Recognition of Human Cytomegalovirus by Human Primary Immunoglobulins Identifies an Innate Foundation to an Adaptive Immune Response

Gary R. McLean, Ole A. Olsen, Ian N. Watt, P. Rathanaswami, Kevin B. Leslie, John S. Babcook and John W. Schrader

*J Immunol* 2005; 174:4768-4778; doi: 10.4049/jimmunol.174.8.4768
http://www.jimmunol.org/content/174/8/4768

**References** This article cites 44 articles, 14 of which you can access for free at: 
http://www.jimmunol.org/content/174/8/4768.full#ref-list-1

**Subscription** Information about subscribing to *The Journal of Immunology* is online at: 
http://jimmunol.org/subscription

**Permissions** Submit copyright permission requests at: 
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at: 
http://jimmunol.org/alerts
Recognition of Human Cytomegalovirus by Human Primary Immunoglobulins Identifies an Innate Foundation to an Adaptive Immune Response

Gary R. McLean, Ole A. Olsen, Ian N. Watt, P. Rathanaswami, Kevin B. Leslie, John S. Babcock, and John W. Schrader

Most primates, including humans, are chronically infected with cospecifically evolved, potentially pathogenic CMV. Abs that bind a 10-aa linear epitope (antigenic determinant 2 site 1) within the extracellular domain of human CMV glycoprotein B neutralize viral infectivity. In this study, we show that genes generated by recombinations involving two well-conserved human germline V elements (IGHV3-30 and IGKV3-11), and IGHJ4, encode primary Ig molecules that bind glycoprotein B at this key epitope. These particular VH, JH, and VH genes enable humans to generate through recombination and N nucleotide addition, a useful frequency of primary Igs that efficiently target this critical site on human CMV and thus confer an innate foundation for a specific adaptive response to this pathogen. The Journal of Immunology, 2005, 174: 4768–4778.

Unlike other proteins, including the pattern-receptors of the innate immune system, Igs are not directly encoded by the genome. Instead, the genes that encode the Ig H and L chains are generated somatically in differentiating B lymphocytes through the combinatorial assembly of a series of inherited genetic elements. The combinational assembly of these genetic elements is one factor that contributes to the diversity of the primary repertoire of Igs (1). L chain genes are generated by recombination of one κ or λ variable (V) gene with one joining (J) segment and a constant (C) region gene encoding either κ or λ L chains. In the case of H chain genes, recombination involves an additional element, a “diversity” (D) gene, and recombination generates VDJ region elements that combine with one of the eight C region genes. Two specialized enzymes, Rag-1 and Rag-2 are responsible for the recombination process, which also involves a series of DNA repair enzymes (2). Stochastic features of the recombination process introduce further diversity into the repertoire of primary Igs, first, through variation in the precise sites at which genetic elements are joined, and, second, through the action of an enzyme, TdT, that randomly removes and inserts nucleotides (N nucleotides) at the junctions of rearrangements (3, 4). Diversity in the primary Ig repertoire is also generated by the combinatorial expression of randomly generated H chain and L chain genes in single B lymphocytes.

These considerations raise the question of how the germline elements (V, D, and J) in the human genome have been selected in evolution. There is evidence suggesting that human Ig loci are under selective pressure from pathogens. Thus, in human populations, many V genes display polymorphisms and appear to be evolving rapidly (5). Conversely, subsets of V gene elements are relatively invariant and occur in all humans. Their conservation would be consistent with their contribution to the generation of combining sites that were either specific for common pathogens or were flexible and poly specific (6). Furthermore, the fact that differences in the flanking recombination signals result in some VH and VL genes being used more frequently than others is also consistent with the notion of selective pressures on certain V genes (7, 8). There is also direct evidence of the preferential use of certain V genes in human Ab responses against pathogens (9–11).

CMV represent an example of a pathogen that was present throughout the evolution of the human immune system, as they are present in all primates and evolved cospecifically with their hosts (12). Some 85% of humans exhibit subclinical infections with human CMV (HCMV) by age 40 (13), and although cellular immunity is probably important for recovery from acute infections, humoral immunity is important in limiting the severity of infection and in protecting the fetus from primary infection (14). The principal target of the humoral immune response to HCMV is the envelope protein glycoprotein B (gB), which contains several dominant antigenic determinants (15). Antigenic determinant 2 site 1 (AD-2S1) is a linear epitope encoded by residues 69–78 of the N-terminal surface fragment of gB, gp116 (16). Abs targeting the AD-2S1 epitope neutralize viral infectivity, probably by blocking a critical function of gp116 in attachment and fusion of virus to the cell (17, 18).

We have found that a series of human Abs specific for AD-2S1 were all derived from primary Igs encoded by single VH, JH, and VH genetic elements, suggesting that these primary Igs might have been able to bind HCMV. To test this, we constructed and expressed synthetic H and L chain genes based on recombination of

The Biomedical Research Centre, University of British Columbia, Vancouver, British Columbia, Canada

Received for publication September 21, 2004. Accepted for publication January 27, 2005.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 This work was supported by a grant from the Canadian Institutes of Health Research.


3 Current address: Abgenix Biopharma, Burnaby, BC, Canada.

4 Current address: Medical Research Council Laboratory of Molecular Biology, Cambridge, U.K.

5 Address correspondence and reprint requests to Dr. John W. Schrader, The Biomedical Research Centre, 222 Health Sciences Mall, University of British Columbia, Vancouver, BC, Canada V6T1Z3. E-mail address: john@brc.ubc.ca.

6 Abbreviations used in this paper: HCMV, human CMV; gB, glycoprotein B; FWR, framework region; gBNT, N terminus of gB; AD-2S1, antigenic determinant 2 site 1.
these germline elements and showed that, provided certain junctional variations were present, these primary Igs bound gB and the AD-2S1 epitope. Our results suggest that these evolutionarily conserved genetic elements can reproducibly contribute to the generation of primary Igs reactive with a neutralizing epitope of the gB protein of HCMV. Because these genetic elements occur in virtually all humans and are preferentially recombined, they potentially confer an innate capability to generate primary Igs that readily bind a key component on a ubiquitous, potentially pathogenic virus.

**Materials and Methods**

**Reagents**

A peptide (SHRNETYNTTLYGDTTGTGJNTTK) containing AD-2S1 epitope of HCMV gB was generated by solid-phase synthesis by Dr. I. Clark-Lewis (Biomedical Research Centre, Vancouver, Canada). The minimal epitope is shown in bold and underlined residues are those altered from the native to aid in solubility. For coating of erythrocytes, a bovine group was added at the amino terminus. Plasmid pHM90-5, containing cDNA-encoding amino acids 28–100 of the gp116 subunit of gB was a gift from Dr. M Mach (Institute for Clinical and Molecular Virology, Erlangen, Germany) (17). To express residues 28–100 of the gp16 subunit of gB on the surface of N50 cells, the plasmid pDisplay (Invirogen Life Technologies) was used.

