The Journal of Immunology

Structures of Preferred Human IgV Genes–Based Protective Antibodies Identify How Conserved Residues Contact Diverse Antigens and Assign Source of Specificity to CDR3 Loop Variation

Steve Bryson, Christy A. Thomson, Louise F. Risnes, Somnath Dasgupta, Kenneth Smith, John W. Schrader and Emil F. Pai

J Immunol 2016; 196:4723-4730; Prepublished online 29 April 2016;
doi: 10.4049/jimmunol.1402890
http://www.jimmunol.org/content/196/11/4723

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Structures of Preferred Human IgV Genes–Based Protective Antibodies Identify How Conserved Residues Contact Diverse Antigens and Assign Source of Specificity to CDR3 Loop Variation

Steve Bryson,[*,†] Christy A. Thomson,[‡,1] Louise F. Risnes,[†,2] Somnath Dasgupta,[§,3] Kenneth Smith,[¶] John W. Schrader,[‡] and Emil F. Pai[*,†,‖,#]

The human Ab response to certain pathogens is oligoclonal, with preferred IgV genes being used more frequently than others. A pair of such preferred genes, *IGVK3-11* and *IGVH3-30*, contributes to the generation of protective Abs directed against the 23F serotype of the pneumonococcal capsular polysaccharide of *Streptococcus pneumoniae* and against the AD-2S1 peptide of the gB membrane protein of human CMV. Structural analyses of Fab fragments of mAbs 023.102 and pn132p2C05 in complex with portions of the 23F polysaccharide revealed five germline-encoded residues in contact with the key component, L-rhamnose. In the case of the AD-2S1 peptide, the KE5 Fab fragment complex identified nine germline-encoded contact residues. Two of these germline-encoded residues, Arg91L and Trp94L, contact both the L-rhamnose and the AD-2S1 peptide. Comparison of the respective paratopes that bind to carbohydrate and protein reveals that stochastic diversity in both CDR3 loops alone almost exclusively accounts for their divergent specificity. Combined evolutionary pressure by human CMV and the 23F serotype of *S. pneumoniae* acted on the IGVK3-11 and IGVH3-30 genes as demonstrated by the multiple germline-encoded amino acids that contact both L-rhamnose and AD-2S1 peptide. *The Journal of Immunology*, 2016, 196: 4723–4730.

Antibodies are composed of two identical H chains and two identical L chains, each encoded by a somatically assembled gene. The epitope-combining site is formed by six variable peptide loops, the CDRs, three from the H chain and three from the L chain. In newly arising human B lymphocytes, the H chain gene is formed by combinatorial recombination of 55 *IGHV* genes with one of 27 *IGHD* genes and 6 *IGHJ* genes, and the L chain gene is formed by assembling one of either 40 *IGKV* or 30 *IGLV* genes, with one of either 5 *IGKJ* or 4 *IGLJ* genes. The V genes encode CDR1 and CDR2 of their respective H chains or L chains, and the recombined D and J elements with their stochastically generated features encode the CDR3s. Additional diversity in the CDR3s is contributed by variation in the precise sites of junctions and the action of DNA repair enzymes that subtract or add nucleotides to the joints between the recombined genes (1–4). The resultant diversity of the CDR3s contributes the diversity of the epitope-combining sites of Igs and further increases their repertoire of binding specificity.

Multiple indications suggest intense evolutionary selection on the human Ig *IGHV*, *IGKV*, or *IGLV* genes (5, 6). Strong evidence supports that many human Abs against particular pathogen epitopes use preferred *IGHV, IGKV*, or *IGLV* genes (7–12). With the ability to generate a multitude of possible different Ig molecules (10^6–10^8), is there any selective advantage in the maintenance and preferential use of certain V genes, and why is a restricted oligoclonal response observed to particular pathogens? It has been speculated that evolutionary pressure from important pathogens could retain certain IgV genes and that IgV genes encode in the germline amino acid residues that contact epitopes on a pathogen (13, 14, 15). However, two necessary conditions apply. First, the Abs encoded by these IgV genes must provide protection against the pathogen, and second, the germline-encoded contacts should facilitate the production of protective Abs.

Streptococcus pneumoniae is a common bacterial pathogen and causes serious disease and death worldwide from conditions such as pneumonia, meningitis, and bacteremia (16). The pneumococcal capsular polysaccharides (PPS) distinguish the different *S. pneumoniae* serotypes, with Abs to PPS aiding in the opsonization of the bacteria

*Princess Margaret Cancer Centre, Toronto, Ontario M5G 1L7, Canada; [†]Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada; [‡]Biomedical Research Centre, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada; [§]Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada; [¶]Oklahoma Medical Research Foundation, Oklahoma City, OK 73104; [‖]Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5S 1A8, Canada; and [#]Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5G 1L7, Canada

[1]Current address: Amgen Inc., Burnaby, BC, Canada.

[2]Current address: Department of Immunology, Centre for Immune Regulation, Oslo University Hospital Rikshospitalet, University of Oslo, Oslo, Norway.

[3]Current address: Department of Chemistry, Alberta Glycomics Centre, University of Alberta, Edmonton, AB, Canada.

ORCIDs: 0000-0002-3381-3695 (J.W.S.); 0000-0002-1162-7242 (E.F.P.).

Received for publication November 17, 2014. Accepted for publication April 1, 2016.

This work was supported by the Canadian Institutes for Health Research–University of British Columbia Strategic Training Program for Translational Research in Infectious Diseases (to C.A.T.), the Canada Research Chairs Program (to E.F.P. and J.W.S.), Canadian Institutes of Health Research Grant MOP-179283, and by the Ontario Ministry of Health and Long-Term Care. Part of the research in this study was performed at the Canadian Light Source (Saskatoon, SK, Canada), which is supported by the Natural Sciences and Engineering Research Council of Canada, the National Research Council Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada, and by the University of Saskatchewan. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Argonne National Laboratory, was supported by the U.S. Department of Energy Contract DE-AC02-06CH11357.

Address correspondence and reprint requests to Dr. Emil F. Pai, University of Toronto, Departments of Biochemistry, Medical Biophysics, and Molecular Genetics, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada. E-mail address: pai@hera.med.utoronto.ca

The online version of this article contains supplemental material.

Abbreviations used in this article: HCMV, human CMV; PPS, pneumococcal capsular polysaccharide; RG, methyl-α-L-rhamnopyranosyl-(1→2)-β-D-galactopyranoside; RGP, methyl-α-L-rhamnopyranosyl-(1→2)-3-(glycer-2-yl-phosphate)-β-D-galactopyranoside.

and their clearance by complement-dependent phagocytosis (17). Vaccination with PPS or protein-conjugated PPS containing up to 23 different serotypes is recommended for small children, the elderly, and other susceptible populations. The *IGHV3-30* and *IGKV3-11* V gene pair (Fig. 1) has been found to be overrepresented in protective Abs targeting the 23F serotype of PPS (Fig. 2A). Although the *IGHV3-30* and *IGKV3-11* V genes are commonly used in Abs [~9 and 4%, respectively (19)], their pairing in mAbs targeting the 23F PPS exceeds their normal distribution.

