Published in final edited form as:

Clin Immunol. 2013 January ; 146(1): 46–55. doi:10.1016/j.clim.2012.11.003.

Contribution of V_H replacement products to the generation of anti-HIV antibodies

Hongyan Liao^a, Jun-tao Guo^d, Miles D. Lange^a, Run Fan^e, Michael Zemlin^f, Kaihong Su^{a,b,c}, Yongjun Guan^g, and Zhixin Zhang, Ph.D.^{a,c,*}

^aDepartments of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE 68198

^bDepartment of Internal Medicine, University of Nebraska Medical Center, Omaha, NE 68198

^cDepartments of Eppley Institute for Research in Cancer, University of Nebraska Medical Center, Omaha, NE 68198

^dDepartment of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, NC 28223

^eDepartment of Microbiology, University of Alabama at Birmingham, AL 35294

^fDepartment of Pediatrics, Philipps Universität Marburg, Germany

^gInstitute of Human Virology and Department of Microbiology & Immunology, University of Maryland School of Medicine, Baltimore, MD 21201

Abstract

V_H replacement occurs through RAG-mediated secondary recombination to change unwanted IgH genes and diversify antibody repertoire. The biological significance of V_H replacement remains to be explored. Here, we show that V_H replacement products are highly enriched in IgH genes encoding anti-HIV antibodies, including anti-gp41, anti-V3 loop, anti-gp120, CD4i, and PGT antibodies. In particular, 73% of the CD4i antibodies and 100% of the PGT antibodies are encoded by potential V_H replacement products. Such frequencies are significantly higher than those in IgH genes derived from HIV infected individuals or autoimmune patients. The identified V_H replacement products encoding anti-HIV antibodies are highly mutated; the V_H replacement “footprints” within CD4i antibodies preferentially encode negatively charged amino acids within the IgH CDR3; many IgH encoding PGT antibodies are likely generated from multiple rounds of V_H replacement. Taken together, these findings uncovered a potentially significant contribution of V_H replacement products to the generation of anti-HIV antibodies.

Keywords

IgH gene; V_H replacement; anti-HIV antibody; gp120; CD4i antibody

* Address correspondence to: Zhixin (Jason) Zhang, Ph.D., Associate Professor, LTC 11706A, Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE 68198-7660, Tel: (402)559-7637; Fax: (402)559-4990; zhangj@unmc.edu.

Conflict of interest statement: The author(s) declare that there are no conflicts of interest.

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

1. Introduction

Neutralizing antibodies are important effectors in humoral immune response against invading bacteria or virus, including HIV [1-3]. Normally, neutralizing antibodies can be induced from our immune system through conventional vaccine approaches and protect us from later infections [1-3]. Upon exposure to HIV, most individuals develop an early anti-HIV immune response and generate specific anti-HIV antibodies. However, this early wave of immune response can only partially inhibit but is not sufficient to prevent HIV spreading and transmission [1-4]. It is very difficult to elicit and maintain effective anti-HIV neutralizing antibody responses in either natural infections or vaccination [1-3].

The difficulties to elicit effective anti-HIV neutralizing response come from both the HIV virus and the immune system. Studies of the viral genome and envelope protein structures of HIV virus indicated that HIV has evolved many strategies, such as the variation of the viral epitope, oligomerization of the viral envelope proteins, glycan shield, and steric interference at the virus-cell interface to evade neutralizing antibody responses [1-5]. Analyses of the Ig genes and protein structures of anti-HIV antibodies also provided valuable information regarding the co-evolution of HIV with the immune system and the potential difficulties to generate anti-HIV neutralizing antibodies. Currently, there are several groups of monoclonal antibodies (mAb) that display neutralizing activities against HIV entry [2, 3]. These include the antibodies that react with the gp41 membrane proximal external regions (MPER) (gp41); the CD4 binding site (CD4b); the gp120 V3 loop (V3); the CD4-induced CCR5 co-receptor binding site on gp120 (CD4i); and the PGT type of highly potent neutralizing antibodies that penetrate the glycan shield to bind the V3 loop (PGT) [2, 3, 6-8]. Detailed analyses of the interaction between these neutralizing antibodies with HIV envelope proteins have provided important information regarding the structural requirements for different anti-HIV neutralizing antibodies. For examples, the IgH CDR3 regions of the CD4i antibodies are essential for interacting with the CCR5 binding pocket in gp120. To facilitate the interaction, CD4i antibodies must have long IgH CDR3 regions with negatively charged amino acids to contact with the positively charged CCR5 binding pocket [9, 10]. Interestingly, several CD4i antibodies have even acquired posttranslational sulfation modifications on the tyrosine residues within their IgH CDR3 regions in mimicry of the CCR5 N-terminal peptide [9, 10]. Tyrosine sulfation modification provides additional negative charges on the IgH CDR3 for CD4i antibodies to interact with the CCR5 binding site. Originally, it has been shown that the F(ab')₂ fractions of CD4i antibodies display broad neutralizing activities against a range of HIV-1 primary isolates [11]. CD4i antibodies can be detected from sera of most HIV patients [4], however their *in vivo* biological functions are still not clear. Antibodies 2F5 and 4E10 represent another group of potent HIV neutralizing antibodies that react with the gp41 membrane proximal external regions (MPER) [6]. Such activity is rarely detected in sera of HIV patients [4]. Recent studies showed that 2F5 and 4E10 antibodies are autoreactive against various self-antigens, including cardiolipin, cytoplasmic proteins, and Ro antigen [6]. Both 2F5 and 4E10 antibodies have relatively long IgH CDR3 regions with multiple positively charged Arg residues. Based on these observations, it has been speculated that potent HIV neutralizing antibodies might be generated through rare recombination events or be eliminated from the antibody repertoire during early B cell development due to their self-reactivities [6]. Recent studies uncovered a new group of PGT type of highly potent neutralizing antibodies [7, 8]. Most of these PGT antibodies have extremely long IgH CDR3 regions, which can penetrate the sugar glycan shield on gp120 to interact with the V3 loop [8]. Interestingly, the PGT type of Abs showed none auto/poly-reactivities.