**Isolation of anti-HCMV B cells**

Peripheral blood was obtained with informed consent from healthy donors seropositive for HCMV. B cells were isolated and plated at 200 cells/well in 96-well microcultures as described previously (19). Wells containing anti-AD-2S1 Abs were identified by ELISA. Single human B cells secreting Abs specific for the AD-2S1-neutralizing epitope of HCMV gB were isolated using an adaptation of the selected lymphocyte Ab method (20).

**RT-PCR amplification of Ig V regions**

The reverse transcription amplification of murine and rabbit V region sequences from single plaque-forming cells has been described previously (20). Human V region sequences were similarly amplified using antisense primers specific for the human Cκ region (RT-Cκ) or human Cγ1 region (RT-Cγ1) in a semi-nested strategy (Table I for full list of primers). Degenerate primers annealing to the leader sequences in the Vβ and Vκ cDNAs were first used in combination with primers specific for either the H chain C H1 region or the Cκ region. A further round of PCR was then performed with primers s.VH-fw and as.JH for the H chain and primers s.Vκ-fw and as.Jκ for the L chain. PCR products of the predicted size were separated by 2% agarose gel electrophoresis, purified using the Quiaquick gel extraction kit (Qiagen), cloned into pCR2.1 (Invitrogen Life Technologies), and sequenced using M13 and M13R primers. Sequences were compared with those in the IMGT sequence database (http://imgt.cines.fr:8104) (21) to determine the nearest germline VH, DH, JH, Vκ, and Jκ gene segments. Accession numbers for amplified V region cDNA sequences are as follows: 2B4V, AJ606108, HH12 Vκ, AJ606109, KE5 Vκ, AJ606110, FA9 Vκ, AJ606111, SE12 Vκ, AJ606112, UF12 Vκ, AJ606113, QGI1 Vκ, AJ606114, 8F9 Vκ, AJ606115, 2B4 Vκ, AJ606116, HH12 Vκ, AJ606117, KE5 Vκ, AJ606118, FA9 Vκ, AJ606119, SE12 Vκ, AJ606120, UF12 Vκ, AJ606121, QGI1 Vκ, AJ606122, and 8F9 Vκ, AJ606123.

**Construction of primary, unmutated anti-HCMV Abs**

The germline IGHV3.30*18 gene was amplified from human placental genomic DNA (Sigma-Aldrich) using primers IGVH3.30*18.sense1 and IGVH3.30*18.antisense (see Table II for full list of primers) and cloned into pCR2.1 (Invitrogen Life Technologies). Sequential rounds of PCR utilizing partially overlapping primers (IGHV3.30*18.sense2, IGHV3.30*18.sense3, IGHV3.30*18.sense4, IGHV3.30*18.antisense) were used to introduce the IGHV3.30*18 leader sequence and also a silent mutation in codon 92 (TGT to TGC) of the IGHV3.30*18 gene, creating a BstII restriction site. The three D-J regions were synthesized in three sections which were annealed together and cloned into pCR2.1. Fragments encoding the D-J regions were excised from pCR2.1, ligated to the germline IGHV3-30 fragment, and cloned into pCR2.1. To construct the L chain germline Vκ region (gLJ11) the Vκ gene IGVK3-11*01 was amplified from human genomic DNA (Sigma-Aldrich), with the leader sequence and Jκ sequence added by sequential rounds of PCR with partially overlapping primers and the final PCR product cloned into pCR2.1. Mutation of the gLJ11 residue L96 from Trp to Val was performed by the QuikChange site-directed mutagenesis protocol (Stratagene) using Pfu polymerase. To construct the germline-based Vκ regions with the Jκ5 sequence, the gLJ11 cDNA was used as a template followed by PCR with primers IGVK3-11*01.sense4 and Jκ5.antisense. We also constructed the L8F9 Vκ region with Jκ5 using the same 3′ primer and PCR conditions, resulting in the modification of three residues (V96I, K103R, and V104L).

**Expression and purification of recombinant Abs**

The L and H chain V regions were expressed using the human Ig expression vectors pAG6422 and pAH4604 (22) or the pLC-huC expression vectors (23). Human IgG1 was purified from culture supernatants from stably expressing clones of NS0 or transiently transfected HEK293 cells using a protein A-Sepharose affinity column. Igs were quantified by ELISA (see below).

### Table I. Primers for RT-PCR amplification of anti-HCMV Vβ and Vκ regions

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Nucleotide Sequence&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RT primers</strong></td>
<td></td>
</tr>
<tr>
<td>RT-Cγ1</td>
<td>agtagagctcttgagaagcttcctctga</td>
</tr>
<tr>
<td>RT-Cκ</td>
<td>cttcacaagtctacttacttacttacttact</td>
</tr>
<tr>
<td>Sense primers</td>
<td></td>
</tr>
<tr>
<td>s.8F9/VH</td>
<td>ttgcctagagctcttaagctgtggttctg</td>
</tr>
<tr>
<td>s.8F9/Vκ</td>
<td>gttcagatctctgctctgctgctgctgct</td>
</tr>
<tr>
<td>s.Vκ leader.3</td>
<td>ggtgggaagttctctgccctgctgctgctgctg</td>
</tr>
<tr>
<td>s.Vκ leader.3</td>
<td>ggtgggaagttctctgccctgctgctgctgctg</td>
</tr>
<tr>
<td>s.Vκ fw.3</td>
<td>ggtgggaagttctctgccctgctgctgctgctg</td>
</tr>
<tr>
<td>s.Vκ fw</td>
<td>ggtgggaagttctctgccctgctgctgctgctg</td>
</tr>
<tr>
<td>Antisense primers</td>
<td></td>
</tr>
<tr>
<td>s.8F9/VH</td>
<td>taaagcttgtaagagctttctagctagctagctagctagctag</td>
</tr>
<tr>
<td>s.8F9/Vκ</td>
<td>taaagcttgtaagagctttctagctagctagctagctagctag</td>
</tr>
<tr>
<td>s.CH1.1</td>
<td>ccacccacagctagctagctagctagctagctagctagctag</td>
</tr>
<tr>
<td>s.Cγ1</td>
<td>gaggagagagagagagagagagagagagagagagagagagagag</td>
</tr>
<tr>
<td>s.Jκ</td>
<td>gcagctagctagctagctagctagctagctagctagctagctag</td>
</tr>
<tr>
<td>s.Jκ</td>
<td>gcagctagctagctagctagctagctagctagctagctagctag</td>
</tr>
</tbody>
</table>

