PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Genetic and structural basis of the human anti-α-galactosyl antibody response

David B. Langley[a], Peter Schofield[a], Damien Nevoltris[a], Jennifer Jackson[a], Katherine J. L. Jackson[a], Tim J. Peters[a], Melanie Burk[b], Jacqueline M. Matthews[c], Antony Basten[a,d], Christopher C. Goodnow[a,e,f], Sheryl van Nunen[b,g], Joanne H. Reed[a,d,1,2], and Daniel Christ[a,d,1,2]

Edited by Ian Wilson, The Scripps Research Institute, La Jolla, CA; received January 18, 2022; accepted May 10, 2022

Humans lack the capacity to produce the Galα1–3Galβ1–4GlcNAc (α-gal) glycan, and produce anti-α-gal antibodies upon exposure to the carbohydrate on a diverse set of immunogens, including commensal gut bacteria, malaria parasites, cetuximab, and tick proteins. Here we use X-ray crystallographic analysis of antibodies from α-gal knockout mice and humans in complex with the glycan to reveal a common binding motif, centered on a germline-encoded tryptophan residue at Kabat position 33 (W33) of the complementarity-determining region of the variable heavy chain (CDRH1). Immunoglobulin sequencing of anti-α-gal B cells in healthy humans and tick-induced mammalian meat anaphylaxis patients revealed preferential use of heavy chain germline IGHV3-7, encoding W33, among an otherwise highly polyclonal antibody response. Antigen binding was critically dependent on the presence of the germline-encoded W33 residue for all of the analyzed antibodies; moreover, introduction of the W33 motif into naive IGHV3-23 antibody phage libraries enabled the rapid selection of α-gal binders. Our results outline structural and genetic factors that shape the human anti-α-galactosyl antibody response, and provide a framework for future therapeutics development.

alpha-galactose | antibody | germline restriction | mammalian meat allergy

Significance

The immune response to Galα1–3Galβ1–4GlcNAc (α-gal) has been linked to human allergy and inflammation, and to resistance to xenotransplantation and infection. Here we investigate structural and genetic features that shape the antibody response to α-gal. Using crystallography, single-cell sorting, and sequencing, we observe a conserved sequence motif (W33 motif) in the complementarity-determining region of the antibody heavy chain. The germline-encoded aromatic residue directly contacts the α-gal antigen, analogous to what has been observed in broadly neutralizing antibodies in HIV and other infectious diseases. Our findings outline common binding modes and germline usage underpinning the α-gal response, providing molecular insights and guidance for drug development efforts.

Human serum immunoglobulin (Ig) naturally contains a high proportion (~1%) of antibodies directed against α-gal, a terminal Galα1–3Galβ1–4GlcNAc moiety linked to proteins or ceramide (1, 2). The carbohydrate is found in high titers on the surface of cells of most mammals but is notably absent in old world monkeys, humans, and other great apes (3). This difference is due to the gene responsible for addition of the terminal galactose, α1,3 galactosyltransferase (α1,3GT, or *GGTA1*) having been inactivated in hominids some ~20 million to 30 million years ago (4). This likely occurred as a survival mechanism against the emergence of an α-gal-containing pathogen (5–8). Sustained antibody responses to the antigen are likely driven by continued exposure to commensal gut bacteria that also express α-gal on their exteriors (9).

Anti-α-gal antibodies in human sera were first observed in blood pathologies β-thalassemia (10) and sickle cell anemia (11), where distorted red cells were coated in antibodies associated with senescent red cells which contain reduced amounts of sialic acid at their exteriors, thus revealing cryptic α-gal epitopes (12). It was postulated that these antibodies helped remove old, damaged, or misshapen cells via phagocytic macrophages, or destruction via complement. Inadvertent defucosylation of the penultimate galactose moiety of the B-type blood antigen would similarly expose a terminal α-gal moiety within the H-antigen scaffold (13), perhaps explaining lower baseline anti-α-gal titers in B-positive individuals (14, 15).

In 2007, one of us (S.v.N.) described an unexpected link between the development of mammalian meat allergy (MMA) and prior tick bites (16, 17), later shown by Commins et al. (18) to be due to specific IgE to Galα1-3Gal. More recently, it has become apparent that sensitization to the α-gal epitope is of key importance to cetuximab-induced anaphylaxis (an α-gal-containing monoclonal antibody produced in murine NS0 cells) (19), MMA (18), tick anaphylaxis (TA) (20), and the rejection of xenotransplants (21). Although the mechanism by which ticks might prime a response leading to MMA remains unclear (22), the phenomenon, also known as α-gal syndrome, is believed to relate to the presence of endogenous α-gal in tick saliva (23, 24). It has also recently been demonstrated that tick saliva, in and of itself, enhances the α-gal B cell response by acting as an effective adjuvant (25). Recently, anti-α-gal IgE has been reported as a possible risk factor for coronary artery disease (26). In contrast to the detrimental effects of the response in allergy and inflammation, high serum levels of anti-α-gal IgM have been shown to be protective against malaria, in both observational human studies and animal challenge experiments; indeed, it has been suggested that protection against malaria may

Author contributions: D.B.L., A.B., C.C.G., S.v.N., J.H.R., and D.C. designed research; D.B.L., P.S., D.N., J.J., J.M.M., and J.H.R. performed research; M.B. and S.v.N. contributed new reagents/analytic tools; D.B.L., P.S., D.N., K.J.L.J., T.J.P., J.H.R., and D.C. analyzed data; and D.B.L., P.S., A.B., C.C.G., S.v.N., J.H.R., and D.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]J.H.R. and D.C. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: joanne.reed@sydney.edu.au or d.christ@garvan.org.au.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2123212119/-/DCSupplemental.

Published July 8, 2022.

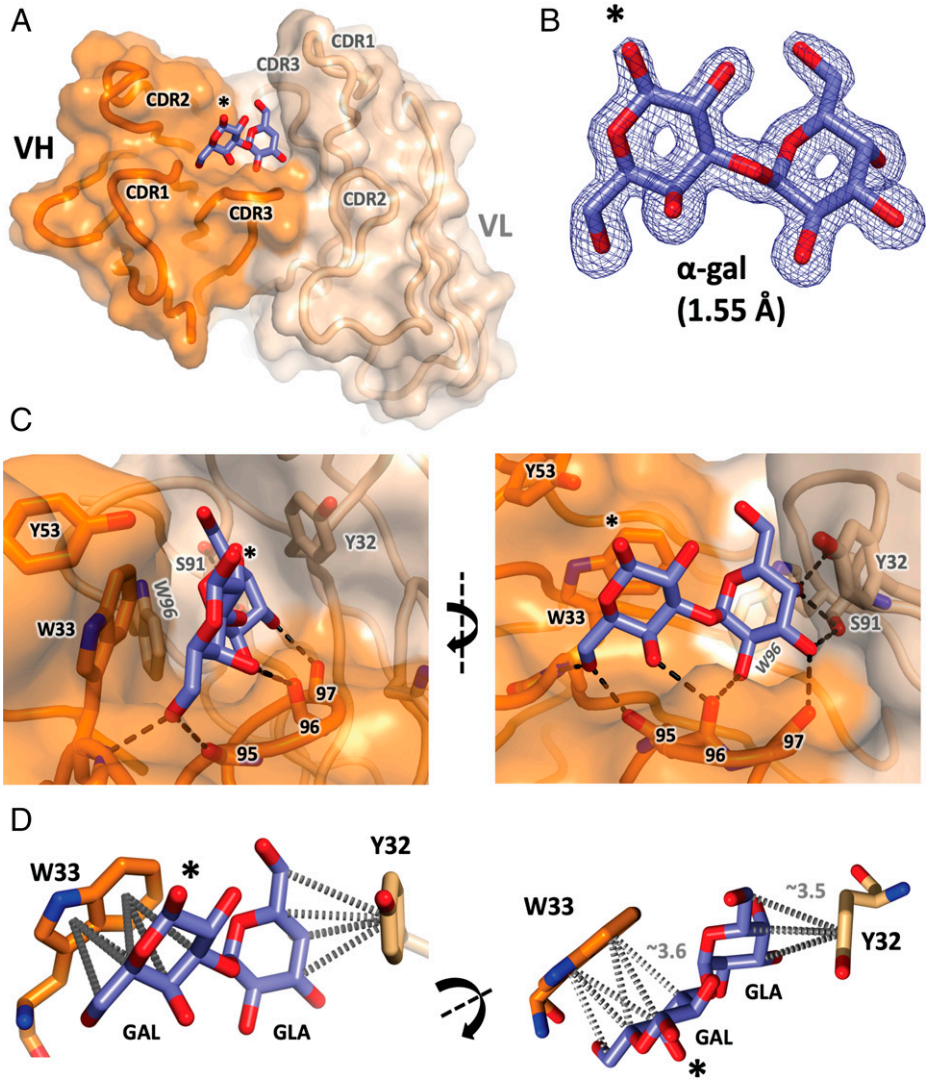

Fig. 1. Crystal structure of M86 Fab in complex with α-gal at 1.55-Å resolution. (*A*) The antigen-binding face of the Fab arm of the M86 antibody (V_H and V_L domains colored orange and tan, respectively) is complexed with the α-gal disaccharide (α1-3 galactobiose, blue and red sticks). The free hemiacetyl hydroxyl, via which attachment to other sugars is facilitated, is indicated by the asterisk. (*B*) Composite omit map (contoured at 1.5 SDs above the mean) for the α-gal moiety allows the disaccharide to be placed unambiguously in the electron density. (*C*) Close-up view of the antibody–ligand interface. Residues of interest are marked with Kabat numbering. Hydrogen bonds directly linking protein to α-gal are shown as dashed lines and are detailed in *SI Appendix*, Table S2. Two perspectives are shown. For detailed stereo views, see *SI Appendix*, Fig. S1. (*D*) Carbon pi bonds between the centroids of the aromatic rings and the carbon moieties within the α-gal sugar are shown by gray dashed lines [image produced using Arpeggio software (71)]. Approximate shortest distances between the planes of the ring and nearest α-gal carbon are shown as gray text. The GAL and GLA components of the α1-3 galactobiose molecule are indicated.

have facilitated the loss of α1,3 galactosyltransferase activity during human evolution (27).