Human CMV (HCMV) is a prevalent virus that has coevolved with humans (20). Generally, 50–90% of the population is chronically infected with HCMV, with the initial infection causing mild symptoms (21). However, if a woman is infected with HCMV for the first time while pregnant, it can be a major cause of damage to the fetus (22). Similarly, for immunocompromised individuals or those individuals undergoing organ transplantation, HCMV is a threat. Thus, HCMV has been a major vaccine target by the World Health Organization for the past 20 y. When infected with HCMV, humans can mount a neutralizing Ab response against the gB envelope protein of the virus. Within the gB protein lies a linear epitope, AD-2S1, that neutralizing Abs target, disrupting the ability of the virus to fuse with host cell membranes (15, 23). Interestingly, independently originating human mAbs targeting AD-2S1 preferentially use the same pair of IgV genes, *IGHV3-30* and *IGKV3-11* (15), as those found for the 23F serotype of *S. pneumoniae*.

In this study, we describe the crystal structures of the Fab fragments of a hypermutated *S. pneumoniae* Ab, Fab 023.102, complexed with the core of its epitope, L-rhamnose, the disaccharide methyl-α-L-rhamnopyranosyl-(1→2)-β-D-galactopyranoside (RG), and the repeating unit of the PPS, methyl-α-L-rhamnopyranosyl-(1→2)-3-(glycer-2-yl-phosphate)-β-D-galactopyranoside (RGP) (Fig. 2B). We also present the structure of the L-rhamnose complex of the Fab fragment of another anti-23F Ab, pn132p2C05 (C05). Using methodology described by Smith et al. (17), this authentic human Ab with naturally paired H and L chains was isolated after vaccination with Pneumovax 23 from a donor different from the one providing 023.102. Finally, as a second representative of an anti-HCMV Ab, we show the crystal structure of the AD-2S1 Ag complex of the KE5 Fab, an Ab that also utilizes *IGKV3-11* and *IGHV3-30* genes but *IGHD* and *IGKJ* genes different from those used by a previously crystallized anti-HCMV Fab, 8F9 (24). In the present study, we provide structural evidence that the same germline IgV genes (*IGHV3-30* and *IGKV3-11*) can assume different roles in protective Abs against two very different Ags on two very different pathogens because of their distinct configurations when recombined with diverse CDR3 loops. Thus, evolutionary pressure by pathogens to retain certain IgV genes can collaborate with two stochastic processes, the recombination of CDR3 loops and somatic hypermutation, in the production of protective Abs.

Materials and Methods

Sample preparation

Recombinant human Fab 023.102 was expressed in *Escherichia coli* XL1-Blue cells (Agilent Technologies, Santa Clara, CA) using plasmid pARC (D. Reason, Children's Hospital Oakland Research Institute, Oakland, CA). The cells were grown in Super Broth media (32 g/l tryptone, 16 g/l yeast extract, and 5 g/l NaCl) with 0.1 mg/ml ampicillin at 37°C until early log phase (OD_{600} = 0.4–0.5). Isopropyl-β-D-thiogalactoside was added to 1 mM, the temperature was reduced to 26°C, and the culture was left to incubate overnight (20 h). The cells were collected by centrifugation and lysed by sonication. The Fab 023.102 protein was purified from the soluble cell extract using Profinity IMAC Ni-charged resin (Bio-Rad) followed by Superdex 200 size-exclusion chromatography (GE Healthcare). Prior to the Superdex 200 column run, the sample was reduced and alkylated by incubating it with 10 mM DTT, followed by 50 mM iodoacetamide, preventing inter-Fab disulfide formation. The final sample, representing a yield of 0.5–1 mg/l culture, was dialyzed against 20 mM Tris/HCl (pH 8.0) and concentrated to 8 mg/ml.

Recombinant human Fab pn132p2C05 was expressed in *E. coli* Shuffle cells (New England Biolabs) using plasmid pET-Duet (Novagen). The cells were grown in Super Broth media (32 g/l tryptone, 16 g/l yeast extract, and 5 g/l NaCl) with 0.1 mg/ml ampicillin at 37°C until early log phase (OD_{600} = 0.3–0.4). Isopropyl-β-D-thiogalactoside was added to 0.1 mM, the temperature was reduced to 18°C, and the culture was left to incubate for 2 d (44 h). The cells were collected by centrifugation and lysed by sonication. The Fab pn132p2C05 protein was purified from the soluble cell extract using KappaSelect affinity media (GE Life Sciences) followed by Mono S ion-exchange chromatography (GE Life Sciences). The final sample, representing a yield of 1–2 mg/l culture, was dialyzed against 20 mM Tris/HCl (pH 8.0) and concentrated to 10 mg/ml.

The Fab fragment of KE5 IgG (15) was prepared by incubation of the intact mAb with 0.05 mg/ml papain in 0.1 M sodium phosphate buffer (pH 6.0), 2 mM EDTA, and 10 mM DTT for 4 h at 37°C and purified by ion chromatography on DE52 cellulose (Whatman) at pH 8.0, followed by a Mono S column (GE Life Sciences) at pH 5.6. After elution with a KCl gradient, Fab KE5 (10 mg/ml) was stored in 20 mM TRIS buffer (pH 8.0).

RG and RGP were synthesized as published (18). For stability, the C1′ hydroxyl of D-galactose was methylated. L-Rhamnose was obtained from Sigma-Aldrich (Mississauga, ON, Canada).

Crystallization/diffraction studies

All crystals were grown using the hanging-drop vapor diffusion method and a protein concentration of ~10 mg/ml. The well solution for Fab 023.102 contained 18% PEG 3350, 0.2 M lithium acetate, and 50 mM sodium citrate buffer (pH 3.0–4.0). Crystals of Fab 023.102 in its free form and complexed with L-rhamnose or methyl-α-L-rhamnopyranosyl-(1→2)-β-D-galactopyranoside were obtained by cocrystallization under identical conditions (10 mM ligand) whereas the Fab 023.102/methyl-α-L-rhamnopyranosyl-(1→2)-3-(glycer-2-yl-phosphate)-β-D-galactopyranoside crystals were obtained by soaking (1 mM ligand). The well solutions for Fab pn132p2C05 and its complex with L-rhamnose (10 mM ligand) as well as Fab KE5 and its AD-2S1 complex (1 mM ligand) contained 0.1 M sodium acetate (pH 4.5), 20% PEG 1500, 30% Jeffamine ED-2001 (pH 7.0), and 0.2 M $(NH_4)Cl$, respectively. All crystals were flash frozen in boiling nitrogen.

All diffraction data were collected at −180°C: for Fab 023.102 on a Rigaku MicroMax 007 rotating copper anode using Osmic VariMax optics and a mar345 detector, data were processed with HKL2000 (25); for the free form of Fab pn132p2C05 on a Rigaku MicroMax 007 rotating copper anode using Osmic VariMax optics and a Rigaku R-Axis IV++ detector, data were processed with Crystal Clear software (Molecular Structure); for the Fab pn132p2C05/L-rhamnose complex on the Advanced Photon Source beamline 23ID-D with λ = 0.979 Å using a MarMosaic 300 detector, data were processed with HKL2000 software (25); and for Fab KE5 at the Canadian Light Source CMCF-ID (08ID-1) beamline with λ = 1.03 Å using a Rayonix MX300HE detector, data were processed using XDS (26).

Structure determination/refinement

All crystal structures were determined by molecular replacement using the program package Phaser (27); those of Fab 023.102 and Fab KE5 with Protein Data Bank entry 3EYF and Fab pn132p2C05 with Fab 023.102 as search models, respectively. The resulting solutions were refined using CNS (28) and Phenix (29), modeled in Coot (30), and are displayed using PyMol (31). The residue numbers follow the Kabat scheme (32). Data collection, structural refinement, and Ramachandran statistics are described in detail in Table I.