<sup>a</sup> Sequences are written 5′ to 3′ with restriction sites for subsequent cloning underlined and initiation codons in bold. Nucleotides in brackets show degeneracies such that the primers contain a mixture of products with the various nucleotides indicated in that position.
ELISA
All Abs were from Southern Biotechnology Associates. Coating with peptides or anti-human κ (in PBS) and blocking (2% BSA in PBS) of plates was performed overnight at 4°C. Dilutions of Abs (in PBS containing 1% BSA) were added and plates were left overnight at 4°C or for 1 h at 37°C. Similar results were obtained under both conditions, suggesting that equilibrium was reached. Washing (with PBS-0.05% Tween 20) was performed overnight at 4°C. Dilutions of Abs (in PBS containing 1% BSA) were added and plates were left overnight at 4°C or for 1 h at 37°C. Similar results were obtained under both conditions, suggesting that equilibrium was reached. Washing (with PBS-0.05% Tween 20) was performed overnight at 4°C. Dilutions of Abs (in PBS containing 1% BSA) were added and plates were left overnight at 4°C or for 1 h at 37°C. Similar results were obtained under both conditions, suggesting that equilibrium was reached. Washing (with PBS-0.05% Tween 20) was performed quickly using the SkanWasher 300 (Skatron). The OD at 405 nm was read in a SpectraMAX 190 spectrophotometer (Molecular Devices).

Flow cytometry analysis
Parental NS0 cells and NS0 cells engineered to stably express the AD-2S1 epitope on the cell surface were suspended at ~10^6/ml in PBS containing 0.5% FCS and the indicated Igs for 45 min on ice. They were washed once with ice-cold PBS-0.5% BSA and then stained with anti-human κ-chain FITC conjugate (Southern Biotechnology Associates). Fluorescence was analyzed using a FACScan (BD Biosciences) and the data were analyzed using CellQuest software.

Results
Analysis of H and L chain genes in single cells producing Abs to AD-2S1: evidence for frequent use of single germline VH and V<sub>e</sub> genetic elements
Human B cells were polyclonally activated in 96-well microcultures at low cell densities (200 cells/well) so that only a low frequency of wells (≤4%) generated Abs specific for AD-2S1. Thus, the cells that secreted AD-2S1-reactive Abs in these wells were likely to be clones. ELISA of supernatants identified wells containing clones of B cells secreting anti-AD-2S1 Abs. Using blood from donor 1, who had a serum titer of anti-AD-2S1 Abs of 1/40000, we identified 111 wells containing anti-AD-2S1 IgG from a total of 2976 wells. Poisson statistics indicated that this corresponded to a frequency of AD-2S1-specific B cells of 1/5305. We used the selected lymphocyte Ab method (20) to clone cDNA encoding the variable regions of Abs specific for AD-2S1 from individual B cells from 29 of these positive wells. Cells were harvested from these cultures and individual cells that were secreting the AD-2S1-specific Abs were identified using a hemolytic plaque assay based on sheep erythrocytes coated with a peptide containing the AD-2S1 epitope. Using glass micropipettes we isolated multiple single AD-2S1-specific plaque-forming cells. We used RT-PCR to amplify cDNA encoding the rearranged V regions of matching H and L chains from the single cells from each of these wells. Analysis of the sequences encoding the VH and V<sub>e</sub> regions confirmed that the cells from these 29 clones of Ab-forming cells were derived from eight distinct clones of memory B lymphocytes (Figs. 1 and 2). Of the 29 sets of cDNAs sequences, 16 were identical to those isolated from clone HH12, 4 to those from clone SE12, 3 to those from clone 8F9, and 2 to those isolated from FA9, while the remaining sequences (from 2B4, KE5, UF12, and QG1) were each unique. From some wells we cloned cDNA encoding VH and V<sub>e</sub> regions from multiple cells. In all cases the sequences of cDNAs from different AD-2S1-specific Ab-secreting cells from the same well were identical, demonstrating that the in vitro activation of B cells did not result in accumulation of somatic mutations. Alignments of the sequences of the VH and V<sub>e</sub> regions of these eight clonotypes indicated that all were closely related. Moreover, comparison with the sequences of germline VH and V<sub>e</sub> genes indicated that in every case, the most closely related germline VH gene was IGHV3-30*18, the most closely related J<sub>H</sub> gene was IGHJ4*02, and the most closely related germline L chain gene was IGKV3-11*01. However, there were differences in the usage of DH and J<sub>V</sub> segments that divided the eight clonotypes into two discrete families that must have arisen from at least two distinct rearrangements of these germline genetic elements. This implies that the clones we studied must have arisen from at least two pro B-cells that, after initial activation, gave rise to two families of somatically mutated progeny. Members of the larger family, termed here family I, (containing the clonotypes 2B4, HH12, KE5, FA9, SE12, and UF12) were characterized by mutated H chain genes derived from primary H chain genes generated by rearrangement of the germline VH gene IGHV3.30*18 with the D<sub>H</sub> segment.
FIGURE 1. Alignments of predicted amino acid sequences of the variable regions of L chains. A, Predicted amino acid sequences of the V\textsubscript{L} regions of Abs from families I and II derived from donor 1 and from the previously described human anti-HCMV mAb ITC88 in combination with the germline-based constructs gL(J1), gL(W96V), and gL(J5), and the germline elements IGKV3-11*01, IGKJ1*01, IGKJ4*01, and IGKJ5*01. Numbering and CDR (shaded areas) are as defined by Chothia et al. (45). Germline elements are as described in the ImMunoGeneTics sequence database (21). B, Detailed alignments of LCDR3 regions. Shown are the nucleotide sequences of the 3' end of the germline IGKV3-11*01 element and the beginning of the germline IGKJ element used in each family, aligned with the nucleotide and predicted amino acid sequences of representative anti-HCMV Abs from each of families I (2B4, KE5) and II (8F9, QG1). The junctional residue Pro\textsubscript{95a} is highlighted in the shaded area. Numbering and LCDR3 definition (residues 89–99) is according to Chothia et al. (45). Germline elements are as described in the IMGT sequence database (21). *ITC88 Ab described by Ohlin et al. (26).