Investigation of α-gal reactivity has relied on monoclonal antibodies derived from mice immunized with rabbit reticulocytes, which are abundantly coated with this glycan (28). Antibodies with superior specificity were identified by immunizing α-gal (α1,3GT) knockout mice (29). When four IgM clones were sequenced, all derived from heavy-chain gene VH22.1 and light-chain gene VK5.1, suggesting that certain antibody gene families dominate recognition. One of these clones, M86, has been widely employed as a prototypical anti-α-gal reagent (21). In humans, anti-α-gal antibodies are found as IgG, IgM, and IgA isotypes (30), and, when nine α-gal-binding clones were sequenced (31), eight had heavy chains derived from the V_H3 gene family, and the other from the V_H1 family, suggesting that the repertoire from which binders are derived might be restricted.

Here we investigate and structurally characterize the basis of antibody heavy-chain gene restriction of the α-gal response. We analyzed structural and genetic mechanisms of α-gal recognition by solving the crystal structure of the mouse M86-α-gal complex, and characterize the affinity and structures of human anti-α-gal antibodies isolated from α-gal-binding B cells of healthy individuals and MMA patients.

Results

Structure of M86 in Complex with α-gal. To investigate the structural basis of α-gal recognition by the M86 monoclonal antibody, the Fab arm was crystallized in the presence of the α-gal terminal disaccharide (α1-3 galactobiose). The structure shows the α-gal moiety sandwiched within the antigen-binding site (Fig. 1; see *SI Appendix*, Table S1 for diffraction data and model refinement statistics). The α-gal binding pocket is lined predominantly by CDR1 and CDR3 of the V_H domain and CDR3 of the V_L domain (see *SI Appendix*, Fig. S1 for details). Collectively, V_H and V_L residues bury ∼300 Å^2 of the α-gal surface (*SI Appendix*, Table S2). Two neighboring Fabs within the crystal make additional minor contributions, burying a further ∼106 Å^2 of α-gal

(*SI Appendix*, Fig. S1 and Table S2). The largest contribution from any single residue is V_H CDR1 tryptophan 33 [W33, Kabat numbering used throughout (32, 33)], which buries 48 Å^2. The centroids of its electronegative aromatic side chain project toward one face (the α-face) of the carbohydrate, forming a carbon pi interface (Fig. 1*D* and *SI Appendix*, Fig. S2). The aromatic side chain of V_L tyrosine 32 (Y32) forms a similar carbon pi interface with the terminus of the ligand (Fig. 1*D*). The main chain of W33 also contributes one of eight Fab ligand hydrogen bonds (*SI Appendix*, Table S2). Additional hydrogen bonds contact water molecules that then contact the ligand (*SI Appendix*, Fig. S1). The bulk of hydrogen bonds directly linking α-gal to the Fab utilize donors/acceptors that are part of the main chain, rather than side-chain residues; in particular, a very short V_H CDR3 loop (residues 95 to 97) forms the base of the binding pocket, contributing hydrogen bonds with four hydroxyl groups along one edge of the disaccharide. The pocket is additionally lined by another tryptophan, V_L W96, burying another 31 Å^2 (Fig. 1*C* and *SI Appendix*, Fig. S1 and Table S2). Overall, binding is achieved by a combination of hydrogen bonds contacting the buried edge of the sugar, and carbon pi interactions between the sides of the sugar and the electronegative pi faces of adjacent hydrophobic residues.

α-gal Reactive B Cells in Controls and MMA Donors. Having identified key features of α-gal recognition in M86, we next sought to identify human anti-α-gal antibodies. Flow cytometry was used to screen human blood for α-gal specific B cells (*SI Appendix*, Fig. S3*A*). CD19$^+$ α-gal$^+$ B cells were identified in 14 healthy donors evaluated (0.53 ± 0.46% of total CD19$^+$ B cells). Higher, but more variable, frequencies were observed in four patients with MMA (2.1 ± 2.2%) (*SI Appendix*, Fig. S3*B*). Immunophenotyping was performed on a subset of seven controls and three MMA patients (*SI Appendix*, Fig. S3*C–E*). In both groups, α-gal$^+$ B cells were present in both naïve and memory compartments. There was a higher frequency of unswitched memory B cells among α-gal$^+$ B cells compared to α-gal$^-$ B cells, while switched memory B cell proportions were comparable.

Germline Usage of α-gal Reactive B Cells. Germline and somatic diversity of human anti-α-gal antibodies were analyzed by single-cell sorting B cells binding fluorescently labeled α-gal and Sanger sequencing paired Ig heavy- and light-chain genes in six controls and four patients with MMA. Heavy-chain isotype usage in α-gal$^+$ B cells from both patients and controls revealed a response dominated by IgM (Fig. 2*A*). Immunophenotyping data confirmed that IgM was derived from both naïve and unswitched memory B cells. The α-gal$^+$ B cell sequences showed increased usage of IgG isotypes and reduced IgA, relative to α-gal$^-$ B cells (Fig. 2*A*). No B cells with IgE isotype were identified in α-gal$^+$ or α-gal$^-$ B cells for patients and controls. Evaluation of Ig heavy-chain variable (IGHV) region usage in α-gal$^+$ B cells from both controls and MMA patients revealed a highly polyclonal response comprising multiple IGHV gene families (Fig. 2*C*; individual profiles in *SI Appendix*, Fig. S4). To identify IGHV regions significantly enriched in α-gal$^+$ B cells, we compared the frequency of each V region (as a percentage of all α-gal$^+$ or α-gal$^-$ B cells sequenced) for each individual (healthy donor and patient). Only one IGHV region was statistically significantly increased in α-gal$^+$ B cells compared to α-gal$^-$ B cells: IGHV3-7 ($P = 0.037$). Notably, IGHV3-7 encodes amino acid W33 in CDRH1, a key contributor in M86 recognition of α-gal (Fig. 1). Moreover, tryptophan usage at position 33 was enriched in α-gal$^+$ B cells in both controls (by 164%) and patients with MMA (by 157%), where it was the third most common residue, while, in α-gal$^-$ B cells, it was the fourth most common (Fig. 2*B*). Additionally, W33 was always germline encoded, not resultant from enrichment by somatic hypermutation. To determine whether differences in somatic hypermutation between α-gal$^+$ and α-gal$^-$ B cells existed, we compared the frequency of IGHV region mutation in these populations. Healthy controls showed a slight increase in somatic mutations across α-gal$^+$ B cells (3.9 ± 5.3%) compared to α-gal$^-$ B cells (2.5 ± 3.9%) ($P < 0.001$); in contrast, patients with MMA showed equivalent frequencies of V-region mutation (4.0 ± 5.0% vs. 3.7 ± 5.2%, respectively; $P = 0.48$).

Complementing W33, the peptide backbone of CDRH3 also interacts prominently with M86 (Fig. 1). Evaluation of CDRH3 across our dataset again revealed a highly polyclonal response, with only one example of clonality (in one MMA patient, two α-gal$^+$ B cells shared IGHV, IGHJ, and CDRH3 sequences). There were no significant differences in CDRH3 length when comparing α-gal$^+$ B cells to α-gal$^-$ B cells in controls (15.6 vs. 16.0 amino acids) or patients with MMA (16.1 vs. 15.8 amino acids).

Since α-gal$^+$ B cells were shown to contain more unswitched memory B cells compared to α-gal$^-$ B cells (*SI Appendix*, Fig. S3*D*), data from four donors and three MMA patients were normalized using indexed flow sorting, enabling equivalent populations of α-gal$^+$ and α-gal$^-$ B cells to be compared. We noted the same bias for 3-7 α-gal$^+$ in switched memory B cells

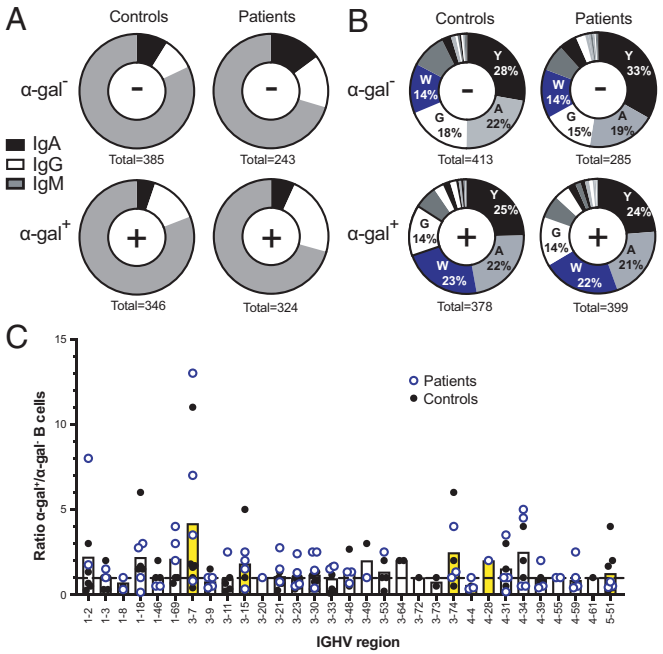

Fig. 2. Ig heavy-chain isotyping and sequencing. (*A*) Ig heavy-chain isotype of single-cell-sorted α-gal$^-$ and α-gal$^+$ CD3$^-$ CD19$^+$ lymphocytes from controls and patients with MMA. The number of cells with sufficient constant region coverage to reveal isotype are shown below each pie graph. (*B*) Frequency of amino acids at position 33 of Ig CDRH1 in α-gal$^-$ or α-gal$^+$ B cells. The four most common amino acids are depicted: tyrosine (Y), alanine (A), tryptophan (W), and glycine (G). Numbers below pie charts indicate the total number of B cells with sufficient CDRH1 coverage to identify amino acid position 33 from each group. (*C*) IGHV usage identified in B cells from MMA patients (open blue circles) and controls (black dots). Each data point represents the ratio of α-gal$^+$/α-gal$^-$ cells attributable to a particular germline for each patient or control. The dashed line represents equivalent numbers of α-gal$^+$ and α-gal$^-$ B cells (ratio = 1). Data points above the line indicate an IGHV bias for α-gal$^+$ B cells. The bars represent the mean α-gal$^+$/α-gal$^-$ ratios for patients and controls combined. IGHVs containing W33 in CDRH1 are shaded yellow. Only germlines for which both α-gal$^+$ and α-gal$^-$ B cells were detected are displayed. Patient-specific germline counts from which these data were derived are shown in *SI Appendix*, Fig. S4.

