of OD405 on diluent-coated plates, duplicate measurements). C. Indirect immunofluorescence of human kidney sections incubated with mAb A (4 μg/ml) and human IgM control (8 μg/ml), visualized with labeled anti-human IgM secondary antibody. D. Patients' anti-GBM IgG inhibit binding of human mAb to alpha3(IV)NC1 collagen. Results are shown for IgG from two patients with active anti-GBM nephritis. Percent inhibition is expressed relative to the OD405 for mAb alone. Each sample was plated in triplicate or quadruplicate, and the assay was repeated twice.

	V	N or P	D-gene	N or P	J-gene	length (a.a.)
mAb B	A R	E S N G D F T R T	Y Y D Y V W G S Y R Y T D	F G P N	Y Y Y Y G M D V	36
mAb C	A S	Y D R	Y D Y V W G S Y R Y	Q	Y Y F D Y	21
mAb D	A K D	W	Y Y D F W S G Y	P G P D	Y Y G M D V	22
mAb E	A R D	L G	Y Y D F W S G Y	W V S A L	D Y	20
mAb A	A R	G L K	L R F L E W L	P T G G G	Y Y Y Y Y G M D V	26
mAb G	A R D	G G R T L R	G G R	Y G G G G E S	Y Y Y Y Y G M D V	28

Figure 3.

Heavy chain CDR3 amino acid sequences of human anti- α 3(IV)NC1 mAb. Genetic origins of the nucleotides that encode the amino acids are indicated, including the IGHV, IGHD, and IGHJ gene segment and intervening N-region nucleotides and P-nucleotides. Gene origins are delineated according to the ImMunoGeneTics (IMGT)/JunctionAnalysis feature of IMGT/V-QUEST, and compared to the closest corresponding germline V, D, and J region gene segment alleles as identified in the IMGT reference database (Giudicelli et al., 2004; Yousfi Monod et al., 2004). The mAb are ordered to facilitate comparison of IGHD and IGHJ gene segment contributions, with spaces introduced to optimize alignments.

A. Binding to Hep-2 cells (ANA)

B. Binding to ssDNA and MPO

Figure 4.

Polyreactivity of human anti- $\alpha 3(\text{IV})\text{NC1}$ collagen mAb. A. Indirect immunofluorescence using human epithelial type 2 (HEp-2) cells incubated with human mAb A, C, or D, visualized with labeled anti-human Ig, and incubated with positive control human serum showing anti-nuclear antibody (ANA) pattern (lower right panel). B. Shown is antigen binding as measured in ELISA using concentration-matched (1 $\mu\text{g}/\text{ml}$) human mAb, for myeloperoxidase (MPO (left) and ssDNA (right)). The value is mean OD405 on antigen after subtraction of OD405 on diluent-coated wells, based on duplicate samples. The

positive control antibody OD value was 0.5 for binding to MPO and 2.8 for binding to ssDNA.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1
Genetic Origins of Human Monoclonal Antibodies Reactive with Alpha3(IV)NC1 Collagen.

	Heavy chain				Light chain			
	IGH genes/alleles				HCDR3	IGK/IGL genes/alleles		
mAb	V-gene	D-gene	J-gene		length	Isotype	V-gene	J-gene
A	V1-8*01	D3-3*01	I6*02		26	kappa	KV2-28*01	KJ1*01
B	V3-48*04	D3-16*02	I6*02		36	lambda	LV1-40*01	LJ1*01
							LV1-40*02	
C	V3-21*01	D3-16*02	I4*02		21	ND	ND	ND
	V3-21*02							
D	V3-23*01	D3-3*01	I6*02		22	kappa	KV1-8*01	KJ1*01
	V3-23*04							
	V3-23D*01							
E	V1-69*01	D3-3*01	I4*02		20	lambda	LV1-40*01	LJ3*02
	V1-69D*01				LV1-40*02			
G	V4-61*02	D2-15*01	I6*02		28	lambda	ND	ND

Legend:

Sequences were determined as described in Methods, and gene segments assigned and heavy chain CDR3 (HCDR3) length delineated according to the ImMunoGeneTics (IMGT) IMGT/JunctionAnalysis feature of IMGT/V-QUEST.

ND, not determined.