A

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2B4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKV3-11*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ1*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ4*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ5*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QG1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8F9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKV3-11*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ1*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ4*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ5*01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th></th>
<th>92</th>
<th>93</th>
<th>94</th>
<th>95</th>
<th>96</th>
<th>97</th>
<th>98</th>
<th>99</th>
<th>100</th>
<th>101</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2B4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKV2-33*18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ4*02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QG1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8F9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKV2-33*18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKJ4*02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 2. Alignments of predicted amino acid sequences of variable regions of H chains. A, The predicted amino acid sequences of the V\textsubscript{H} regions of Abs from families I and II from donor 1 and from the previously described human anti-HCMV mAb ITC88 along with the germline elements IGHV3-30*18 and IGHJ4*02 and the germline-based V\textsubscript{H} regions gH(M), gH(A), and gH(D). Numbering and CDR definitions (shaded areas) are as defined by Chothia et al. (45). Germline elements are as described in the IMGT sequence database (21). B, Detailed alignments of HCDR3 regions. Shown are the nucleotide sequences of the 3' end of the germline IGHV3-11*01 element and the beginning of the germline IGHJ4 element used in each family, aligned with the nucleotide and predicted amino acid sequences of representative anti-HCMV Abs from each of families I (2B4, KE5) and II (8F9, QG1). Numbering and HCDR3 definition (residues 95–102) is according to Chothia et al. (45). Germline elements are as described in the IMGT sequence database (21). *ITC88 Ab described by Ohlin et al. (26).
IGHD3-22*01 and the J4_3 segment IGHJ4*02. The mutated L chain genes of family I were derived from primary L chain genes generated by rearrangement of the germline V_L gene IGKV3-11*01 and the J_L segment IGKJ6*01. Two clonotypes, 8F9 and QG1, were assigned to family II. They used the same germline V_H and J_H gene segments as in the Abs of family I, but were derived from a different D_H element, IGHD2-5*01. There were significant differences in the nucleotide sequences of their H chain genes that could be interpreted as indicating that they originated from two independent rearrangement events. However, given the similarities in the N nucleotide-encoded elements in HCDR3, we consider that the most parsimonious interpretation of the data is that they arose from the same rearrangement event and underwent extensive somatic mutation. Likewise the L chain genes were derived from primary L chain genes generated from the same germline V_L gene used in family I, IGKV3-11*01, recombinated with a different J_L element, IGKJ1*01. Thus, the usage of different D_H regions in the ancestral germline H chain genes correlated with the use of different J_L elements in the L chain genes. The D_H reading frame of both families of clonotypes encoded hydrophilic amino acids, which are derived from the most commonly used reading frame for these D_H segments (24).

Interestingly, the reading frame for IGHD2-15*01 of family II Abs encoded paired Cys residues that were not mutated in either clonotype, perhaps because they form a disulfide bond that stabilizes the H3 loop. Such “sculpting” of the H3 loop by paired Cys residues is a feature of human V_H regions (25). The fact that all 29 of the clones from donor 1 that expressed anti-AD-2S1 Abs used the same V_H and V_L gene elements indicates that Abs derived from these V genes dominated the immune response to AD-2S1 in this individual.

We expressed the V regions of examples of clonotypes of each family in combination with the C regions of IgG1 and tested these Abs for their ability to bind AD-2S1 and to neutralize the infectivity of two strains of HCMV, a laboratory strain AD169, and a clinical isolate, 3170 (Fig. 3). Both Abs bound to the AD-2S1 peptide and both neutralized the infectivity of HCMV, confirming previous reports (17, 18) that the binding of Abs to the AD-2S1 epitope neutralized infection by HCMV. The IC_{50} for the 8F9 IgG1 was ~0.8 µg/ml and for the 2B4 IgG1 was 7.0 µg/ml.

Significantly, we noted that the V region sequences of a previously reported human mAb ITC88, which was also directed against the AD-2S1 epitope (26), indicated that it too was derived from primary Ig genes generated by recombination of the same germline V_L gene (3-11*01), the same J_H segment (4*02), and an allele of the same germline V_H gene (3-30*04) (Figs. 1 and 2). As with the case with the two families of Abs in our donor, another D_H segment (3-9*01) and another J_L element (4*01) had been used. Nevertheless there were similarities in the residues (Y/F, S/T) encoded by the 3′ end of these three D_H genes.

Table III summarizes these data that demonstrate the highly restricted usage of germline sequences in anti-gB, AD-2S1-specific Abs derived from a minimum of three distinct rearrangement events. The common features are the use of the same V_H (IGHV3-30), J_H (IGHJ4*02), and V_L (IGKV3-11*01) germline elements, the differences lying in the usage of D_H and J_L. Finally, from a second donor we cloned and sequenced cDNAs encoding the H chains made by 13 single cells (derived from 42 ELISA-positive wells from a total of 1152 wells) that were secreting Abs specific for the AD-2S1 epitope. This second donor had seroconverted just before venesection and had a serum anti-AD-2S1 titer of only 1/1200. Remarkably, 11 of the 13 H chain cDNAs closely resembled those isolated from our first donor in that they used the same germline V_H gene (3-30*18) and the same J_H, 4*02 (data not shown). Analysis of the level of somatic hypermutation in the identified V_H 3-30 genes revealed a mean of 23.9 ± 1.8 mutations, similar to that from donor 1 (30.5 ± 6.1).

**Evidence for the importance of amino acid residues encoded by codons generated by junctional events**

Both families of clonotypes from donor 1 exhibited striking similarities in those amino acids in HCDR3 that were probably encoded by N nucleotides added stochastically by TdT (Fig. 2B). In H chain genes from family I, nontemplated nucleotides were likely to have generated five codons at the 5′ end of D_H and four codons at the 3′ end. In family II, the additions 5′ to D_H were more limited than in family I, but those 3′ to D_H were similar to those seen in family I. Thus, the SGLL junctional motif at the 3′ end of D_H in family II exhibited close homology with the SGLI junctional motif at the 3′ end of D_H in family I (Table III and Fig. 2). Moreover, analysis of the N nucleotide sequences encoding the HCDR3 region of the ITC88 Ab reported by Ohlin et al. (26) also revealed similarities in the residues attributed to N nucleotide addition. Thus, the four functional codons added 3′ to D_H in the V_H gene of ITC88 encoded the identical motif, SGLI, that was seen in family I (Table III and Fig. 2). Moreover, the H chain genes we sequenced from our second donor also exhibited a very similar junctional motif (SGLP) 3′ to D_H (data not shown). There was also a conserved junctional motif 5′ to D_H with the motif DG occurring in Abs from family II and ITC88, and the homologous motif EG in Abs from family I (Table III).

There were also similarities in the stochastic elements of the rearrangement events that had generated the various L chain genes (Fig. 1B). Thus, the junction of IGKV3-11*01 and IGKJ resulted in the generation of the same residue, Pro in L, in the CDR3 of L chains in the three independent rearrangement events that gave rise...
to Igs of families I and II and ITC88. This junctional residue is encoded by nucleotides from both IGKV3-11*01 and IGKJ and occurs frequently in rearranged human κ L chains (27, 28), presumably the result of an in-frame recombination event. Although ProL95a was found in all VH regions we sequenced, a study investigating the effects of L chain shuffling with the H chain of ITC88 demonstrated that this residue was not required for high-affinity recognition of AD-2S1 (29).