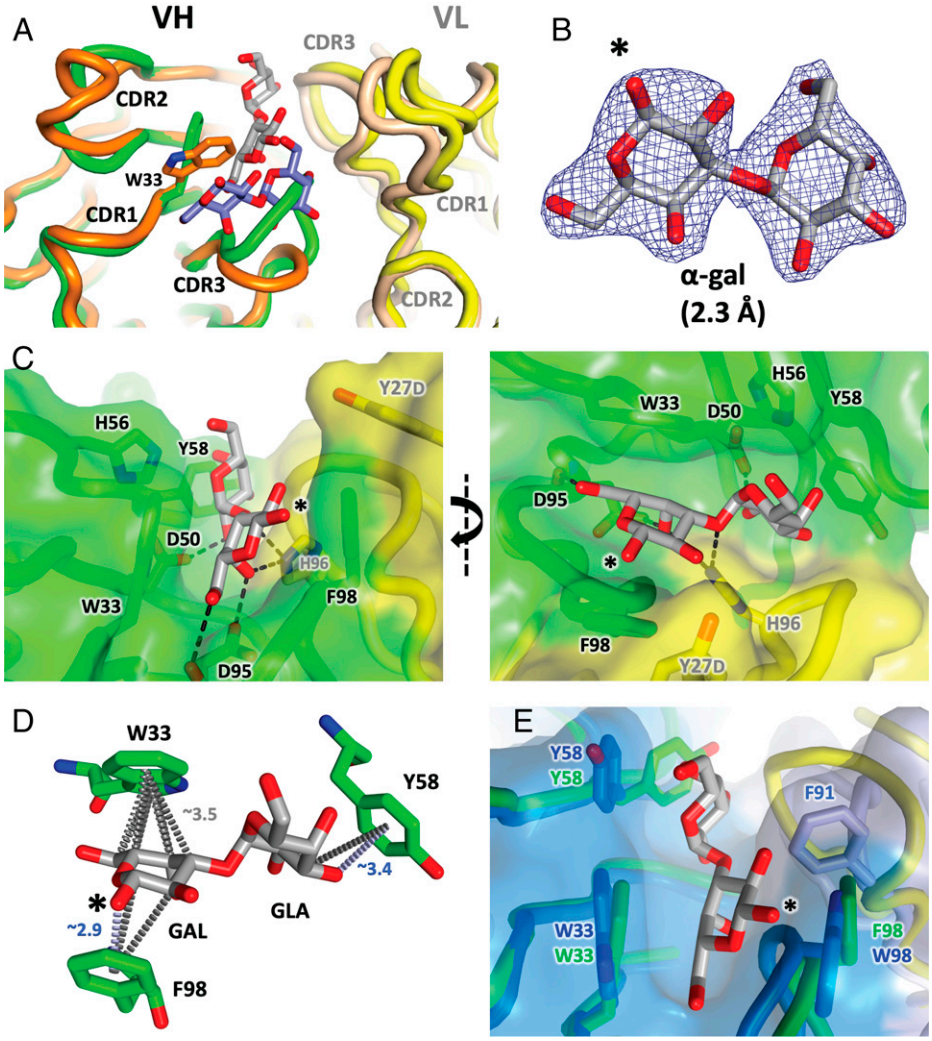

Fig. 3. HKB7 and JEC1 Fab structures. (*A*) HKB7 structure (V_H and V_L chains colored green and yellow, respectively, and α-gal colored gray) superposed with the M86 structure (orange/tan V_H/V_L, and blue α-gal). A superposition about the α-gal moiety is shown in *SI Appendix*, Fig. S7. (*B*) Composite omit electron density (contoured at 1.0 SD above the mean) for the α-gal moiety. The free hemiacetyl hydroxyl, via which attachment to other sugars is facilitated, is indicated by the asterisk. (*C*) Antibody–ligand interface. Hydrogen bonds directly linking protein to α-gal are shown as dashed lines (detailed in *SI Appendix*, Table S5). Two perspectives are shown. For detailed stereo views, see *SI Appendix*, Fig. S8. (*D*) In this structure, all three significant pi interfaces involve heavy-chain residues (W33, Y58, and F98). Carbon pi bonds between the centroids of the aromatic rings and the carbon moieties within the α-gal sugar are shown by dashed gray lines. Oxygen donor pi bonds are shown by blue dashed lines. Approximate nearest-atom distances between α-gal and the aromatic centroids are indicated. The GAL and GLA components of the α1-3 galactobiose molecule are indicated. (*E*) HKB7 structure (green/yellow cartoon and sticks) superposed with JEC1 Fab structure (dark/light blue, sticks and surface). The α-gal molecule from the HKB7 structure (gray and red sticks) sits in the JEC1 cleft, defined predominantly by V_H residues 33, 58, and 98.

in both cohorts (*SI Appendix*, Fig. S5, and individual profiles in *SI Appendix*, Fig. S6). Moreover, 3-7 usage increased as cells differentiated from naïve to memory cells, for both healthy controls (11% naïve, 14% unswitched, and 25% switched) and MMA patients (2% naïve, 13% unswitched, and 20% switched) (*SI Appendix*, Fig. S5), while, in α-gal$^-$ B cells, there was little change (controls: 6% naïve, 1% unswitched, and 7% switched; MMA patients: 4% naïve, 3% unswitched, and 5% switched).

Structures of Human anti-α-gal Antibodies. To determine whether human antibodies derived from α-gal$^+$ B cells with 3-7 and W33 shared the same α-gal recognition as M86, structures were solved for two human IGHV 3-7-derived α-gal binding clones, HKB7 and JEC1. These antibodies were isolated from two patients, one with MMA and one with both MMA and TA (*SI Appendix*, Tables S3 and S4, Donor IDs 11 and 1). Both IgM and IgG isotypes of HKB7 were identified and JEC1 was IgM.

In the HKB7 structure (Fig. 3), clear evidence of α-gal-derived electron density was observed in the antibody binding cleft (Fig. 3*B*). When superposed with the M86 protein fold (Fig. 3*A*, and with the α-gal moiety in *SI Appendix*, Fig. S7; M86, orange/tan; HKB7, green/yellow), the CDRH3 regions occupy different conformational space, reflecting both different lineage (murine vs. human) and the highly polyclonal CDRH3 responses observed. However, the overall CDRH1 binding mode of HKB7 is closely related to that of M86, with a binding cleft lined by aromatic residues, dominated by V_H residues W33, Y58, and F98 (see Fig. 3*C* and *SI Appendix*, Fig. S8 for stereo depictions, and *SI Appendix*, Table S5 for buried surface). Hydrogen bonds are mostly evident at the interface between the sugar and the base of the cleft. The buried surface area, ~340 Å^2, is similar to that for M86, with the bulk contributed by W33 and F98 (both ~50 Å^2; *SI Appendix*, Table S5). As for M86, the centroids of aromatic rings participate in carbon pi and oxygen donor pi interactions (Fig. 3*D*). By contrast, in the JEC1 structure, residual electron density at the antibody binding cleft better fitted 2-methyl-2,4-pentanediol (MPD; a component of the crystallographic screen present at extremely high concentration; see

SI Appendix, Fig. S9). However, the cleft is, otherwise, highly similar to that of HKB7, lined by the side chains of V_H aromatic residues 33, 58, and 98. Indeed, placement of α1-3 galactobiose as a consequence of superposition with HKB7 provided an adequate fit into the JEC1 cleft (Fig. 3*E*; HKB7 in green/yellow, JEC1 in dark/light blue, α-gal sticks from HKB7).

Affinity Measurements and Site-Directed Mutagenesis. Next, we investigated in vitro binding affinities of the structurally characterized antibodies as well as another human α-gal binding clone carrying W33, HKD8 (germline IGHV3-15, IgA isotype, sequence comparison in *SI Appendix*, *Sequences*), both in monovalent (1:1) and multivalent binding formats using biolayer interferometry (BLI) (Fig. 4). Both formats required a longer (14 carbon vs. 3 carbon) linker tethering the α-gal component and the bovine serum albumin (BSA) substrate. When monovalent soluble Fabs were interrogated, the antibodies all bound to α-galactobiose-conjugated BSA (α-galactobiose-BSA) immobilized on the sensor with equilibrium binding constants in the micromolar range (M86 < HKD8 < HKB7 ≪ JEC1; Fig. 4, *Left*). This affinity range was confirmed by isothermal calorimetry performed with the M86 Fab and uncoupled α-galactobiose, which reported an affinity of ∼100 μM (*SI Appendix*, Fig. S10). Such a low monovalent affinity range is commonly observed for antibody carbohydrate interactions. In marked contrast, when soluble multivalent α-galactobiose-BSA (∼25 α-galactobiose per BSA) was interrogated by immobilized biotinylated Fab fragments on the surface of the biosensor, as a collective, the antibodies bound more potently (Fig. 4, *Center Left*). To further investigate the role of W33, alanine substitutions were introduced at position V_H33 (W33A); these substitutions abolished binding, even under multivalent conditions (Fig. 4, *Center Right*). Similar results indicating the importance of W33 were obtained when α-galactobiose-BSA was substituted with α-gal coupled to human serum albumin (α-gal-HSA; *SI Appendix*, Fig. S11). Further, when the α-galactobiose-BSA analyte was substituted with *N*-acetyllactosamine–BSA (α-gal missing the terminal galactose moiety), binding was also abolished (Fig. 4, *Right*).