Results

Interactions between Fabs 023.102 and C05 and their common 23F carbohydrate Ag

We obtained diffraction data from crystals of the free Fab 023.102 (Fig. 1) and its complexes with the epitope building blocks L-rhamnose, the disaccharide RG, and the core RGP (Fig. 2), which incorporates the most important antigenic features of the complete 23F Ag (16). We also collected data from Fab C05 crystals with and without L-rhamnose. Data collection and refinement statistics are presented in Table I. In both the Fab 023.102/RGP and Fab C05/L-rhamnose complexes, a molecule of L-rhamnose sits in a superimposable position, in the same tight crevice formed by the

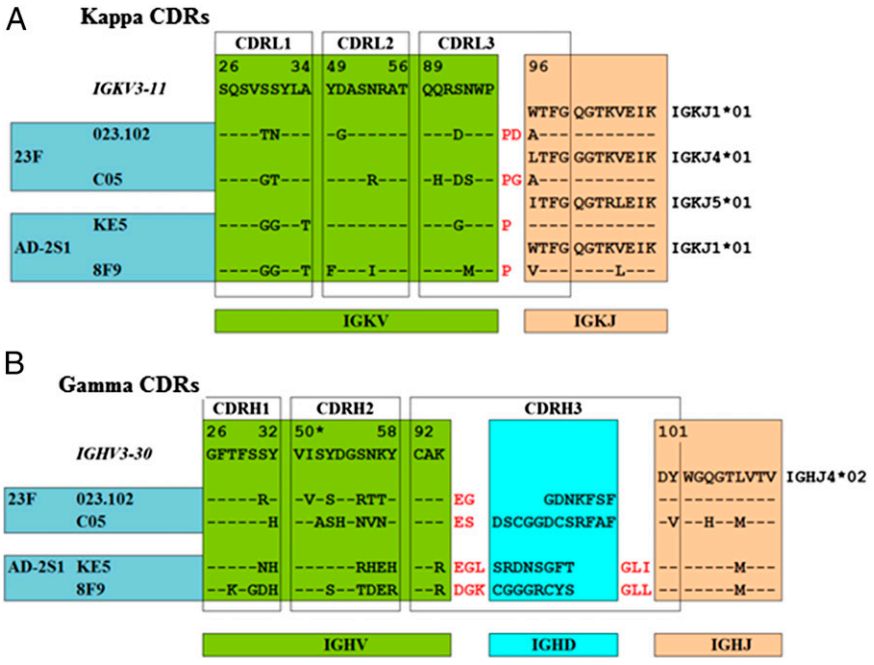

FIGURE 1. Sequences of the CDRs of anti-HCMV and anti–*S. pneumoniae* Abs. (**A**) Sequences of the κ L chains. (**B**) Sequences of the γ H chains. Residues listed in red letters result from the addition of untemplated bases by the enzyme TdT. *, residue Ser^{52}A(H) in CDRH2. Gene names and numbers are available at: http://www.imgt.org.

hypervariable loops CDRL3, CDRH2, and CDRH3 (Fig. 3A, 3B; Supplemental Figs. 1, 2). A network of hydrogen bonds to CDRL3 and CDRH3, together with van der Waals interactions with CDRH2, defines the specificity of L-rhamnose recognition by Fabs 023.102 and C05. A primary contact between the paratopes and L-rhamnose is formed by the germline-encoded CDRL3 residue Arg^{91}(L), interacting with the L-rhamnose C2′ and C3′ hydroxyl groups (bond distances, 3.0 and 2.8 Å, respectively). The L-rhamnose C6 methyl group in both complexes forms van der Waals contacts with the side chains of germline-encoded CDRH2 residue, Val^{50}(H). In Fab 023.102, germline-encoded residue Ser^{52}(H) is also part of this pocket, but in Fab C05 this residue is somatically mutated to alanine. The axial C2′ hydroxyl of L-rhamnose approaches the backbone carbonyls of germline-encoded Asn^{93}(L) (2.6 Å) whereas the ether oxygen O5 of L-rhamnose nudges up against the germline-encoded Trp^{94}(L) and stochastically encoded Pro^{95}A(L), key residues of CDRL3 that form an inter–side chain stacking arrangement. In coordination with germline-encoded residue Arg^{91}(L), CDRH3 junctional residue Glu^{95}(H) forms strong hydrogen bonds between the equatorial hydroxyls at C3′ and C4′ (2.6 and 2.8 Å, respectively). The pocket is completed by somatic mutations present in both 23F-directed Abs, Thr^{56}(H) in 023.102 and Val^{56}(H) in C05. Finally, the hydroxyl group connected to the anomeric carbon of rhamnose is in the α-conformation and points away from the Fab surface. Thus, in both Fabs, germline-encoded residues in CDRL3 and CDRH2 contact L-rhamnose, the major determinant of antigenicity, and shape the tight crevice in concert with Pro^{95}A(L) and the stochastically encoded CDRH3. In the Fab 023.102/RGP complex, electron density corresponding to the D-galactose portion was lower, suggesting that this part of the Ag adopts a number of similar conformations. Consequently, no stable contacts to the Fab can be assigned. The glycerol phosphate group linked to the C3′ position of the D-galactose moiety, however, occupies clear, unambiguous electron density. It is positioned next to Lys^{100}(H) of the CDRH3 loop. Crystal packing forces seem to prevent it from approaching the lysine amino group more closely than the 3.5 Å observed.

In addition to the RGP complex with Fab 023.102, we also analyzed Ag-free Fab 023.102 as well as its complexes with L-rhamnose and the disaccharide RG. All of these structures show no obvious changes in protein conformation or ligand–protein interaction. A general understanding of 23F specificity has therefore been established. Our results also demonstrate a clear directionality of the 23F polymer on the surface of Fabs 023.102 and C05.

Interactions between Fab KE5 and its AD-2S1 Ag

The KE5 paratope binds the AD-2S1 peptide (ETIYNTTLKY) in an extended conformation (Fig. 3C, Supplemental Fig. 3). The peptide backbone of the Ag encircles the CDRH3, forming a network of hydrogen bonds with CDRH3 backbone atoms. Additionally, the AD-2S1 peptide's side chains contact both L and H chain CDRs of Fab KE5. Fig. 3C illustrates the many germline-encoded residues that make important contacts with the AD-2S1 Ag peptide. These residues include Tyr^{32}(L) from CDRL1, Tyr^{49}(L) and Asp^{50}(L) from CDRL2, Arg^{91}(L), Trp^{94}(L), and Pro^{95}(L) from CDRL3, as well as Ser^{52}(H) and Tyr^{52}A(H) from CDRH2.

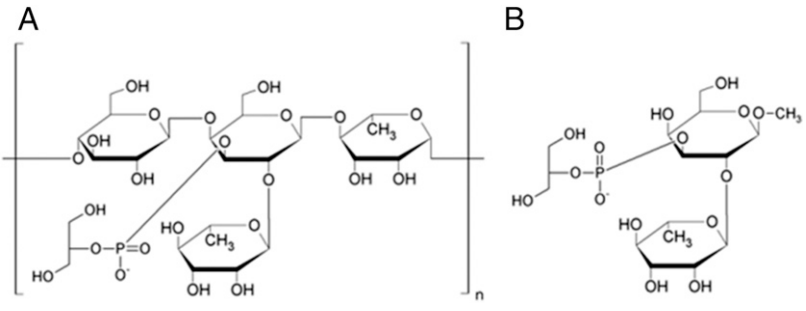

FIGURE 2. Structure of the polysaccaride of *S. pneumoniae* serotype 23F. (**A**) Schematic structure of the repeating unit of the pneumococcal polysaccharide of *S. pneumoniae* serotype 23F (16). (**B**) Schematic structure of the methylated core disaccharide RGP (18). These figures were created using ChemSketch (ACD Labs).