Evidence for convergent evolution during somatic hypermutation of independently generated H chains and L chains

Alignment of the predicted protein sequences encoded by the germline IGHV3-30*18 gene with those encoded by the H chain genes of clonotypes from both families (Fig. 2) indicated that somatic mutation had generated multiple replacement mutations. These occurred not only in the CDRs, but also in the framework regions (FWRs). The same was the case when the sequences of the L chains were aligned with that encoded by the germline IGKV3-11*01 gene (Fig. 1), although there were fewer mutations. Thus, the frequency of mutations in the L chains averaged 5.1% and in the H chains 9.2% (Table IV). Comparisons of the ratios of R:S mutations, indicated that replacement mutations occurred more frequently in the CDRs than in the FWRs, consistent with the expected effects of Ag-driven selection for replacements in the CDRs. On average, for the L chains, the values of R:S were 2.2 in FWRs and 3.6 in CDRs and for H chains were 1.4 in FWRs and 2.6 in CDRs. Thus, although the L chains exhibited fewer mutations, there was also evidence of a relative lack of replacement mutations at certain positions in CDRs 1 and 2, suggesting that these residues may be particularly important for Ag binding. For example, in the LCDR1, residues 25–29 remained germline as did residues 50–52 in the LCDR2. Similar positions in the HCDR1 (residues 26–29) and HCDR2 (residues 52–54) also remained germline. We also observed that in all clonotypes from donor 1, the germline encoded Lys (codon AAA) residue at position H94 of FWR3 was replaced with Arg (codon AGA). This may reflect a polymorphism in the 3-30*18 germline gene, the residues at both L31 and L32 are extremely variable in the canonical Ig structure and are encoded by a hypermutable AGY codon (30). In the LCDR3, AsnL93 was repeatedly replaced by Ser in clones of both families.

There was also evidence of a relative lack of replacement mutations at certain positions in CDRs 1 and 2, suggesting that these residues may be particularly important for Ag binding. For example, in the LCDR1, residues 25–29 remained germline as did residues 50–52 in the LCDR2. Similar positions in the HCDR1 (residues 26–29) and HCDR2 (residues 52–54) also remained germline. We also observed that in all clonotypes from donor 1, the germline encoded Lys (codon AAA) residue at position H94 of FWR3 was replaced with Arg (codon AGA). This may reflect a polymorphism of the 3-30*18 gene in this individual rather than selection for this mutation. Lys and Arg are both positively charged amino acids and frequently occur in other Abs at this position.

Notably, there was also clear evidence of Ag-driven selection of somatic mutations in the FWRs of the H chains. Residues H13, H35, H47, H79, and H80 in H chains derived from independent rearrangements were often replaced by similar residues (e.g., in FWR2 MetH135 was mutated to IsoH135 or ValH135 in both families of Abs, Fig. 2A). It will be interesting to test the hypothesis that the frequently occurring replacements in FWRs influence Ag binding and their selection represents convergent evolution of combining sites with high affinity for gB.

Table IV. Analysis of somatic mutations in V<sub>H</sub> and V<sub>L</sub> sequences of anti-HCMV Abs

<table>
<thead>
<tr>
<th>Clone</th>
<th>% Identity V&lt;sub&gt;L&lt;/sub&gt;&lt;sup&gt;a&lt;/sup&gt;</th>
<th>% Identity V&lt;sub&gt;H&lt;/sub&gt;&lt;sup&gt;b&lt;/sup&gt;</th>
<th>V&lt;sub&gt;L&lt;/sub&gt; R/S&lt;sup&gt;c&lt;/sup&gt;</th>
<th>V&lt;sub&gt;H&lt;/sub&gt; R/S&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>FWR</td>
<td>CDR</td>
</tr>
<tr>
<td>2B4</td>
<td>94.1</td>
<td>88.4</td>
<td>2.3</td>
<td>2.5</td>
</tr>
<tr>
<td>FA9</td>
<td>95.8</td>
<td>89.8</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>HH12</td>
<td>93.8</td>
<td>88.1</td>
<td>3.5</td>
<td>5.0</td>
</tr>
<tr>
<td>UF12</td>
<td>92.4</td>
<td>88.8</td>
<td>1.0</td>
<td>4.0</td>
</tr>
<tr>
<td>KE5</td>
<td>95.5</td>
<td>90.5</td>
<td>4.0</td>
<td>2.5</td>
</tr>
<tr>
<td>SE12</td>
<td>93.8</td>
<td>79.9</td>
<td>0.5</td>
<td>7.0</td>
</tr>
<tr>
<td>QG1</td>
<td>93.8</td>
<td>88.1</td>
<td>1.0</td>
<td>4.5</td>
</tr>
<tr>
<td>8F9</td>
<td>92.4</td>
<td>83.0</td>
<td>2.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Average</td>
<td>94.0</td>
<td>87.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Percentage nucleotide identity to IGKV3-30*18.

<sup>b</sup> Percentage nucleotide identity to IGHV3-30*18.

<sup>c</sup> Ratio of R:S mutations (133 total mutations from 2592 total nucleotides).

<sup>d</sup> Ratio of R:S mutations (244 total mutations from 2664 total nucleotides).

<sup>e</sup> HCDR3 not included in analysis.
Ancestral unmutated primary Igs bind to HCMV gB

It is extremely improbable that by chance alone, four randomly selected primary rearrangements from three individuals (donor 1, donor 2, and donor of ITC88) would have involved the same combination of IGHV3-30, IGJH4*02, IGK3-11*01, and the junctional SGL motif that occurred in the progenitors of all four of the clonotypes that we have documented here. Thus, our observations strongly suggest that in the humans from which these Igs were derived, primary Igs encoded for by genes generated from these germline elements dominate the repertoire to the AD-2S1 epitope. Indeed, given that IGHV3-30, IGJH4*02, and IGK3-11*01 (or closely related alleles) are present in all humans and are frequently used in encoding Igs, it is probable that the primary repertoire to AD-2S1 will include Igs derived from these same genetic elements in all humans. The simplest interpretation of these observations is that these primary Igs would, when expressed as the receptors on naive B lymphocytes, bind gB with sufficient avidity to initiate the process of clonal expansion, affinity maturation, and somatic mutation. However, it is also possible that these primary Igs lacked significant affinity for gB and that the newly generated B lymphocytes on which they were expressed were initially activated to clonal expansion by other Ags, with the ability to bind gB only emerging as a result of somatic mutation.