Phage Display Selections. Given the apparent importance of W33, we further investigated whether introduction of the residue would endow α-gal binding capacity onto an otherwise different antibody germline. For this purpose, we utilized a previously reported antibody phage library based on IGHV3-23 (34), a commonly used germline in human antibody therapeutics. More specifically, we utilized Kunkel mutagenesis to modify position 33 in the library, restricting diversity to tryptophan or, alternatively, alanine only (*SI Appendix*, Table S6). The libraries were then selected against α-gal, using different carrier proteins for round 1 and round 2 of the selections (chemically modified BSA and mouse laminin, which is naturally rich in α-gal, respectively). This experiment indicated the enrichment of α-gal (cross-specific) signal after two rounds of selection for the W33 library only (gain of function) (Fig. 5*A*). Three of the selected clones were further characterized by sequencing (*SI Appendix*, *Sequences*), which revealed that, while the clones displayed diverse CDRH3 sequences, they shared a common "WXG" motif within CDRH1 (positions 33/35). As had been observed for the human- and mouse-derived antibodies, site-directed mutagenesis of the W33 position into A33 abolished binding to α-gal antigen (>100-fold signal reduction) (Fig. 5*B*).

Discussion

A structural appreciation of antibody–carbohydrate interactions is crucial to understanding how the immune system interacts with antigens containing carbohydrate entities. Early attempts to understand antibody–carbohydrate interfaces suggested that three or four carbohydrate units were required to interact with antigen-binding sites (35). Initial characterization of the α-gal epitope determined a requirement for three carbohydrate units, thereby defining the canonical α-gal entity (Galα1–3Galβ1–4GlcNAc) (29). The murine and human antibody-α-gal structures outlined here suggest that only the terminal two α1-3-linked galactosyl sugars are absolutely required, with the free-hemiacetyl hydroxyl (through which attachment to further sugars is mediated) positioned at the exterior of the binding groove (Figs. 1 and 3). Indeed, removal of the terminal galactose moiety abolished binding (Fig. 4, *Right*). In both mouse and human structures, aromatic residues line a shallow elongated cleft into which the terminal disaccharide slots, with hydrogen bonds connecting the spine of the sugar to the bottom of the cleft (Figs. 1 and 3).

The galactose moieties at the terminus of the α-gal epitope are not stereochemically equivalent. Examinations of monosaccharides bound by protein crystal structures (36–38) show that the α-face of β-D-galactose (Protein Data Bank [PDB] identifier GAL; the penultimate sugar of the α-gal epitope, indicated by an asterisk in Figs. 1 and 3 and *SI Appendix*, Fig. S2) outwardly projects a greater concentration of unencumbered electropositive CH atoms than its β-face, or by either the α- or β-faces of α-D-galactose (PBD identifier GLA, the terminal galactose of the α-gal moiety). Further, the more electropositive faces of monosaccharides are more likely to stack against the electronegative pi planes of aromatic residues especially tryptophan and to a lesser extent tyrosine and phenylalanine. Consistent with this thesis, in the mouse-derived M86 structure, it is the α-face of the penultimate GAL moiety that is positioned with its pyranose ring plane stacked against that of the side chain of W33, forming multiple carbon pi contacts (Fig. 1*D* and *SI Appendix*, Fig. S2). The β-face of this GAL moiety is relatively free of contact with the antibody. In contrast, despite lacking an overtly electropositive α- or β-face, the terminal GLA moiety still contributes carbon pi hydrogen bonds to the side chain of the neighboring tyrosine residue (Y32), but via a smaller end-on pose utilizing the triangular terminus (CH atoms at the C3, C4, and C5 positions) of the $^4C_1$ chair conformation of the pyranose ring (Fig. 1*D* and *SI Appendix*, Fig. S2). In the human-derived HKB7 structure, a similar binding theme is adopted. The penultimate GAL is positioned with its more electropositive α-face against the ring plane of W33 (Fig. 3*D*). In contrast to M86, the β-face of the sugar makes additional carbon pi, as well as oxygen pi, contacts with the side chain of F98. The terminal GLA moiety also contributes a mix of carbon and oxygen pi bonds, in this case with Y58, but, like M86, using a more end-on pose (Fig. 3*D*).

The PDB currently houses ∼120 structures of antibodies bound to carbohydrate ligands [SAbDab structural antibody database (39)]. A nonredundant set of ∼10, including our α-gal structures, employ a tryptophan residue at V_H33 (*SI Appendix*, Table S7). The indole face of W33 is positioned adjacent to a variety of sugars, where it contributes its electronegative pi face toward an electropositive face or end of the adjacent sugar, often in combination with additional CDR-derived aromatic residues. In several cases, galactose units are part of the ligand but not significantly contacted by the antibody. In only one case does W33 bind a galactose moiety [part of the $Lewis^X$ antigen, PDB entry 1uz8 (40)], with the α-face of the single GAL unit juxtaposed similarly as in our α-gal structures (*SI Appendix*, Table S7). Tryptophan aside, ∼15 structures contain a tyrosine residue at position V_H33, although in only one case does Y33 instigate carbon

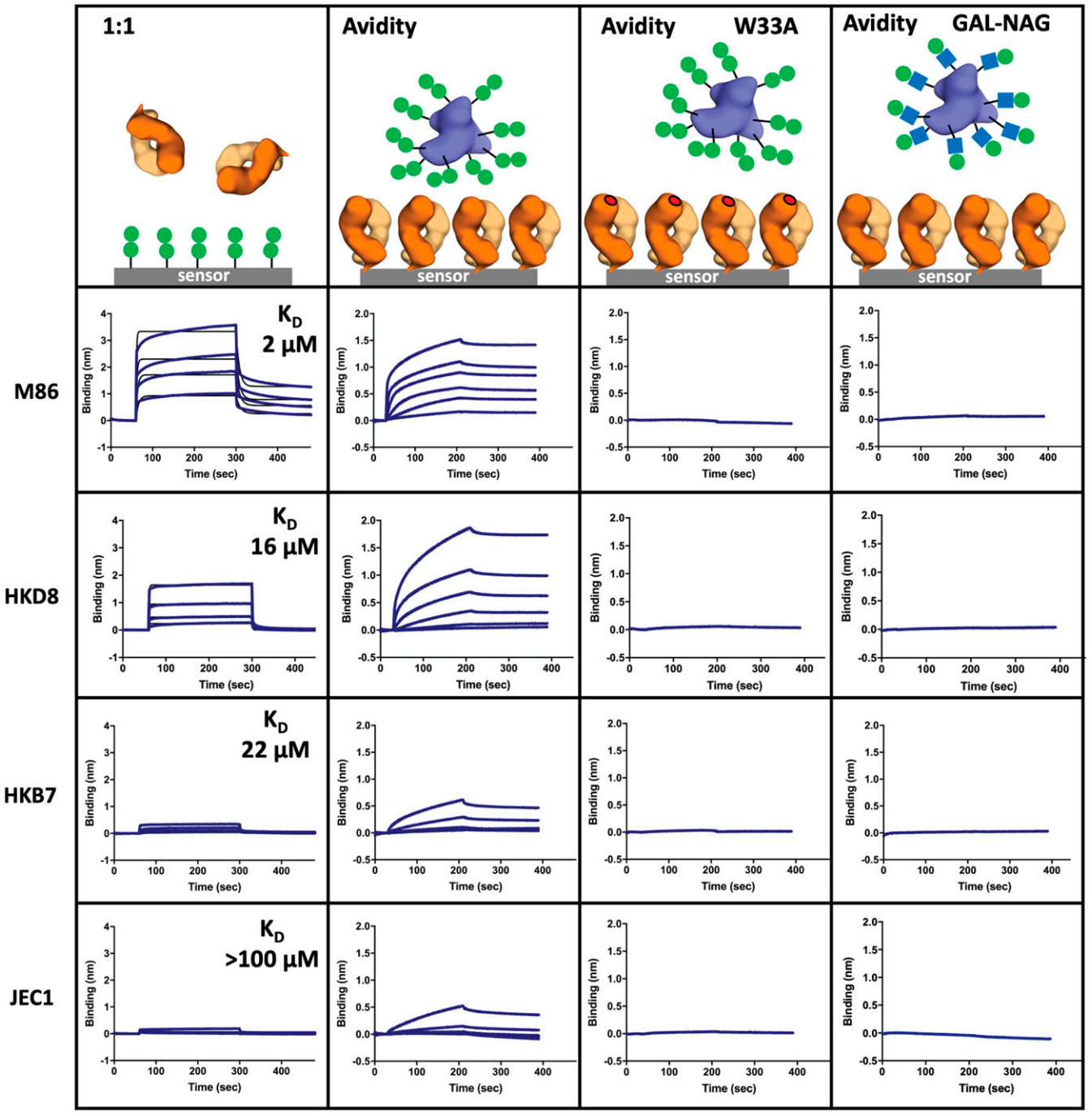

Fig. 4. Anti-α-gal affinity as determined by bilayer interferometry. Fab formats assayed comprised M86 (mouse hybridoma), HKD8 and HKB7 (human MMA patient), and JEC1 (human MMA and TA patient). At the top are cartoon depictions of experimental formats used to measure stoichiometric (1:1) and avidity-mediated binding. The 1:1 binding utilized biotinylated BSA-galactobiose conjugate attached to the biosensor (*Left*). Fab concentrations of 24, 12, 6, and 3 µM were incubated with the sensor. Avidity was gauged by flipping the format; biotinylated Fab was loaded onto streptavidin biosensors (heavy and light chains; orange and light orange) and used to capture multivalent BSA-galactobiose (GLA–GAL; green spheres) (*Center Left*) or BSA-*N*-acetyl-lactosamine (GAL–GlcNAc (PBD identifier NAG [N-Acetylglucosamine]); green sphere + blue square) conjugates (*Right*). For all four antibody Fabs, the W33A heavy-chain point mutant abolished binding (*Center Right*), and none were capable of binding *N*-acetyl-lactosamine, equivalent to the α-gal epitope missing the terminal galactose moiety.

pi contacts with the galactose unit [Lewis^{Y} antigen, PDB entry 1cly (41)].