Table I. X-ray diffraction data collection and refinement statistics

	Fab 023.102	Fab 023.102 RAM	Fab 023.102 RG	Fab 023.102 RGP	Fab C05	Fab C05 RAM	Fab KE5	Fab KE5 AD-2S1
Data collection								
Space group	$P2_12_12_1$	$P2_1$	C2	C2	C2	C2	$P2_1$	$P2_1$
Cell dimensions								
a, *b*, *c* (Å)	66.0, 72.6, 94.2	73.8, 66.7, 109.3	144.4, 66.9, 119.1	143.9, 66.5, 118.2	132.1, 48.8, 80.0	132.2, 48.7,79.3	73.9, 79.5, 86.3	43.3, 65.5, 67.1
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 101.3, 90.0	90.0, 111.9, 90.0	90.0, 111.6, 90.0	90.0, 120.8, 90.0	90.0, 120.2, 90.0	90.0, 113.3, 90.0	90.0, 96.6, 90.0
Resolution (Å)	1.90 (1.95–1.90)	2.00 (2.05–2.00)	2.30 (2.35–2.30)	2.10 (2.20–2.10)	1.80 (1.85–1.80)	1.51 (1.55–1.51)	1.70 (1.75–1.70)	1.60 (1.65–1.60)
R_{sym}[a]	0.053 (0.39)	0.053 (0.29)	0.040 (0.24)	0.049 (0.26)	0.032 (0.059)	0.032 (0.147)	0.076 (0.51)	0.048 (0.49)
$I/\sigma I$	21.4 (2.4)	17.7 (3.0)	24.9 (2.9)	16.4 (2.8)	44.7 (9.0)	29.2 (6.1)	19.0 (3.8)	19.8 (3.2)
Completeness (%)	97.3 (75.4)	98.1 (87.0)	94.9 (67.6)	90.5 (67.4)	98.6 (97.5)	99.7 (97.4)	96.8 (95.5)	99.6 (99.9)
Redundancy	3.3 (1.6)	1.8 (1.4)	1.8 (1.1)	1.7 (1.2)	1.8 (1.8)	1.9 (1.8)	7.3 (7.4)	3.7 (3.7)
No. molecules in asymmetric unit	1	2	2	2	1	1	2	1
Refinement								
Resolution (Å)	17.0–1.90	17.0–2.00	17.0–2.30	17.0–2.10	25.2–1.80	41.5–1.51	43.0–1.70	43.0–1.60
No. reflections	35,499	70,327	44,921	55,118	40,216	67,826	97,584	48,944
R_{work}/R_{free}	0.218/0.243	0.224/0.236	0.243/0.255	0.249/0.262	0.172/0.196	0.179/0.206	0.217/0.224	0.213/0.226
No. atoms								
Protein	3191	6318	6293	6311	3290	3264	6744	3431
Ligand/ion	N/A	22	44	64	18	19	N/A	5
Water	273	312	111	319	332	350	687	289
B factors								
Protein	24.7	36.2	50.3	40.4	19.4	20.3	15.3	16.7
Ligand/ion	N/A	34.2	77.9	68.5	20.9	19.4	N/A	18.1
Water	30.8	39.8	48.3	43.9	26.1	27.3	24.4	25.1
RMSD								
Bond lengths (Å)	0.0064	0.0061	0.0051	0.0077	0.0070	0.0070	0.0041	0.0059
Bond angles (°)	1.60	1.58	1.58	1.72	1.15	1.16	1.36	1.60
Ramachandran plot								
Favored regions (%)	97.1	96.4	95.6	94.7	95.9	95.6	95.4	96.4
Allowed regions (%)	2.4	3.1	3.6	4.7	3.9	3.9	3.7	3.0
Outliers (%)	0.49	0.50	0.75	0.62	0.23	0.48	0.92	0.68
PDB ID	4HIE	4HIH	4HII	4HIJ	4PTT	4PTU	4HH9	4HHA

Values in parentheses are for highest resolution shell.
[a] $R_{sym} = \Sigma_{hkl} \Sigma_i(|I_{hkl,i} - \langle I_{hkl}\rangle|)/\Sigma_{hkl} \Sigma_i I_{hkl,i}$, where $\langle I_{hkl}\rangle$ is the average of symmetry-related observations.
AD-2S1, ETIYNTTLKY; PDB ID, Protein Data Bank ID (http://www.rcsb.org); RAM, rhamnose; RMSD, root mean square deviation.

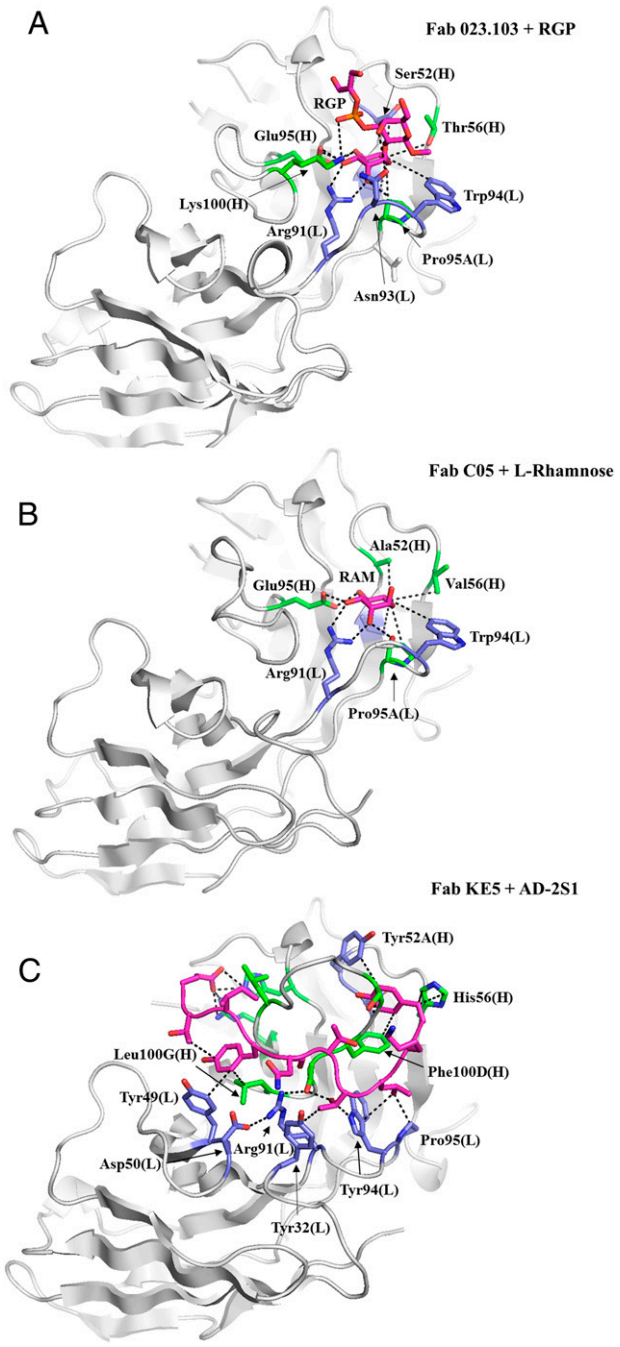

FIGURE 3. Structures of Abs 023.102, pn132p2C05, and KE5 with bound ligands. (**A**) The methylated core disaccharide RGP bound to the paratope of Ab 023.102. (**B**) L-Rhamnose bound to the paratope of Ab pn132p2C05. (**C**) AD-2S1 peptide (ETIYNTTLKY) bound to the paratope of Ab KE5. Carbon atoms of germline-encoded residues are colored blue. Carbon atoms of non–germline-encoded residues are colored green (oxygen atoms, red; nitrogen atoms, dark blue). For ligands, carbon atoms are colored magenta (oxygen atoms, red; nitrogen atoms, dark blue, phosphate atom, orange) Dashed lines indicate specific contacts. See Supplemental Figs. 1–3 for details.