Therefore to determine directly whether these primary Igs would bind HCMV gB or AD-2S1, we synthesized and expressed models of these primary Igs. We constructed genes by combining the germline VH, VH, JH, DH, and Jc elements that we deduced encoded the primary ancestor of clonotype 8F9. The germline-based L chain construct, termed gL(J1) encoded the junctional residues that were generated by the conjunction of nucleotide motifs occurring in 8F9. The germline-based gH(A), the added sequences encoded alanine substitutions in the junctional motif, and the somatically mutated 8F9 L chain with the junctional motif, and the somatically mutated 8F9 L chain, bound relatively strongly, and detectable binding occurred in all of the clonotypes we observed. In essence, these H chain genes were comprised of IGHV3-30*18 and IGJH4*02 flanking the DH segment used in 8F9 (IGHD2-15*01), with the addition of the sequences corresponding to the regions immediately 5' and 3' to IGHD2-15*01 in 8F9 that were presumably encoded by N nucleotide additions. In the first of these variants, termed gH(M), these added junctional sequences encoded the 5' and 3' junctional motifs seen in 8F9. In the second, termed gH(D), contained the entire IGHD2-15*01 segment plus four spacing nucleotides to maintain the correct reading frame. The predicted protein sequences are shown in Fig. 2A. Each of these germline-based VH and VH region genes was fused with the appropriate C region gene to create vectors encoding κ L chains or γ1 H chains, which when coexpressed generated IgG1 molecules. We could also express chimeric IgG1 Igs in which germline-based H or L chains were combined with their mutated counterparts from the 8F9 clonotype (H-8F9 and L-8F9). The various Abs were expressed in human 293 cells and purified from supernatants using protein A.

We tested the ability of these unmutated, primary Igs to bind to a fragment of the N terminus of gB (gBNT) that corresponded to aa 28–100, contained the AD-2S1 epitope, and was expressed on the surface of the mammalian cell line N50. As shown in Fig. 4A, the unmutated, primary Ig gH(M)gL(J1) exhibited clear binding to gBNT. As expected, the somatically mutated 8F9 Ab bound better, as evidenced by the ~200-fold lower concentration required for detectable binding. Comparison of the three unmutated, primary IgG1s that differed in the presence or absence of the H chain junctional motif occurring in 8F9 indicated that detectable binding depended on the presence of this motif (Fig. 4B). Analyses of chimeric Abs in which germline-based H chains were combined with the mutated 8F9 L chain indicated that the mutated L chain contributed significantly to binding and its use resulted in higher levels of fluorescence than seen with the primary Ig using the gL(J1) L chain. A chimera, gH(M):L8F9, which contained the germline-based H chain with the junctional motif, and the somatically mutated 8F9 L chain, bound relatively strongly, and detectable binding occurred at concentrations that were only 20-fold lower than those required with 8F9 itself. Further evidence that the somatic mutations in the 8F9 L chain promoted binding to gB came from observations on chimeras involving gH(A), the germline-based H chain in which the junctional motif had been interrupted by alanine residues. Pairing of gH(A) with the somatically mutated 8F9 L chain resulted in binding to gBNT, whereas pairing with an unmutated, germline-based L chain, gL(J1) did not (Fig 4B).

We also used ELISA to compare the ability of 8F9 and the unmutated, primary Igs to bind to the 17-aa peptide containing the AD-2S1 epitope. These results indicated that, in contrast to 8F9, the unmutated, primary Igs had a relatively weak ability to bind the peptide (Fig. 5). The best of the unmutated, primary IgG1s was gH(D)gL(J1) but, compared with 8F9, the concentrations required for detectable binding to the peptide were 10,000-fold greater. These observations indicated that the mutations that were selected for during the process of affinity maturation that generated 8F9 resulted in much greater increases in binding to the peptide than to
the folded gBNT polypeptide expressed on the surface of mammalian cells. As with binding to gBNT, there was evidence that mutations in the L chain were important for binding. Thus, chimeras of unmutated, primary gH(M), gH(A), or gH(D) H chains and the 8F9 L chain bound better to the peptide than did any of the entirely unmutated, primary Igs (Fig. 5B). The most effective pairing was seen between gH(M) and L8F9, in which case the concentrations of Ig required for binding to peptide were only ~30-fold lower than those of the 8F9 IgG1 itself. Comparison of this chimera with those containing gH(A) or gH(D) provided further evidence for the importance of the GLL junctional motif.

**Germline J<sub>μ</sub>-encoded residues improve primary Ab binding to gBNT**

In considering the possible basis for the great improvement in binding to gBNT or AD-2S1 seen when the germline gH(M) was combined with the mutated L8F9 rather than the unmutated, primary gL(J1), we noted that, while J<sub>μ</sub>1 encodes a bulky Trp residue at position L96, in both the mutated anti-AD-2S1 clonotypes that used J<sub>μ</sub>1 this residue was invariably mutated to Leu or Val (Fig. 1). Residue L96 in the CDR3 of the L chain is known to be structurally important and to form contacts with Ag in other human Abs (31). Moreover, shuffling experiments in which random L chains were combined with the H chain of ITC88 identified only one AD-2S1-binding Ab encoded by a gene incorporating J<sub>μ</sub>1 (29). Significantly, in this L chain, residue L96 was mutated from Trp to Gly, providing further evidence that Trp in this position negatively affects binding to gB. We therefore examined the effect of mutating the Trp at L96 in gL(J1) to the Val residue that is found at this position in 8F9 (Fig. 6A). Combination of this germline-based L chain containing a single somatic mutation at L96 (gL(W96V)) with the germline-based gH(M) H chain resulted in a striking improvement in binding, both to the peptide and to gBNT expressed on the surface of mammalian cells (Fig. 6B). Indeed, a chimeric Ab combining the 8F9 H chain with the germline-based L chain gL(W96V) was only 10-fold less effective in our assays for binding to the AD-2S1 peptide or gBNT than was 8F9 with its multiply mutated L chain (Fig. 6C).

This demonstration that the presence of the bulky Trp at L96 disfavored binding to gB or the peptide epitope raised the possibility that unmutated, primary Igs derived from germline J<sub>μ</sub> segments that encoded smaller side chains at L96 might bind better to gB than those encoded by IGKJ1. To test this hypothesis, we made a germline-based L chain construct (gL(J5); Fig. 1) in which IGKV3-11*01 was fused to the IGKJ5*01. The latter was the J<sub>μ</sub> segment used in the germline ancestor of the clonotypes of family I and encodes Ile at residue L96. It is noteworthy that this germline Ile was not somatically mutated in any of the six hypermutated Abs in this family, consistent with the notion that it was critical for binding to gB and the peptide.