Antibody interactions aside, the PDB houses another ∼1,000 protein–carbohydrate structures. This ensemble contains lectins, galectins, agglutinins, adhesins, and carbohydrate modification enzymes, and informs how non-Ig designs bind carbohydrate. Dozens bind a disaccharide whereby a GAL moiety is paired with α-D-glucopyranose (i.e., lactose), 2-acetamido-2-deoxy-β-D-galactose (T-antigen or galacto-*N*-biose), 2-acetamido-2-deoxy-α-D-galactose (TF-antigen), or 2-acetamido-2-deoxy-β-D-glucose (*N*-acetyllactoseamine; essentially α-gal missing the terminal GLA). In contrast to our antibody structures (shallow groove accommodating two carbohydrate units), the binding mode for most of these is a dimple predominantly accommodating just one carbohydrate residue, in most cases, GAL and involving an aromatic residue (mostly tryptophan) and a carbon pi interface with the α-face of the sugar [e.g., PDB entries 5cbl (42) and 4owl (43)]. Many structures contain galactose units as part of the trisaccharide or tetrasaccharide human blood group antigens A, B, and H, or variants such as Lewis^{A}, Lewis^{B}, Lewis^{X}, or Lewis^{Y}. These ligands

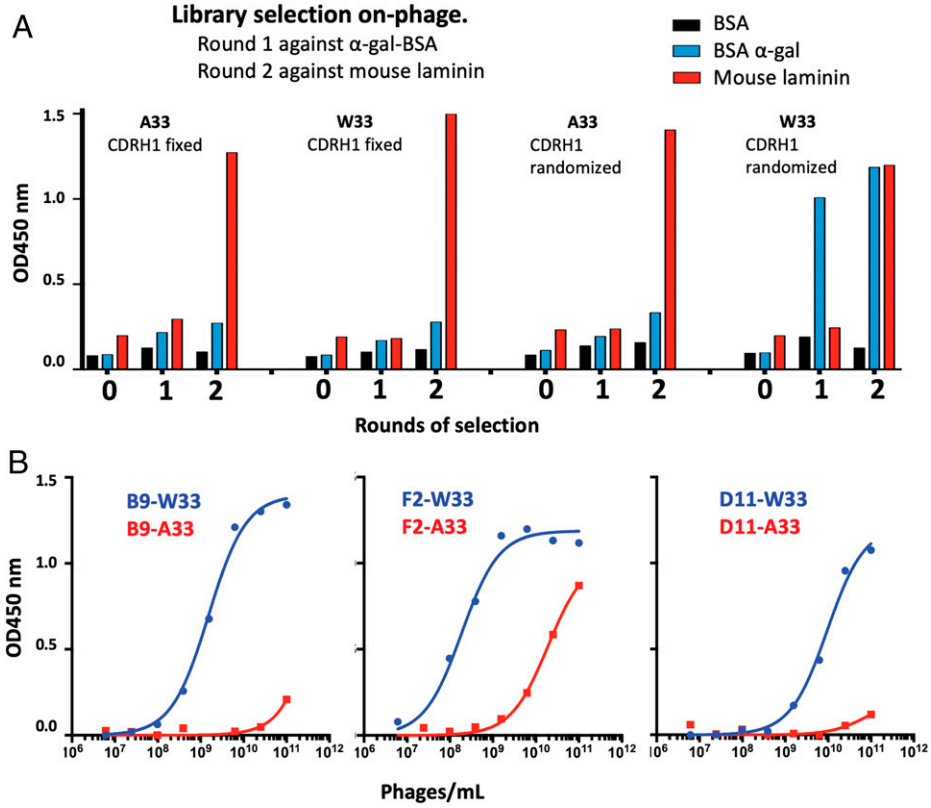

Fig. 5. V_H W33-primed phage libraries and anti-α-gal affinity. (*A*) Polyclonal phage ELISA response of scFv libraries containing V_H position 33 fixed to alanine or tryptophan through two rounds of selection. (*B*) The binding capacity of three V_H W33 containing anti-α-gal binding clones (B9, F2, and D11), could be abolished by a heavy-chain W33A mutation. OD450 nm, optical density at 450 nm.

are typically nonlinear, and the carbohydrate binding pocket is often small, with the interface dominated by a more deeply buried sugar, often fucose (FUC), with GAL or GLA components often contributing little contact [e.g., PDB entry 2obt (44)]. In other cases, especially B-group blood antigen binders (where the ligand includes GLA–GAL, as with α-gal), a larger binding pocket employs an aromatic residue (usually tryptophan) stacked against the α-face of the GAL unit [e.g., PDB entry 2ygm (45)]. Contrasting modes of binding (antibody vs. nonantibody) are evident in $Lewis^X$ structures; nonantibody $Lewis^X$ binders interface ligand primarily via the FUC component of the sugar, while, in the antibody binder, GAL is stacked against W33 employing carbon pi interactions. Outside of B-group blood antigen binders, PDB entries with consecutive galactose residues are infrequent and don't always employ the same α1-3 linkage as found in α-gal, although most involve a tryptophan–GAL interface (*SI Appendix*, Table S8); however, unlike the antibody structures, most employ a small binding pocket accommodating just the terminal (proximal) sugar unit (usually GAL, but sometimes GLA). In the cases of some glycosyl hydrolases, two or three consecutive GAL residues of poly-GAL substrates are captured in the substrate entry channels, each similarly interfaced with its own tryptophan residue [PDB entries 6q3r (46) and 2ccr (47), respectively].

Six structures contain trisaccharides either identical or very similar to α-gal itself (*SI Appendix*, Table S8, bottom rows). Two of these employ no carbon pi interfaces, and two employ carbon pi interfaces but bind primarily through the terminal GLA residue, while two unrelated fungal lectins bind in a manner similar to our antibody structures (superposed with M86 about the central GAL in Fig. 6). One of these binds via a shallow depression where significant contact is limited to the central GAL [Fig. 6*A*; PDB entry 1ule (48)], while the other employs a shallow elongated cleft that accommodates the terminal GLA–GAL residues [Fig. 6*B*; PDB entry 2iho (49)]. As with M86, in both cases, the GlcNAc (PBD identifier NAG) component at the reducing end of the sugar projects away from the surface and makes little contribution to binding. Both lectins employ a similarly positioned tryptophan (W72 or W35), which partakes in carbon pi interactions with the α-face and edge of the penultimate sugar, similar to the mode observed in the case of W33 presented by both the murine and human α-gal binding antibodies. Hence, the mode of binding displayed by our antibody-derived α-gal binders is similar to at least some of the non-Ig modes of α-gal binding noted to date.

The monovalent equilibrium dissociation constants (K_D) determined for M86, HKD8, HKB7, and JEC1 binding to α-gal are in the micromolar range (Fig. 4), below those commonly observed for antibody–protein interactions but typical for many antibody–carbohydrate interactions (50). In part, this reflects the small surface areas buried at the interfaces (~300 $Å^2$ to 350 $Å^2$). It may also reflect thermodynamics, wherein modest enthalpic gains due to binding are offset by high entropy costs associated with constraining the torsional flexibility of glycosidic bonds, and the rearrangement of water molecules associated with the solvated ligand and the binding cleft. Antigen accessibility is also known to be important (51), and a linker (14 atoms as compared to 3) between the galactobiose and the carrier protein (BSA) was absolutely required to enable binding. More potent binding was observed when assaying the antibodies for multivalent binding (Fig. 4), in agreement with the observation that the bulk of the α-gal response in mice (29) and humans is of the (pentameric, or hexameric in the absence of the J chain) IgM isotype (Fig. 2*A*). Lectins, similarly, are often

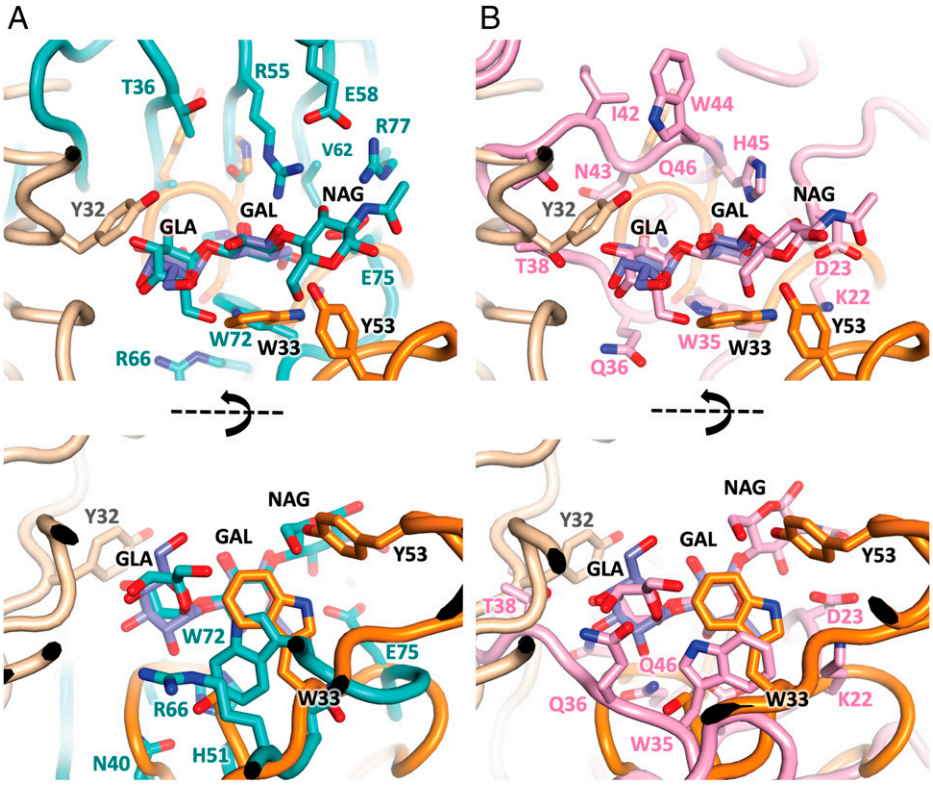

Fig. 6. Antibody and lectin α-gal binding modes. The M86 structure (orange Ig domains and blue galactobiose) superposed about the central GAL moiety with (*A*) galectin CGL2 from *Coprinopsis cinerea* (teal cartoon and ligand, PDB entry 1ule), and (*B*) lectin MOA from *Marasmius Oreades* (pink cartoon and ligand, PDB entry 2iho). Both employ a shallow elongated cleft and employ a tryptophan (W72 and W35, respectively) analogous to W33 in the M86 structure, adjacent to the α-face of the penultimate GAL unit. Perspectives from above (*Top*) and the side (*Bottom*) are shown.

presented as surface-associated clusters, underlying the importance of avidity to protein–carbohydrate interactions (52).