Specificity for the AD-2S1 peptide is dictated primarily by these germline-encoded residues. The Tyr side chain of peptide residue P4 sits in a hydrophobic pocket formed by germline-encoded CDRL2 residue Tyr^{49}(L), the CDRH3 backbone, and junctional residue Leu^{100}G(H), whereas the backbone oxygen of P5-Asn contacts the extended side chain of germline-encoded amino acid Arg^{91}(L). A chloride ion (not shown) is centered between the P5-Asn backbone nitrogen and germline-encoded residues Asp^{50}(L) and Arg^{91}(L).

The methyl group of the P7-Thr side chain stacks against the aromatic ring of the germline-encoded residue Tyr^{32}(L) whereas its hydroxyl group interacts with the ring nitrogen of germline-encoded Trp^{94}(L) and the backbone oxygen of Thr^{100}E(H) from CDRH3. Trp^{94}(L), together with germline-encoded Pro^{95}(L), creates a hydrophobic groove that holds the side chain of P8-Leu. Finally, the P10-Tyr side chain inserts into a deep pocket formed by the CDRH2 and CDRH3 loops. At the base of the pocket sits CDRH3 residue Phe^{100}D(H) engaging in an edge-on aromatic–aromatic interaction with the P10-Tyr ring. One wall of the pocket is comprised of backbone atoms of the CDRH3; the opposite wall is formed by the side chains of germline-encoded residues Ser^{52}(H) and Tyr^{52}A(H), which engage in van der Waals interactions with the Ag peptide, and by somatically mutated residue His^{56}(H), whose side chain lies parallel to the P10-Tyr ring centered on the β-carbon.

In KE5, the junctional residue Glu^{95}(H) is critical for CDRH3 stabilization, forming hydrogen bonds with the backbone nitrogen atoms of Thr^{100}E(H) and Gly^{100}F(H), which are part of an opposing junctional Ser/Thr-Gly-Leu-Leu/Ile motif, encoded by stochastically generated DNA, conserved in the anti–AD2-S1 Abs, and located at the C-terminal base of the CDRH3 (Fig. 1) (15, 24). Its residues not only stabilize the conformation of the CDRH3 loop but also form part of the hydrophobic pockets that interact with the AD-2S1 peptide side chains of P4-Tyr and P8-Leu (24).

Discussion

The crystal structures of the complexes of Fab 023.102/RGP and C05/L-rhamnose reveal the basis for the high specificity of these Abs for L-rhamnose, the key antigenic component of the PPS of *S. pneumoniae* serotype 23F (33–35). Binding interactions are mostly achieved through contacts with the germline-encoded amino acids Arg^{91}(L), Trp^{94}(L), and Val^{50}(H), and also Ser^{52}(H) Asn^{93}(L) in Fab 023.102, with further stochastically generated amino acids also contributing (33). Residues of CDRL3, CDRH2, and CDRH3 encircle a deep crevice, tailoring interactions closely to this specific carbohydrate molecule. Germline-encoded residue Arg^{91}(L) in CDRL3 and stochastically encoded Glu^{95}(H) in CDRH3 are placed so that their side chains complement the axial and equatorial L-rhamnose hydroxyls. The CDRH2 loop forms a small pocket to accommodate the C6 methyl group of L-rhamnose engaging the sugar in several van der Waals interactions, thereby elegantly promoting its binding over that of L-mannose (33). *IGHV3-30* germline-encoded residues in CDRH2 are prominent in forming the pocket for L-rhamnose, explaining this sugar's much weaker interactions with anti-23F Abs that use *IGKV3-11* paired with other *IGHV* genes (33). The paratope complementarity for L-rhamnose in Fabs 023.102 and C05 is dictated almost entirely by germline- and stochastically encoded residues already present in their primary Ab ancestors. The only exception is germline-encoded residue Asn at position 56 of the CDRH2 loop (Fig. 1), which is mutated to threonine in Fab 023.102 and to a comparably shaped valine in Fab C05. Presumably, the methyl groups of these shorter side chains are better suited to accommodate the L-rhamnose C6 methyl group. The same change has been found in other anti-23F Abs, isolated from different donors, which also use *IGHV3-30* and *IGKV3-11* (10). The D-galactose moiety of the 23F polysaccharide sits outside this pocket, with its exact position not well defined in crystals of the complex. The axial 4′-hydroxyl of D-galactose is near the side chain of CDRL3's germline-encoded Asn^{93}(L) (3.2 Å), an interaction that may well be strengthened upon binding of the full-length Ag. The (3-*O*)-phosphate-2-glycerol portion of RGP, however, does provide significant additional information. The electron density for the phosphate group, another major binding determinant of the 23F polymer, is unambiguous and positions the ion directly above CDRH3

residue Lys^{100}(H). Interestingly, despite using a different, much longer *IGHD* segment that includes a disulfide bond, Fab C05 has an arginine residue in a superimposable position to interact with the 23F phosphate. The residues of the C-terminal portion of the CDRH3 loop in Fab 023.102 are Lys-Phe-Ser-Phe, very similar to the equivalent residues in Fab C05, Arg-Phe-Ala-Phe, suggesting this motif is important in 23F binding by positioning a positively charged residue to form an electrostatic interaction with the phosphate moiety. Underlining the importance of these residues in ligand specificity, the related protective human anti-23F mAb, PVAX2-p8 E03 using *IGHV3-30* and *IGKV3-11* (17) with the CDRH3 sequence TKEGAPPGKYAFDI, contains a junctional glutamic acid in the same position as Glu^{95}(H) and a Lys-Tyr-Ala-Phe motif, with the lysine residue in a similar position as Lys^{100}(H). These results suggest a model for the binding of the full 23F polymer consistent with ideas put forward by Reason and Zhou (33). Fab 023.102 interacts with one of the repeating four-sugar building blocks of the 23F polymer with distinct directionality, specifically selecting the antennary L-rhamnose moiety over the internal backbone L-rhamnose. The crystal structures predict that the missing backbone carbohydrate, D-glucose, will most likely be placed above the folded-in CDRH3 loop, and the internal backbone L-rhamnose will be positioned in contact with the remaining part of CDRL3, specifically the flat surface of germline-encoded Trp^{94}(L).