We expressed an unmutated, primary Ig, gH(M):gL(J5), in which gL(J5) was paired with the gH(M) H chain and compared it with 8F9 or its germline ancestor gH(M):gL(J1) in terms of binding to gBNT or the AD-2S1 peptide. Remarkably, the unmutated, primary gH(M):gL(J5) Ab bound only ~10-fold less effectively to gBNT and 30-fold less effectively to the peptide than did the fully mutated 8F9 (Fig. 7). The use of J<sub>μ</sub>5 rather than J<sub>μ</sub>1 in unmutated, primary Igs resulted in an even greater improvement (~100-fold) in binding to the peptide than did in binding to gBNT on the surface of cells (Fig. 7C). Indeed, a chimeric Ig in which the 8F9 H chain was combined with the gL(J5) L chain behaved like 8F9 in binding to the peptide in the ELISA (Fig. 7D). This suggests that the amino acid replacements elsewhere in the 8F9 L chain might not be critical for high-avidity binding to the peptide, although further more detailed analyses may still reveal differences in the kinetics of the interactions of these two Abs with the peptide or the native protein. Certainly these data indicate that the unmutated, primary gH(M):gL(J5) Ab was extremely effective in recognizing both the gB protein and the linear epitope. The strong binding of the gH(M):gL(J5) Ab to gBNT suggests that the human genome contains germline elements that, when recombined with the appropriate junctional residues, encode Igs that not only have sufficient avidity for Ag to trigger B cell activation and thus initiate

---

**FIGURE 5.** Unmutated primary Igs bind weakly to the AD-2S1 epitope. Binding to the AD-2S1 peptide was determined by ELISA. Shown is the reactivity of IgG1 chimeric Igs in which various H chains, H8F9 (○), gH(M) (●), gH(A) (♦), or gH(D) (▲) were paired with the L chain variable regions of gL(J1) (A) or L8F9 (B). Inset, Protein sequence of HCDR3s with the residues predicted to be encoded by N nucleotides in Ab 8F9 in shaded boxes.

**FIGURE 6.** Mutation of L96 in gL(J1) from Trp to Val improves binding to the AD-2S1 epitope. A. Amino acid sequences of LCDR3 of L8F9, gL(J1), and gL(W96V) showing differences in residue L96. Reactivity to AD-2S1 peptide assessed by ELISA with Igs composed of gH(M) (B) or H8F9 (C) H chains and various L chains, L8F9 (○), gL(J1) (▲), or gL(W96V) (●). Insets, FACS analysis of Igs (5 μg/ml) gH(M;gL(W96V) (B) and H8F9;gL(W96V) (C) binding to NS0-gBNT cells of secondary Ab alone (shaded) and both primary plus secondary Abs (unshaded).
somatic mutation, but are themselves likely to be capable of effector activity. Certainly, there is evidence that in mice germline-based Igs bind avidly to vesicular stomatitis virus and can protect against infection (32–34).

**Discussion**

These findings demonstrate that the human germline contains genetic elements that recombine to form genes that encode primary Igs with significant ability to bind a critical structure on an important human pathogen. They suggest that adult human Abs that bind the AD-2S1-neutralizing epitope on the gB envelope protein of HCMV are predominantly encoded by genes derived from recombination events involving a single VH gene, IGHV3-30*18, and a constant or J gene, IGKV3-11*01. Unmutated, primary Igs of this structural family bound this epitope with as little as 10-fold less effectiveness than an extensively mutated, neutralizing Ab.

Although the human VH locus exhibits great polymorphism, some VH genes, including IGHV3-30, are present in most if not all humans. This is consistent with the idea that some VH genes have been fixed by selection for their fitness in contributing to genes that encode unmutated, primary Abs able to recognize common pathogens (6). Interestingly some 70% of humans exhibit a serial duplication of IGVH3-30 that results in four additional copies that differ from IGHV3-30 by only a few bases (35). It is probable all of the genes in this expanded IGHV3-30 minifamily contribute to encoding similar unmutated, primary Igs, effectively raising the frequency with which IGHV3-30 will be represented in this initial repertoire. The IGKV3-11 gene is also well conserved among humans and is one of the most frequently used V genes (7, 27). Thus, all humans are likely to have the genetic capacity to generate, in concert with the stochastic elements of the recombination process, genes that encode primary Igs that belong to this structural minifamily and have affinity for gB. The frequency with which primary Igs that are encoded by IGHV3-30, IGKV3-11, and IGHJ4*02 are generated is increased by the fact that these elements are all used frequently (7, 27, 36). The recombination signal sequences of IGHV3-30 have evolved to mediate its preferential rearrangement (8), and it is possible that its participation in the generation of the subfamily of primary Igs with affinity for AD-2S1 could have been one of the selective pressures.

Our observation of the ability of primary Igs encoded by IGHV3-30 and IGKV3-11 to bind gB depended on specific junctional sequences that were generated by stochastic events during gene rearrangement greatly reduces the frequency with which gB binding of primary Igs of the minifamily we have described will be generated. Even given the high rates of usage of IGHV3-30, IGKV3-11, and IGHJ4, the requirement for incorporation of the specific junctional nucleotides means that estimates of the frequency of the generation this minifamily of primary Igs become extremely low (~1 in 100 million). This raises the question of whether new B cells bearing such Igs would be generated at a useful frequency. One difficulty is that we have no data on how many rearrangement events occur each day in humans. However, our observation that in a single donor, the anti-AD-2S1 Abs of this minifamily were derived from a minimum of two separate rearrangement events implies that such primary Igs were generated at a biologically significant rate. Thus, considerations of the limited life span of newly generated B cells that fail to encounter their cognate Ag and antigenic competition that leads to dominance of clones with higher affinity receptors (37) make it likely that the rearrangements that generated these two clones of newly generated B cells both occurred during a short interval before or during that individual’s initial encounter with HCMV. This is because once a response to gB by one B lymphocyte clone had been initiated, somatic mutation and affinity maturation would result in the secretion of increasing concentrations of Abs of increasingly higher affinity. Thereafter, any newly generated B cells expressing primary Ig receptors that bind AD-2S1 would be outcompeted by these high-affinity serum Abs and memory B cells. Since naive B cells have a short life unless they encounter their specific Ag, this implies that the rearrangements that generated these two clones of primary B cells occurred within a short period of time. It is worth noting that the threshold required to initiate B lymphocyte activation and somatic hypermutation appears to be low ($K_a > 10^6 \text{ M}^{-1}$) (38), a fact which increases the numbers of B lymphocytes that have rearranged IGHV3-30 and IGKV3-11 that could have a biologically significant affinity for gB. Although rearrangements of other V genes could well generate unmutated, primary Abs that bind the AD-2S1 region of gB, our findings suggest that primary Igs that could outcompete primary Igs of the structural family we have described must arise much less frequently. Our results thus predict that the Abs with the structural features we have described will dominate the human immune response to AD-2S1.