Flow cytometry revealed α-gal$^+$ B cells in healthy controls at a frequency consistent with levels reported previously (~1%) (2), although individual MMA patients with 3 to 4% were also observed (*SI Appendix*, Fig. S3*B*). A higher proportion of α-gal$^+$ B cells were unswitched memory (relative to α-gal$^-$ B cells), which have an increased propensity to differentiate into antibody-secreting plasmablasts compared to IgG-switched counterparts (53), indicating that α-gal$^+$ unswitched memory B cells may be important in the early response against α-gal-cloaked pathogens such as malaria. Sequencing of α-gal$^+$ B cells from controls and MMA patients revealed a polyclonal heavy-chain gene response. However, it is possible that a portion of the polyclonal response observed may result from nonspecific receptors recognizing a range of carbohydrate moieties, thus contributing to a low affinity α-gal response. Nevertheless, IGHV germline 3-7 antibodies were increased, relative to other IGHV, particularly in the memory compartment (Fig. 2 and *SI Appendix*, Fig. S5). Here, 3-7 was used by 11 to 25% of α-gal$^+$ memory B cells (Fig. 2), in contrast to a prior study using massively parallel B cell Ig sequencing in which 3-7 was used by less than 2% of mature naïve or switched memory B cells and ~5% of IgM memory B cells (54).

Comparison of total α-gal$^+$ and α-gal$^-$ B cells from controls and patients confirmed a significant increase in the usage of 3-7, which contains V_H W33. These data support prior reports of human anti-α-gal clones using 3-7 (31), and of immunized α-gal knockout mice preferencing a heavy chain encoding W33 (30) (VH22.1, including M86 characterized here; germline alignments shown in *SI Appendix*, *Sequences*), suggesting that, as for humans, distinct antibody germlines preferably yield more α-gal binders (21).

However, W33 is not the sole determinant predisposing α-gal binding. Indeed, five other IGHV germlines encode W33 (3-15, 3-74, 4-4, 4-28, and 5-51, colored yellow in Fig. 2*C* and *SI Appendix*, Fig. S5), yet are often only moderately enriched, and lack the detectable bias for α-gal binding noted for 3-7. This might reflect sample size, or differences in other positions critical to binding, such as V_H position 50. HKB7 and JEC1 contain aspartic acid and asparagine at V_H 50. In HKB7, the projected side chain of V_H D50 contributes a hydrogen bond to α-gal (Fig. 3*C*). Other W33-encoding IGHV germlines contain different residues at position 50, such as arginine (3-15, 3-74), glutamate (4-4), or tyrosine (4-28), whose larger side chains might sterically occlude the surface pocket characterized in our structures, and thus impede binding. Indeed, the diversity of α-gal binders observed, combined with insights from other protein–glycan structures (*SI Appendix*, Tables S7 and S8), suggests that different combinations of residues might achieve low-affinity binding to a variety of carbohydrate ligands.

The α-gal antibody response was remarkably similar in healthy individuals and MMA patients, reporting similar IGHV germlines, rates of somatic hypermutation, CDRH3 lengths, and B cell phenotype. Moreover, anti-α-gal IgE, associated with MMA, were also absent in patients and healthy donors. These data may be explained by the overall low frequency of IgE-expressing B cells in blood, the possibility that pathogenic IgE B cells preferentially localize in tissue, and the timing of sampling (e.g., time since exposure to mammalian meat and anaphylactic episode). A better appreciation of factors that prompt class switching to IgE will be required to more fully understand the pathogenicity of anti-α-gal antibodies.

The findings outlined here provide evidence that germline restriction, and, in particular, the presence of an IGHV3-7

germline-encoded tryptophan residue (W33), characterizes the human anti-α-gal antibody response. We observed not only that this residue was essential for binding for all of the analyzed antibodies (as indicated by mutation) but also that the introduction of a W33 residue into an otherwise unrelated germline (IGHV3-23) endowed α-gal binding onto an existing naïve antibody phage display library. The ability to produce antibodies with α-gal reactivity in a defined fashion may enable development of therapeutic anti-α-gal antibodies, for instance, for protection against malaria or for inhibition of pathogenic anti-α-gal antibodies in meat anaphylaxis and xenotransplant rejection. With IGHV3-23 representing by far the most commonly used germline in human antibody therapeutics (including Herceptin, Humira, Avastin, and others) (55), our findings may open up strategies toward antibody therapeutics development for α-gal targeted human applications including in MMA, malaria, and xenotransplantation.

Materials and Methods

Fab Expression and Purification. The Fab heavy- and light-chain sequences of mouse monoclonal M86 (29) were cloned into separate pCEP4 expression vectors (Thermo Fisher Scientific) via *Kpn*I and *Bam*H1 restriction sites. Fab arms of HKB7 and JEC1 were cloned accordingly. Heavy chains were C-terminally hexahistidine tagged. Transient expression involved transfecting Expi293 cells using a 2:1 ratio of heavy- to light-chain vectors (Thermo Fisher Scientific). Fabs were purified from cell culture supernatant using HisTrap FF Crude columns (GE Healthcare), then purified via gel filtration using an S200 26/60 column plumbed with 25 mM Tris (pH 8.0), 100 mM NaCl, as the running buffer. Fab fractions were pooled and concentrated to ∼1.5 mg/mL with spin filters (Amicon Ultracel 10KMWCO, Merck).

Cohorts. Noncoagulated blood was obtained from the Australian Red Cross Blood Service ($n = 12$), from additional controls unexposed to tick bites ($n = 2$), and from patients with MMA ($n = 3$) or MMA and TA ($n = 1$) (*SI Appendix*, Tables S3 and S4). MMA was diagnosed only after recurrent episodes of anaphylaxis delayed by 3 h to 6 h following the ingestion of mammalian meat, associated history of tick bite, and demonstration of α-gal-specific IgE at titers of <0.35 kUA/L. Diagnosis was not based on the presence of serum anti-α-gal IgE alone (56). Peripheral blood mononuclear cells (PBMCs) were prepared with Ficoll Histopaque and cryopreserved in Roswell Park Memorial Institute Medium (RPMI) 1640, 50% heat-inactivated human serum, and 10% dimethyl sulfoxide. The study was approved by the human research ethics committees of St Vincent's Hospital and Royal North Shore Hospital; written informed consent was obtained from all participants.

Fluorescence-Activated Cell Sorting. Alexa Fluor 647 BSA-α-gal conjugate was prepared from Alexa Fluor 647 NHS ester (Thermo Fisher Scientific) and α-galactobiose-BSA (Dextra) as per the manufacturer's instructions. Conjugate (0.25 μM to 0.5 μM) was premixed with fluorescently conjugated antibodies (CD19, CD20, IgD, CD27, CD38, viability dye, CD3, CD14) and added to thawed PBMCs for 30 min at 4 °C. $CD19^+$ α-gal^+ or $CD19^+$ α-gal^- cells were single-cell sorted using a fluorescence-activated cell sorting Aria III (BD Biosciences) into 96-well plates containing 5 μL of 10 mM Tris, pH 8.0, and 1 unit/μL RNasin. Plates were stored at −80 °C for 1 wk to 4 wk prior to processing.

Immunoglobulin Sequencing. Paired heavy- and light-chain immunoglobin genes were amplified from single sorted B cells using forward primers specific for heavy- and light-chain variable region leader sequences and reverse primers targeting α, γ, μ, ε, κ, and λ constant regions as described (57). PCR products were Sanger sequenced using 3′ nested PCR primers (57) and analyzed using MiXCR software (58). Selected paired heavy- and light-chain sequences were expressed as Fabs as outlined above. Raw sequencing reads of heavy-chain V regions are available in *SI Appendix*.

Fab-α-gal Complexes. The α1-3 galactobiose (Dextra) was dissolved in water and spiked into Fab samples such that the molar ratio of α-gal:Fab was ∼30:1, after which samples were concentrated (10KMWCO spin filters) to ∼6 mg/mL to 8 mg/mL for crystallization.

Crystallography. Crystallization experiments used a Mosquito robot (TTP Labtech) to combine equal volumes (400 nL) of protein and well solution in a 96-well sitting-drop format. M86 crystals grew in SaltRx condition E10 (Hampton Research). Crystals were optimized in a 24-well hanging-drop format combining equal volumes (2 μL) of protein and well solution (1.5 M $NaH_2PO_4.H_20.K_2HPO_4$, pH 8.0). JEC1 crystals grew in JCSG+ condition D5 (Molecular Dimensions) and were optimized using 100 mM Hepes (pH 7.4), 65% (vol/vol) MPD. HKB7 crystals grew in JCSG+ condition C2 (1M $LiCl_2$, 100 mM citrate [pH 4.0] 20% [wt/vol] PEG6000). Crystals were plunge frozen into liquid nitrogen.

Diffraction data were recorded at Australian Synchrotron beamlines MX1 (JEC1) and MX2 (M86, HKB7) using ADSC Quantum 315r (MX1) or Dectris Eiger X16M detectors (MX2); 360° sweeps of data were recorded in 1° oscillations (MX1), or deconvoluted into 3,600*0.1° images (MX2). Reflections were indexed and integrated using iMOSFLM (M86 and JEC1) (59) or X-ray detector software (HKB7) (60). Space groups were determined with Pointless (61). and scaling and merging were performed with AIMLESS (62), both components of CCP4 (63). Structures were determined by molecular replacement using Phaser (64), where a single Fab arm (PDB 1n8z) was first split into variable and constant pairings (VH+VL and CH1+CL1). Rigid body and restrained B-factor refinement were performed with a combination of REFMAC5 (65) and Phenix.Refine (66), interspersed with manual inspection and modification using Coot (67). Waters were modeled if satisfying hydrogen bond networks. Electron density consistent with a α-gal moiety was immediately evident in the M86 and HKB7 structures. The α1-3 linked galactobiose was modeled into this feature late in refinement. In the case of JEC1, residual density existed; however, being smaller in volume, it more likely represented multiple conformations of MPD, which comprised 65% (vol/vol) of the crystallization condition (not included in the final model). The HKB7 structure could be solved in two related space groups, C2221 and C2 (*b* angle measuring 90.00° or 90.91°), with one and two Fabs present in the asymmetric unit. Although models refined similarly, sufficient differences between Fab molecules for the lower symmetry (C2) space group (with relaxed noncrystallographic symmetry restraints) prompted PDB submission of the C2 model.