KE5 is an anti–AD-2S1 Ab that also utilizes *IGHV3-30*, *IGHJ4*, and *IGKV3-11*. However, KE5 employs a different *IGHD* gene, *IGHD3-22*01*, whereas 8F9 uses *IGHD2-5*01* (Fig. 1) (15). In the Ag-free Fabs of both KE5 and 8F9, the extended conformations of the CDRH3 loops are stabilized by tight intra-CDRH3 interactions. In the KE5 paratope, a hydrogen bond at the loop apex between the side chains of Asp^{100}(H) and Ser^{100}B(H) and in its 8F9 equivalent, an intra-CDRH3 disulfide bond (24), achieve this goal. In the presence of the Ag peptide, however, the CDRH3 conformations are slightly different from each other and from the one adopted in Fab 8F9, indicating an intrinsic mobility of CDRH3 (24). The presence of such stabilizing elements will reduce flexibility in extended CDR loops, thus minimizing the entropic penalty when forming a rigid complex with the Ag. Not only do 8F9 and KE5 CDR3H adopt very similar conformations, but their backbone atoms also display the same interactions with the AD-2S1 peptide. Despite presenting different sequences in their CDRH3s, the sites of most interactions with the Ag and the same structural features are seen in both AD-2S1–binding Abs. With the exception of the N- and C-terminal amino acids, the AD-2S1 peptide conformations in both the Fab KE5 and Fab 8F9 complexes are virtually identical.

Germline-encoded residues of Fabs 023.102 and C05 that interact with the core disaccharide, especially the L-rhamnose part, are maintained and somatic mutations from affinity maturation again support germline-encoded contacts. In the Fab KE5/AD-2S1 complex, as had been seen in the 8F9/AD-2S1 structure (24), many residues encoded by the germline V genes make key contacts with Ag peptide side chains, and affinity maturation serves to support these germline-encoded contacts. A closer inspection of the residues encoded by the L chain V gene *IGKV3-11* reveals two key basic and aromatic germline-encoded residues, Arg^{91}(L) and Trp^{94}(L), which are conserved in all four Fab molecules and make important Ag contacts (10, 15, 23, 24, 33, 34). The side chain of Arg^{91}(L) sits at the base of CDRL3 and CDRH3, forming hydrogen bonds with both the L-rhamnose and AD2-S1 Ags. In both structures, its position is stabilized by interactions with residues contributed by the CDR loops. In Fab 023.102 and Fab C05, the Arg^{91}(L) guanidinium group forms a bridge between the backbone oxygens of Pro^{95}A(L) of CDRL3 and Lys^{100}(H) of CDRH3. A water molecule, deep in the binding pocket, links Lys^{100}(H)'s backbone carbonyl with the Nε of Arg^{91}(L). A similar situation is found in the anti–AD-2S1 Fabs KE5 and 8F9 (24). Again, a buried water molecule stabilizes the position of the guanidinium group of germline-encoded Arg^{91}(H). In this case, the water is bound to the hydroxyl of CDRL1 residue Thr^{34}(L) and the backbone nitrogen of germline-encoded Asp^{50}(L) of CDRL2. The side chain of germline-encoded Asp^{50}(L) also binds the germline-encoded Arg^{91}(L) guanidinium group, bridging to the backbone oxygen of Thr^{100}E(H) at the base of CDRH3.

Germline-encoded Trp^{94}(L) makes direct contacts with both peptide and sugar Ags. In Fab KE5, as well as Fab 8F9 (24), it forms a surface that interacts with the side chain of P8-Leu of the AD-2S1 peptide, whereas in Fabs 023.102 and C05 it contributes to one side of the L-rhamnose–binding crevice. The conformations of these two residues, however, as well as those of the entire CDRL3, are very different in the two Fabs. The insertion of a single junctional residue, Asp^{95}B(L) in Fab 023.102 or Gly^{95}B(L) in Fab C05, between the stretches coded for by *IGKV3-11* and *IGKJ*, directly adjacent to the Trp-Pro-Pro motif at the apex of the CDRL3 loop, causes an entirely different backbone conformation, leading to an almost 180° rotation of the indole ring of germline-encoded Trp^{94}(L) between the two Abs. The Asp^{95}B(L) side chain of Fab 023.102 does not contact the Ag; instead, it forms a hydrogen bond to its own backbone amide nitrogen, which should further stabilize the CDRL3 conformation.

Fig. 4 illustrates the superposition of α-carbon traces of the variable domains of Fab 023.102, Fab C05, Fab KE5, and Fab 8F9. The four proteins, despite a few somatic mutations in their V genes, are structurally very similar, except for the conformations of the CDRL3 and CDRH3 loops. The pair of preferred V genes, *IGHV3-30* and *IGKV3-11*, of which *IGHV3-30* encodes CDRH1 and CDRH2 and *IGKV3-11* encodes CDRL1, CDRL2, and part of CDRL3, with corresponding *IGHD*, *IGHJ*, and *IGKJ* genes, binds to the AD-2S1 Ag of the gB glycoprotein of the HCMV and to the 23F serotype of *S. pneumoniae* of PPS with different CDRH3 and CDRL3. Analysis of the full stochastically synthesized portions of the CDRH3 and 3′ part of CDRL3 (34) revealed that the CDR3s dictate specificity. mAbs 8F9 and KE5 use very similar paratopes in binding to the gB protein (15, 23) whereas the paratopes of

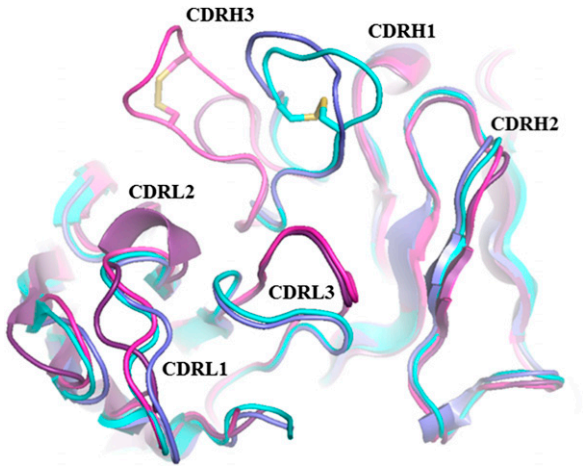

FIGURE 4. Superimposition of the backbone tracings of anti-HCMV and anti–*S. pneumoniae* Abs. α-Carbon backbone tracings of the variable regions of anti-23F Fabs 023.102 (purple), C05 (magenta), anti-HCMV Fabs KE5 (dark blue), and 8F9 (cyan). Also displayed are the disulfide bonds (yellow) of the CDRH3 loops of Fab C05 and Fab 8F9. Major conformational differences are restricted to the CDR3 loops. Note that the conformation of the CDRL1 of Fab 023.102, probably due to disorder, deviates from the consensus conformation.

023.102 and C05, which bind to the 23F serotypes of PPS, are also very similar but quite different from those of mAbs 8F9 and KE5 (Fig. 4). These findings were consistent with results in mice suggesting that it is the role of the CDR3 to provide Ag specificity (36, 37).

For two species to coevolve, there has to be adaptive genetic change in both species as a result of the interaction. In a human–pathogen interaction, alleles of the pathogen may be selected for their ability to increase infectivity, whereas those of the human would be selected to resist infection (38, 39). For example, vertebrate MHC diversity is thought to have evolved to impede pathogens such as HCV by targeting conserved peptides important for pathogen infectivity (40, 41). These changes do not always come without disadvantages, however, with MHC restriction being correlated with autoimmune disease (42). Similarly, the human Ab response must balance the ability to respond to multiple pathogens with the potential for unwanted cross-reactivity and subsequent deletion.