The notion that particular germline elements are used preferentially in encoding Abs that bind small structures like organic chemicals or peptides is well established (39–41). Our novel contribution is that the epitope we studied is a critical site on an important, cospecifically evolved pathogen and that the Abs defined by the use of these preferentially used V genes neutralize infectivity in multiple HCMV variants (26). Thus, it would be surprising to us if one of the pressures that has stabilized the evolution of IGHV3-30 and IGKV3-11 was not selection against loss of this ability to reliably generate primary Igs that bind gB/AD-2S1 of HCMV. The AD-2S1 region of gB appears to be important in fusion of the viral envelope with the host cell, and the structural constraints imposed by such a function may account for the relative invariability of AD-2S1 in strains of HCMV (42). Indeed, there is a striking similarity between the AD-2S1 epitope of HCMV gB and the homologous region of chimpanzee CMV gB, EYIF/NRTLIP (common residues in bold, similar residues in italics) (43). Chronic infection with cospecifically evolving species of CMV was a constant feature of primate evolution (12). It is intriguing that inspection of the
chimpanzee genome reveals genes that encode peptides closely related to those encoded by the human IGHV3-30 (93% similarity) and IGKV3-11 (97% similarity).

There are two nonmutually exclusive mechanisms through which particular V genes could confer fitness in defense against pathogens. One is that these V genes have been selected for their increased fitness in generating genes that encode Igs that bind to components of important pathogens. The second is that V genes have been selected for their versatility in encoding poly-specific, flexible combining sites in primary Igs. There is evidence that other pathogens, in addition to HCMV, have exerted selective pressure to maintain the IGHV3-30 and IGKV3-11 gene elements. Thus, the IGHV3-30 gene is also used frequently to encode Abs that bind to the carbohydrates of Streptococcus pneumoniae type 23F (11). Indeed, there is evidence that human Abs to type 23F pneumococcal polysaccharide are frequently derived from unmutated, primary Igs that are very closely related to those we have shown bind gB/AD-2S1 and that, like them, utilize IGHV3-30, IGHJ4*02, and IGKV3-11, as well as, in some instances, IGKJ1*01 (11). The only differences between these unmutated, primary Igs and the gH(M):gL(J1) Ig we described in this study would be those encoded by the differences in the D\(_\text{H}\) segments and N nucleotide additions. This suggests that IGHV3-30 and IGKV3-11 may have been selected for their versatility in synergizing with junctional and N nucleotide diversity to generate genes that encode a broad repertoire of combining sites. This is consistent with crystallographic and kinetic studies that show the binding sites of unmutated, primary Igs exhibit considerable conformational flexibility that enables them to interact with a broad range of Ags (39, 40, 44). Interestingly, modest reactivity with the capsular carbohydrates of Cryptococcus neoformans and S. pneumoniae was observed with the primary Ig gH(A):gL(J1) that lacks the junctional motif we found in anti-AD-2S1 Abs and binds poorly to gB/AD-2S1 (G. R. McLean and J. W. Schrader, unpublished observations). Future work on the structural details of the interaction of IGHV3-30- and IGKV3-11-encoded residues with gB and AD-2S1 may elucidate the mechanisms through which these genetic elements encode critical features of combining sites that bind gB, while retaining the versatility to encode features of paratopes that bind different Ags, such as carbohydrates.

Our observations do provide some insights into the structural basis of the preadaptation of unmutated, primary Igs of this mini-family for recognition of the AD-2S1 region of gB. Although limited, our analyses point to considerable flexibility in the usage of D\(_\text{H}\) and J\(_\text{S}\) segments, with each of the three families of Abs using a different combination of D\(_\text{H}\) and J\(_\text{S}\) segments. This suggests that amino acids encoded by IGHV3-30, IGHJ4*02, and IGKV3-11 have important roles in shaping the paratope for the AD-2S1 region of gB. Likewise, the observation that the HCDR3 regions of clonotypes generated from distinct rearrangement events all exhibited homology of the amino acid motifs encoded by untemplated N nucleotides suggests that these motifs are also important in Ag binding. This was confirmed by the demonstration that only the unmutated, primary Ig gH(M):gL(J1) that included the junctional motif seen in 8F9 exhibited detectable binding to gB. Substitution of this motif with a full-length D\(_\text{H}\) element lacking the 8F9 junctional motif abolished binding. Likewise in the case of hybrid Igs in which germline-based H chains were combined with the somatically mutated L chain of 8F9, the absence of the junctional motifs seen in the 8F9 H chain markedly reduced the binding to gB. We have restricted our analyses to Abs of the IgG isotype, which dominate the later stages of the immune response to HCMV. In ongoing studies we will determine the function of IgM primary Igs of this mini-family, both as secreted Abs and as cell surface receptors. That similar somatic mutations were selected for during the affinity maturation of Abs originating from distinct rearrangement events implies convergent evolution. This is intriguing and demonstrates that, despite the stochastic basis of the process of affinity maturation, the independently generated high-affinity Abs nevertheless exhibited significant structural similarities. Thus, in conjunction with two stochastic processes, junctional diversity during recombination and somatic hypermutation during affinity maturation, these germline elements provided an innate foundation for the generation of high-affinity binding sites with predictable structural features. Although it cannot be formally proven, it is likely that the innate foundation of the response to AD-2S1 on HCMV that is conferred by IGHV3-30 and IGKV3-11 has been shaped by selective pressure from HCMV and other pathogens. That generation of Igs that bind gB requires not only rearrangements of these particular V genes, but also additional stochastic events, does not of course negate the proposition that these V genes can be subject to selective pressure from HCMV as long as the stochastic processes yield a biologically significant number of positive outcomes. It would not be surprising if there were other examples where particular V genes provide a foundation for the generation of genes encoding combining sites that are “preadapted” to neutralize pathogens or their products.

Acknowledgments
Thanks to Brendon Trotter for technical assistance, Andy Johnson for flow cytometry expertise, and Chantal Cho and Christy Thomson for critical reading of this manuscript.

Disclosures
The authors have no financial conflict of interest.

References


ipate in secondary immune responses when higher affinity competition is reduced. J. Exp. Med. 196:1215.


40. Romesberg, F. E., B. Spiller, P. G. Schultz, and R. C. Stevens. 1988. Immuno-


44. Manivel, V., N. C. Sahoo, D. M. Salunkhe, and K. V. Rao. 2000. Maturation of an antibody response is governed by modulations in flexibility of the antigen-com-
bining site. Immunity 13:611.