Affinity Measurements. Kinetic association and dissociation constants of the interactions between M86, JEC1, HKD8, and HKB7 antibodies and conjugated α-gal variants were measured using BLI using streptavidin biosensors on a BLItz system (ForteBio). When testing 1:1 binding affinity, biotinylated α-galactobiose-BSA (14-atom linker, Dextra) was loaded onto biosensors, and binding with different concentrations of Fab was measured. Kinetics were determined by global fitting using BLItzPro1.2. When testing binding in an avidity context, biotinylated Fab was loaded onto biosensors, and binding was measured with different concentrations of α-galactobiose-BSA, *N*-acetyl-lactosamine-BSA (14 atom linker, Dextra), or α-gal-HSA (14 atom linker, Carbosynth). In all cases, sensor tips were prewashed in unconjugated BSA.

Isothermal Titration Calorimetry. Protein was dialyzed overnight at 4 °C into phosphate-buffered saline (PBS) and filtered (0.22 μM) prior to use. The α1-3 galactobiose (Dextra) was dissolved in the same buffer. Isothermal titration calorimetry was performed at 25 °C using an i200 MicroCal microcalorimeter (ATA Scientific). Ligand was used at a concentration of 1 mM, and M86 Fab was used at a concentration of 52 μM. Baseline data were measured by titration of ligand into buffer and subtracted from experimental data. Data were analyzed using Origin 7.0 isothermal titration calorimetry Data analysis software (MicroCal).

Phage Library Construction. The Garvan2 scFv library (based on human heavy-chain DP47 [IGHV3-23] and light-chain DPK9 germlines) (75), was used to construct further libraries by Kunkel mutagenesis (68). For each library, modifications were restricted to CDRH1. Position 33 was either Tryptophane (W) or Alanine (A), with CDR diversity at other positions as in the parental library (*SI Appendix*, Table S6).

Phage Selections. Two rounds of phage display selections were carried out essentially as previously described (69), alternating the antigen between α-gal-conjugated BSA (α-gal-BSA, 14 atom linker, Dextra) and mouse laminin (Sigma-Aldrich), a glycoprotein rich in α-gal. In round 1, 10^{12} phages from each library were deselected against 3% (wt/vol) unconjugated BSA, then selected against

96-well MaxiSorp (Nunc) ELISA plate wells coated with 250 nM α-gal-BSA and blocked with PBS supplemented with 0.1% Tween-20 and 4% skim milk powder (MPBST). Bound phages were eluted with trypsin, then used to infect TG1 bacteria from which phages were harvested for the next round. Phage outputs for W33 and A33 libraries were 5×10^8 and 2×10^8, respectively. For round 2, 10^{12} phage recovered from round 1 were similarly blocked, then selected against wells coated with 100 nM mouse laminin. Outputs for W33 and A33 libraries were 4×10^9 and 4×10^8, respectively.

Phage ELISA. For phage ELISA, plate wells were coated with 150 nM BSA or α-gal-BSA overnight at 4 °C, then blocked for 2 h with MPBS; 1×10^9 purified phages were blocked in MPBST and incubated in wells for 1 h. Plates were washed in PBST, incubated with horseradish peroxidase–conjugated anti-M13 antibody (GE Healthcare) for 1 h, and washed 3 times with PBST. Plates were then incubated with 3,3′,5,5′-Tetramethylbenzidine substrate (Perkin-Elmer), the reactions quenched with HCl, and then detected at 450 nM (ClarioStar-BMG Labtech).

Statistical Analyses. Cohort level differences in V-region usage between α-gal$^+$ and α-gal$^-$ B cells were tested in patients and controls separately, and then between patients and controls (interaction test) by transforming counts and testing for differential usage using limma voom (70). In this case, individual V-region counts were treated as genes from an RNA sequencing experiment. Relative V-region enrichment of α-gal$^+$ B cells relative to α-gal$^-$ B cells within the same donor was calculated using a c^2 test. In all cases, false discovery correction (Benjamini–Hochberg) was performed on per-region P values, and false discovery rate values of <0.05 were considered significant.

Data Availability. The coordinates for the crystal structures of M86, HKB7, and JEC1 have been deposited in the Protein Data Bank as entries 7UEN (72), 7UEM (73), and 7UEL (74), respectively. Raw sequencing data have been deposited at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject PRJNA851943 and can be accessed at http://www.ncbi.nlm.nih.gov/bioproject/851943 (75). All processed sequencing data is included in the supplementary information.

ACKNOWLEDGMENTS. Crystallographic data were recorded at the Australian Synchrotron, part of the Australian Nuclear Science and Technology Organisation. We also thank the Garvan-Weizmann Centre for Cellular Genomics and Garvan Molecular Genetics facilities and Megan Faulks for technical services, and control and MMA/TA patients who contributed samples to the study. The work was supported by National Health and Medical Research Council (NHMRC) Program Grant 1113904 (C.C.G. and D.C.), NHMRC Project Grants 1142186 (K.J.L.J. and J.H.R.) and 1108800 (C.C.G.), NHMRC Fellowships 1157744 (D.C.) and 1081858 (C.C.G.), and New South Wales Health Early- to Mid-Career Fellowship (J.H.R.).

Author affiliations: [a]Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia; [b]Tick-induced Allergies Research and Awareness Centre, Sydney, NSW 2065, Australia; [c]School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2006, Australia; [d]St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Sydney, NSW 2010, Australia; [e]School of Medical Sciences, University of New South Wales, Sydney, NSW 2052, Australia; [f]Cellular Genomics Futures Institute, University of New South Wales, Sydney, NSW 2052, Australia; and [g]Northern Clinical School, Sydney Medical School, Faculty of Medicine and Health, University of Sydney, Sydney, NSW 2065, Australia