Our findings are consistent with the conclusion that preferred human V genes are under evolutionary selective pressure by multiple evolving pathogens. The germinal center can generate high-affinity Abs but without preference for the protective Abs directed against vulnerable, invariable Ags. Only evolutionary pressure can select the preferred V genes. Indeed, Ab responses to bacterial carbohydrates are T cell–independent and as such generally undergo limited mutation, thus relying on avidity effects for binding to the multiarrayed carbohydrate on the bacterial surface and the underlying binding ability of the germline Ab. Preferred V gene usage to carbohydrate Ags has not only been observed in the human Ab response to *S. pneumoniae* but also to *Haemophilus influenzae* type b where analysis of human Abs targeting HibPS found the repeated pairing of the *IGKV2D-29* (*VK A2*) L chain with the *IGHV3-23* H chain and a canonical CDRH3 (43, 44). Although *IGHV3-23* is prevalently used in human Abs (6–10%), the dominant use of *IGKV2D-29* in response to *H. influenzae* is underscored by its limited use otherwise (0.2–0.5%) (45). Through expression of germline-based Fabs devoid of mutations but combined with the canonical CDR3s, significant binding to HibCP oligosaccharides is observed (46). Similarly, expression of Fabs representing nonmutated Abs from 4-mo-old infants similarly demonstrated the role of these V genes in binding and the requirements of a short CDRH3 with the sequence Gly-Tyr-Gly-Phe/Met/Asp (44). In mouse models of bacterial infection, structural analysis likewise showed conserved germline gene usage in response to carbohydrate Ags (47).

The *IGHV3-30* and *IGKV3-11* V genes are well represented in humans. High-throughput Ab sequencing of naive versus IgM memory versus IgG memory repertoires revealed that *IGHV3-30* and *IGHV3-23* are the most commonly used heavy variable genes, both of which are dominant in the *S. pneumoniae* 23F human Ab response (10, 48). Interestingly, *IGHJ4* usage increased in memory subsets, and those containing *IGHJ4* paired with *IGHV3-30* found increased usage of *IGHD3* family members, the same combination used in KE5.

Adaptable protection against carbohydrate Ags using small numbers of germline genes (49) and an example of cross-reactivity between carbohydrate and single-strand DNA with limited paratope overlap (50) have been reported. We now add data that support a mechanism whereby the *IGHV3-30* and *IGKV3-11* V genes provide an innate foundation for protection to two divergent pathogens by their ability to generate neutralizing mAbs targeting an invariant site on a viral protein necessary for viral infectivity as well as neutralizing mAbs directed at a pneumococcal capsular polysaccharide. Variations in CDR3s then achieve high specificity to particular Ags. Their ability to do so against two pathogens with which humans have coevolved underscores the importance and usage of *IGHV3-30* in both naive and memory B cell subsets. Our results provide an example that illustrates how a limited set of human germline genes can contribute to the creation of binding sites for a variety of highly diverse Ags.

Acknowledgments

We thank D.C. Reason (Children's Hospital Oakland Research Institute) for providing the pARC plasmid and for advice regarding expression of Fab 023.102, and M. Nitz (University of Toronto) for contributions to the synthesis of the carbohydrate analogs. We thank A. Cunningham (University of Toronto) for help with protein crystallization. We are grateful to A. Dong (Structural Genomics Consortium, Toronto, ON, Canada), R. Pfoh (Princess Margaret Cancer Centre), B.T. Eger (University of Toronto), and the staff of the Canadian Macromolecular Crystallography Facility at the Canadian Light Source for help with data collection. We are also grateful to the staff at the X-ray Operations and Research beamline 23ID-D of the Advanced Photon Source, Argonne National Laboratories (Argonne, IL).

Disclosures

K.S. consults with Pamlico Biopharma, Inc. in the area of Ab therapeutics for pneumococcal disease. The other authors have no financial conflicts of interest.

References

1. Desiderio, S. V., G. D. Yancopoulos, M. Paskind, E. Thomas, M. A. Boss, N. Landau, F. W. Alt, and D. Baltimore. 1984. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature* 311: 752–755.
2. Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa. 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33: 717–728.
3. Thai, T. H., M. M. Purugganan, D. B. Roth, and J. F. Kearney. 2002. Distinct and opposite diversifying activities of terminal transferase splice variants. *Nat. Immunol.* 3: 457–462.
4. Alt, F. W., and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions. *Proc. Natl. Acad. Sci. USA* 79: 4118–4122.
5. Boyd, S. D., B. A. Gaëta, K. J. Jackson, A. Z. Fire, E. L. Marshall, J. D. Merker, J. M. Maniar, L. N. Zhang, B. Sahaf, C. D. Jones, et al. 2010. Individual variation in the germline Ig gene repertoire inferred from variable region gene rearrangements. *J. Immunol.* 184: 6986–6992.
6. Watson, C. T., and F. Breden. 2012. The immunoglobulin heavy chain locus: genetic variation, missing data, and implications for human disease. *Genes Immun.* 13: 363–373.
7. Scott, M. G., and M. H. Nahm. 1992. Characterization of the human IgG antibody V_L repertoire to *Haemophilus influenzae* type b polysaccharide. *J. Infect. Dis.* 165(Suppl. 1): S53–S56.
8. Scott, M. G., H. G. Zachau, and M. H. Nahm. 1992. The human antibody V region repertoire to the type B capsular polysaccharide of *Haemophilus influenzae*. *Int. Rev. Immunol.* 9: 45–55.
9. Silverman, G. J., and A. H. Lucas. 1991. Variable region diversity in human circulating antibodies specific for the capsular polysaccharide of *Haemophilus influenzae* type b. Preferential usage of two types of VH3 heavy chains. *J. Clin. Invest.* 88: 911–920.
10. Zhou, J., K. R. Lottenbach, S. J. Barenkamp, A. H. Lucas, and D. C. Reason. 2002. Recurrent variable region gene usage and somatic mutation in the human antibody response to the capsular polysaccharide of *Streptococcus pneumoniae* type 23F. *Infect. Immun.* 70: 4083–4091.
11. Weitkamp, J. H., B. J. Lafleur, H. B. Greenberg, and J. E. Crowe, Jr. 2005. Natural evolution of a human virus-specific antibody gene repertoire by somatic hypermutation requires both hotspot-directed and randomly-directed processes. *Hum. Immunol.* 66: 666–676.
12. Chan, C. H., K. G. Hadlock, S. K. Foung, and S. Levy. 2001. V_H1-*69* gene is preferentially used by hepatitis C virus-associated B cell lymphomas and by normal B cells responding to the E2 viral antigen. *Blood* 97: 1023–1026.
13. Davies, D. R., and G. H. Cohen. 1996. Interactions of protein antigens with antibodies. *Proc. Natl. Acad. Sci. USA* 93: 7–12.
14. Rajewsky, K., I. Förster, and A. Cumano. 1987. Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science* 238: 1088–1094.
15. McLean, G. R., O. A. Olsen, I. N. Watt, P. Rathanaswami, K. B. Leslie, J. S. Babcook, and J. W. Schrader. 2005. Recognition of human cytomegalovirus by human primary immunoglobulins identifies an innate foundation to an adaptive immune response. *J. Immunol.* 174: 4768–4778.
16. Richards, J. C., and M. B. Perry. 1988. Structure of the specific capsular polysaccharide of *Streptococcus pneumoniae* type 23F (American type 23). *Biochem. Cell Biol.* 66: 758–771.

17. Smith, K., J. J. Muther, A. L. Duke, E. McKee, N. Y. Zheng, P. C. Wilson, and J. A. James. 2013. Fully human monoclonal antibodies from antibody secreting cells after vaccination with Pneumovax 23 are serotype specific and facilitate opsonophagocytosis. *Immunobiology* 218: 745–754.
18. Dasgupta, S., and M. Nitz. 2010. Synthesis of a core disaccharide from the *Streptococcus pneumoniae* type 23F capsular polysaccharide antigen. *Carbohydr. Res.* 345: 2282–2286.
19. de Wildt, R. M., W. J. van Venrooij, G. Winter, R. M. Hoet, and I. M. Tomlinson. 1999. Somatic insertions and deletions shape the human antibody repertoire. *J. Mol. Biol.* 294: 701–710.
20. McGeoch, D. J., A. Dolan, and A. C. Ralph. 2000. Toward a comprehensive phylogeny for mammalian and avian herpesviruses. *J. Virol.* 74: 10401–10406.
21. Ho, M. 1982. *Cytomegalovirus: Biology and Infection.* Plenum, New York. doi: 10.1007/978-1-4684-4073-7
22. Nigro, G., S. P. Adler, R. La Torre, and A. M. Best, Congenital Cytomegalovirus Collaborating Group. 2005. Passive immunization during pregnancy for congenital cytomegalovirus infection. *N. Engl. J. Med.* 353: 1350–1362.
23. McLean, G. R., C. W. Cho, and J. W. Schrader. 2006. Autoreactivity of primary human immunoglobulins ancestral to hypermutated human antibodies that neutralize HCMV. *Mol. Immunol.* 43: 2012–2022.
24. Thomson, C. A., S. Bryson, G. R. McLean, A. L. Creagh, E. F. Pai, and J. W. Schrader. 2008. Germline V-genes sculpt the binding site of a family of antibodies neutralizing human cytomegalovirus. *EMBO J.* 27: 2592–2602.
25. Otwinowski, Z., and W. Minor. 1997. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276: 307–326.
26. Kabsch, W. 2010. XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66: 125–132.
27. McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. 2007. Phaser crystallographic software. *J. Appl. Cryst.* 40: 658–674.
28. Brünger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, et al. 1998. *Crystallography & NMR System*: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 54: 905–921.
29. Afonine, P. V., R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, and P. D. Adams. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 68: 352–367.
30. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60: 2126–2132.
31. DeLano, W. L. 2010. *The PyMOL Molecular Graphics System, Version 1.4*. Schrödinger, LLC, New York. Available at: http://www.pymol.org/pymol.
32. Abhinandan, K. R., and A. C. R. Martin. 2008. Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol. Immunol.* 45: 3832–3839.
33. Reason, D. C., and J. Zhou. 2004. Correlation of antigenic epitope and antibody gene usage in the human immune response to *Streptococcus pneumoniae* type 23F capsular polysaccharide. *Clin. Immunol.* 111: 132–136.
34. Thomson, C. A., K. Q. Little, D. C. Reason, and J. W. Schrader. 2011. Somatic diversity in CDR3 loops allows single V-genes to encode innate immunological memories for multiple pathogens. *J. Immunol.* 186: 2291–2298.
35. Alonso de Velasco, E., A. F. Verheul, A. M. van Steijn, H. A. Dekker, R. G. Feldman, I. M. Fernández, J. P. Kamerling, J. F. Vliegenthart, J. Verhoef, and H. Snippe. 1994. Epitope specificity of rabbit immunoglobulin G (IgG) elicited by pneumococcal type 23F synthetic oligosaccharide- and native polysaccharide-protein conjugate vaccines: comparison with human anti-polysaccharide 23F IgG. *Infect. Immun.* 62: 799–808.
36. Xu, J. L., and M. M. Davis. 2000. Diversity in the CDR3 region of V_H is sufficient for most antibody specificities. *Immunity* 13: 37–45.
37. Nguyen, H. P., N. O. Seto, C. R. MacKenzie, L. Brade, P. Kosma, H. Brade, and S. V. Evans. 2003. Germline antibody recognition of distinct carbohydrate epitopes. *Nat. Struct. Biol.* 10: 1019–1025.
38. Lederberg, J. 1999. J. B. S. Haldane (1949) on infectious disease and evolution. *Genetics* 153: 1–3.
39. Woolhouse, M. E., L. H. Taylor, and D. T. Haydon. 2001. Population biology of multihost pathogens. *Science* 292: 1109–1112.
40. Erickson, A. L., Y. Kimura, S. Igarashi, J. Eichelberger, M. Houghton, J. Sidney, D. McKinney, A. Sette, A. L. Hughes, and C. M. Walker. 2001. The outcome of hepatitis C virus infection is predicted by escape mutations in epitopes targeted by cytotoxic T lymphocytes. *Immunity* 15: 883–895.
41. Kubinak, J. L., J. S. Ruff, C. W. Hyzer, P. R. Slev, and W. K. Potts. 2012. Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proc. Natl. Acad. Sci. USA* 109: 3422–3427.
42. Mathieu, A., F. Paladini, A. Vacca, A. Cauli, M. T. Fiorillo, and R. Sorrentino. 2009. The interplay between the geographic distribution of HLA-B27 alleles and their role in infectious and autoimmune diseases: a unifying hypothesis. *Autoimmun. Rev.* 8: 420–425.
43. Pinchuk, G. V., C. Nottenburg, and E. C. Milner. 1995. Predominant V-region gene configurations in the human antibody response to *Haemophilus influenzae* capsule polysaccharide. *Scand. J. Immunol.* 41: 324–330.
44. Lucas, A. H., G. R. McLean, D. C. Reason, A. P. O'Connor, M. C. Felton, and K. D. Moulton. 2003. Molecular ontogeny of the human antibody repertoire to the *Haemophilus influenza*e type B polysaccharide: expression of canonical variable regions and their variants in vaccinated infants. *Clin. Immunol.* 108: 119–127.
45. Glanville, J., W. Zhai, J. Berka, D. Telman, G. Huerta, G. R. Mehta, I. Ni, L. Mei, P. D. Sundar, G. M. Day, et al. 2009. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc. Natl. Acad. Sci. USA* 106: 20216–20221.
46. Hougs, L., L. Juul, A. Svejgaard, and T. Barington. 1999. Structural requirements of the major protective antibody to *Haemophilus influenzae* type b. *Infect. Immun.* 67: 2503–2514.
47. Brooks, C. L., S. Müller-Loennies, S. N. Borisova, L. Brade, P. Kosma, T. Hirama, C. R. Mackenzie, H. Brade, and S. V. Evans. 2010. Antibodies raised against chlamydial lipopolysaccharide antigens reveal convergence in germline gene usage and differential epitope recognition. *Biochemistry* 49: 570–581.
48. Briney, B. S., J. R. Willis, B. A. McKinney, and J. E. Crowe, Jr. 2012. High-throughput antibody sequencing reveals genetic evidence of global regulation of the naïve and memory repertoires that extends across individuals. *Genes Immun.* 13: 469–473.
49. Haji-Ghassemi, O., R. J. Blackler, N. Martin Young, and S. V. Evans. 2015. Antibody recognition of carbohydrate epitopes. *Glycobiology* 25: 920–952.
50. Haji-Ghassemi, O., S. Müller-Loennies, T. Rodriguez, L. Brade, P. Kosma, H. Brade, and S. V. Evans. 2015. Structural basis for antibody recognition of lipid A: insights to polyspecificity toward single-stranded DNA. *J. Biol. Chem.* 290: 19629–19640.