1. U. Galili, B. A. Macher, J. Buehler, S. B. Shohet, Human natural anti-alpha-galactosyl IgG. II. The specific recognition of alpha (1—3)-linked galactose residues. *J. Exp. Med.* **162**, 573–582 (1985).
2. U. Galili, F. Anaraki, A. Thall, C. Hill-Black, M. Radic, One percent of human circulating B lymphocytes are capable of producing the natural anti-Gal antibody. *Blood* **82**, 2485–2493 (1993).
3. U. Galili, M. R. Clark, S. B. Shohet, J. Buehler, B. A. Macher, Evolutionary relationship between the natural anti-Gal antibody and the Gal alpha 1—3Gal epitope in primates. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1369–1373 (1987).
4. U. Galili, K. Swanson, Gene sequences suggest inactivation of alpha-1,3-galactosyltransferase in catarrhines after the divergence of apes from monkeys. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7401–7404 (1991).
5. J. L. Avila, M. Rojas, U. Galili, Immunogenic Gal alpha 1–3Gal carbohydrate epitopes are present on pathogenic American *Trypanosoma* and *Leishmania*. *J. Immunol.* **142**, 2828–2834 (1989).
6. P. M. Repik, J. M. Strizki, U. Galili, Differential host-dependent expression of alpha-galactosyl epitopes on viral glycoproteins: A study of eastern equine encephalitis virus as a model. *J. Gen. Virol.* **75**, 1177–1181 (1994).
7. U. Galili, Significance of the evolutionary α1,3-galactosyltransferase (GGTA1) gene inactivation in preventing extinction of apes and old world monkeys. *J. Mol. Evol.* **80**, 1–9 (2015).
8. B. Ravindran, A. K. Satapathy, M. K. Das, Naturally-occurring anti-alpha-galactosyl antibodies in human Plasmodium falciparum infections–A possible role for autoantibodies in malaria. *Immunol. Lett.* **19**, 137–141 (1988).
9. U. Galili, R. E. Mandrell, R. M. Hamadeh, S. B. Shohet, J. M. Griffiss, Interaction between human natural anti-alpha-galactosyl immunoglobulin G and bacteria of the human flora. *Infect. Immun.* **56**, 1730–1737 (1988).
10. U. Galili, A. Korkesh, I. Kahane, E. A. Rachmilewitz, Demonstration of a natural antigalactosyl IgG antibody on thalassemic red blood cells. *Blood* **61**, 1258–1264 (1983).
11. U. Galili, M. R. Clark, S. B. Shohet, Excessive binding of natural anti-alpha-galactosyl immunoglobin G to sickle erythrocytes may contribute to extravascular cell destruction. *J. Clin. Invest.* **77**, 27–33 (1986).
12. U. Galili, I. Flechner, A. Knyszynski, D. Danon, E. A. Rachmilewitz, The natural anti-alpha-galactosyl IgG on human normal senescent red blood cells. *Br. J. Haematol.* **62**, 317–324 (1986).
13. U. Galili, J. Buehler, S. B. Shohet, B. A. Macher, The human natural anti-Gal IgG. III. The subtlety of immune tolerance in man as demonstrated by crossreactivity between natural anti-Gal and anti-B antibodies. *J. Exp. Med.* **165**, 693–704 (1987).
14. T. Rispens, N. I. L. Derksen, S. P. Commins, T. A. Platts-Mills, R. C. Aalberse, IgE production to α-gal is accompanied by elevated levels of specific IgG1 antibodies and low amounts of IgE to blood group B. *PLoS One* **8**, e55566 (2013).
15. C. Hamsten *et al.*, Red meat allergy in Sweden: Association with tick sensitization and B-negative blood groups. *J. Allergy Clin. Immunol.* **132**, 1431–1434 (2013).
16. S. A. van Nunen, K. S. O'Conner, S. L. Fernando, L. R. Clarke, R. X. Boyle, An association between *Ixodes holocyclus* tick bite reactions and red meat allergy. *Int. Med. J.* **37**, A132 (2007).
17. S. A. Van Nunen, K. S. O'Connor, L. R. Clarke, R. X. Boyle, S. L. Fernando, An association between tick bite reactions and red meat allergy in humans. *Med. J. Aust.* **190**, 510–511 (2009).
18. S. P. Commins *et al.*, Delayed anaphylaxis, angioedema, or urticaria after consumption of red meat in patients with IgE antibodies specific for galactose-alpha-1,3-galactose. *J. Allergy Clin. Immunol.* **123**, 426–433 (2009).
19. C. H. Chung *et al.*, Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *N. Engl. J. Med.* **358**, 1109–1117 (2008).
20. S. A. van Nunen *et al.*, Alphagal and mammalian meat specific IgE in mammalian meat tolerant tick anaphylaxis sufferers. *Allergy Eur. J. Allergy Clin. Immunol.* **69**, 276–277 (2014).
21. F. Naso, A. Gandaglia, L. Iop, M. Spina, G. Gerosa, Alpha-Gal detectors in xenotransplantation research: A word of caution. *Xenotransplantation* **19**, 215–220 (2012).
22. S. A. van Nunen, Tick-induced allergies: Mammalian meat allergy and tick anaphylaxis. *Med. J. Aust.* **208**, 316–321 (2018).
23. R. N. Araujo *et al.*, *Amblyomma sculptum* tick saliva: α-Gal identification, antibody response and possible association with red meat allergy in Brazil. *Int. J. Parasitol.* **46**, 213–220 (2016).
24. A. Cabezas-Cruz *et al.*, Tick galactosyltransferases are involved in alpha-Gal synthesis and play a role during *Anaplasma phagocytophilum* infection and *Ixodes scapularis* tick vector development. *Sci Rep-Uk* **8** (2018).
25. J. L. Chandrasekhar *et al.*, Cutaneous exposure to clinically relevant lone star ticks promotes IgE production and hypersensitivity through CD4$^+$ T cell- and MyD88-dependent pathways in mice. *J. Immunol.* **203**, 813–824 (2019).
26. J. M. Wilson *et al.*, IgE to the mammalian Oligosaccharide galactose-alpha-1,3-galactose is associated with increased atheroma volume and plaques with unstable characteristics - A brief report. *Arterioscler. Thromb. Vasc. Biol.* **38**, 1665–1669 (2018).
27. B. Yilmaz *et al.*, Gut microbiota elicits a protective immune response against malaria transmission. *Cell* **159**, 1277–1289 (2014).
28. U. Galili, C. B. Basbaum, S. B. Shohet, J. Buehler, B. A. Macher, Identification of erythrocyte Gal alpha 1-3Gal glycosphingolipids with a mouse monoclonal antibody, Gal-13. *J. Biol. Chem.* **262**, 4683–4688 (1987).
29. Z. C. Chen, M. Z. Radic, U. Galili, Genes coding evolutionary novel anti-carbohydrate antibodies: Studies on anti-Gal production in alpha 1,3galactosyltransferase knock out mice. *Mol. Immunol.* **37**, 455–466 (2000).
30. R. M. Hamadeh, U. Galili, P. Zhou, J. M. Griffiss, Anti-alpha-galactosyl immunoglobulin A (IgA), IgG, and IgM in human secretions. *Clin. Diagn. Lab. Immunol.* **2**, 125–131 (1995).
31. L. Wang, M. Z. Radic, U. Galili, Human anti-Gal heavy chain genes. Preferential use of VH3 and the presence of somatic mutations. *J. Immunol.* **155**, 1276–1285 (1995).
32. T. T. Wu, E. A. Kabat, An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* **132**, 211–250 (1970).
33. K. R. Abhinandan, A. C. Martin, Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol. Immunol.* **45**, 3832–3839 (2008).
34. K. Dudgeon *et al.*, General strategy for the generation of human antibody variable domains with increased aggregation resistance. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10879–10884 (2012).
35. E. A. Kabat, Heterogeneity in extent of the combining regions of human antidextran. *J. Immunol.* **77**, 377–385 (1956).
36. W. Chen *et al.*, Structural and energetic basis of carbohydrate-aromatic packing interactions in proteins. *J. Am. Chem. Soc.* **135**, 9877–9884 (2013).
37. K. L. Hudson *et al.*, Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **137**, 15152–15160 (2015).
38. I. M. Stanković, J. P. Blagojević Filipović, S. D. Zarić, Carbohydrate - Protein aromatic ring interactions beyond CH/π interactions: A Protein Data Bank survey and quantum chemical calculations. *Int. J. Biol. Macromol.* **157**, 1–9 (2020).
39. J. Dunbar *et al.*, SAbDab: The structural antibody database. *Nucleic Acids Res.* **42**, D1140–D1146 (2014).
40. A. M. van Roon *et al.*, Structure of an anti-Lewis X Fab fragment in complex with its Lewis X antigen. *Structure* **12**, 1227–1236 (2004).
41. P. D. Jeffrey *et al.*, The x-ray structure of an anti-tumour antibody in complex with antigen. *Nat. Struct. Biol.* **2**, 466–471 (1995).
42. J. K. Rustiguel, P. S. Kumagai, M. Dias-Baruffi, A. J. Costa-Filho, M. C. Nonato, Recombinant expression, purification and preliminary biophysical and structural studies of C-terminal carbohydrate recognition domain from human galectin-4. *Protein Expr. Purif.* **118**, 39–48 (2016).

43. K. Kaus, J. W. Lary, J. L. Cole, R. Olson, Glycan specificity of the *Vibrio vulnificus* hemolysin lectin outlines evolutionary history of membrane targeting by a toxin family. *J. Mol. Biol.* **426**, 2800–2812 (2014).
44. S. Cao *et al.*, Structural basis for the recognition of blood group trisaccharides by norovirus. *J. Virol.* **81**, 5949–5957 (2007).
45. M. A. Higgins, E. Ficko-Blean, P. J. Meloncelli, T. L. Lowary, A. B. Boraston, The overall architecture and receptor binding of pneumococcal carbohydrate-antigen-hydrolyzing enzymes. *J. Mol. Biol.* **411**, 1017–1036 (2011).
46. S. Torpenholt, J. C. N. Poulsen, S. J. Muderspach, L. De Maria, L. Lo Leggio, Structure of *Aspergillus aculeatus* β-1,4-galactanase in complex with galactobiose. *Acta Crystallogr. F Struct. Biol. Commun.* **75**, 399–404 (2019).
47. J. Le Nours *et al.*, Investigating the binding of beta-1,4-galactan to *Bacillus licheniformis* beta-1,4-galactanase by crystallography and computational modeling. *Proteins* **75**, 977–989 (2009).
48. P. J. Walser *et al.*, Structure and functional analysis of the fungal galectin CGL2. *Structure* **12**, 689–702 (2004).
49. E. Grahn *et al.*, Crystal structure of the *Marasmius oreades* mushroom lectin in complex with a xenotransplantation epitope. *J. Mol. Biol.* **369**, 710–721 (2007).
50. O. Haji-Ghassemi, R. J. Blackler, N. Martin Young, S. V. Evans, Antibody recognition of carbohydrate epitopes. *Glycobiology* **25**, 920–952 (2015).
51. J. J. Lammerts van Bueren *et al.*, Anti-galactose-α-1,3-galactose IgE from allergic patients does not bind α-galactosylated glycans on intact therapeutic antibody Fc domains. *Nat. Biotechnol.* **29**, 574–576 (2011).
52. W. Van Breedam, S. Pöhlmann, H. W. Favoreel, R. J. de Groot, H. J. Nauwynck, Bitter-sweet symphony: Glycan-lectin interactions in virus biology. *FEMS Microbiol. Rev.* **38**, 598–632 (2014).
53. M. Seifert *et al.*, Functional capacities of human IgM memory B cells in early inflammatory responses and secondary germinal center reactions. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E546–E555 (2015).
54. Y. C. Wu *et al.*, High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. *Blood* **116**, 1070–1078 (2010).
55. T. Jain *et al.*, Biophysical properties of the clinical-stage antibody landscape. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 944–949 (2017).
56. J. de la Fuente, A. Cabezas-Cruz, I. Pacheco, Alpha-gal syndrome: Challenges to understanding sensitization and clinical reactions to alpha-gal. *Expert Rev. Mol. Diagn.* **20**, 905–911 (2020).
57. P. Schofield *et al.*, "Sequencing and affinity determination of antigen-specific B lymphocytes from peripheral blood" in *Antibody Engineering, Methods and Protocols*, D. Nevoltris, P. Chames, Eds. (Humana Press, 2018), chap. 15.
58. D. A. Bolotin *et al.*, MiXCR: Software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).
59. T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. W. Leslie, iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).
60. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
61. P. R. Evans, An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 (2011).
62. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
63. M. D. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
64. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
65. G. N. Murshudov *et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
66. P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
67. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of COOT. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
68. R. Rouet, K. Dudgeon, D. Christ, Generation of human single domain antibody repertoires by Kunkel mutagenesis. *Methods Mol. Biol.* **907**, 195–209 (2012).
69. C. M. Lee, N. Iorno, F. Sierro, D. Christ, Selection of human antibody fragments by phage display. *Nat. Protoc.* **2**, 3001–3008 (2007).
70. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
71. H. C. Jubb *et al.*, Arpeggio: A web server for calculating and visualising interatomic interactions in protein structures. *J. Mol. Biol.* **429**, 365–371 (2017).
72. D. B. Langley, D. Christ, 7UEN, Genetic and structural basis of the human anti-alpha-galactosyl antibody response. Protein Data Bank. https://www.rcsb.org/structure/7UEN. Deposited 22 March 2022.
73. D. B. Langley, D. Christ, 7UEM, Genetic and structural basis for the human anti-alpha-galactosyl antibody response. Protein Data Bank. https://www.rcsb.org/structure/7UEM. Deposited 22 March 2022.
74. D. B. Langley, D. Christ, 7UEL, Genetic and structural basis for the human anti-alpha-galactosyl antibody response. Protein Data Bank. https://www.rcsb.org/structure/7UEL. Deposited 22 March 2022.
75. D. B. Langley *et al.*, Raw sequencing data for "Genetic and structural basis of the human anti-α-galactosyl antibody response." Accession # 851943. NCBI BioProject. https://www.ncbi.nlm.nih.gov/bioproject/851943. Accessed 6 July 2022.