The diversified antibody repertoire is generated through recombination activation gene product (RAG)-mediated V(D)J recombination process [12-14]. Due to the random nature of

V(D)J recombination, many Ig rearrangements might be either non-functional or encoding self reactive antibodies. Thus, early B lineage cells can edit the previously formed Ig genes through secondary recombination on IgH or IgL genes [12, 15-18]. V_H replacement occurs through RAG-mediated secondary rearrangement involving a cryptic recombination signal sequence (RSS) within a previously rearranged V_H(D)J_H joint with a 23 bp RSS from an upstream V_H gene [17-20]. The biological potential of V_H replacement to edit unwanted IgH gene, diversify the IgH repertoire, and rescue B cells with non-functional IgH rearrangement have been demonstrated in mice carrying different knocked-in IgH V(D)J exons within the IgH loci [21-26]. V_H replacement changes almost the entire V_H coding region but retains a stretch of nucleotides from the previous V_H gene as a “footprint” [20]. Two of the intrinsic features of V_H replacement are that V_H replacement elongates the IgH CDR3 region and V_H replacement “footprints” preferentially encode charged amino acids [17, 18, 20]. V_H replacement products contributes to about 5% of the primary B cell repertoire in healthy donors [20]. Our recent analysis of IgH gene sequences from the IMGT database revealed that the frequencies of V_H replacement products are significantly elevated in IgH genes encoding anti-HIV antibodies, especially CD4i and PGT antibodies. This finding led us to explore a potential contribution of V_H replacement products in generation of anti-HIV antibodies.

2. Materials and methods

2.1. Sequence analysis and identification of potential V_H replacement products

For the initial sequence analysis, all the IgH sequences were obtained from the IMGT database. Potential V_H, D_H, and J_H germline gene usages were assigned based on the Igbblast program (<http://www.ncbi.nlm.nih.gov/igblast/>) or V-QUEST program (<http://imgt.cines.fr/textes/vquest/>). Potential V_H replacement products were identified using our previously established method [20]. Briefly, after assignment of the potential usages of the V_H, D_H, and J_H germline genes, the V-D junction regions (N1 regions) were analyzed for the existence of potential V_H replacement “footprints” [20]. The initial screen of the IgH sequences from the IMGT database was conducted by computational analysis to identify stretches of pentameric nucleotide motifs within the V-D junctions that matched with the 3′ ending sequences from V_H germline genes (5 nucleotides with no mismatch). For IgH genes encoding different subgroups of anti-HIV antibodies, we searched for pentameric and hexameric (6 nucleotides with one mismatch) V_H replacement “footprint” motifs within the V-D junctions. The complete group of IgH genes encoding CD4i antibodies was kindly provided by Drs. P. Kwon and C. Huang at NIH.

Mutation analyses of the identified IgH genes are conducted with the V-QUEST program. The replacement versus silence mutation (R/S) ratio was used as an indicator for antigen selection [27].

2.2. Statistical analysis

The frequency of V_H replacement products in IgH genes encoding anti-HIV antibodies was compared with that in the IgH genes obtained from normal peripheral B cells (control), RA and SLE patients, and HIV infected individuals. Statistical significance was determined by two-tailed *Chi-square* test with Yates' correction. Significant difference was determined if the *P* value < 0.05. Extremely significant difference was determined if the *P* value < 0.0001.

2.3. Computational modeling of the IgH CDR3 structures of CD4i and PGT antibodies

The structures of the CD4i and PGT antibody CDR3 regions were analyzed to determine the positions of the residues contributed by the V_H replacement “footprints”. For CD4i or PGT antibodies with unsolved 3-D structures, we used homology modeling technique to predict

the structures of their CDR3 regions. Briefly, the amino acid sequences of CD4i or PGT antibodies (heavy chain) were aligned to their corresponding sequences of the solved structures from Protein Data Bank (PDB) [28], for example, 1RZFH (PDB code, with the fifth letter indicating the chain ID) for E51, 1RZIH for 47e, and 3TYG for PGT128. The models were generated using the homology modeling package, Modeller (version 8v1) [29]. Molecular graphics were generated using the Pymol program (<http://www.pymol.org>).

3. Results

3.1. Elevated frequencies of V_H replacement products in IgH genes encoding anti-HIV antibodies

Our analysis of human IgH sequences from the IMGT database revealed an interesting finding that the frequencies of V_H replacement products are significantly elevated in IgH genes derived from autoimmune diseases, such as rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE); from HIV infected patients; and from IgH genes encoding anti-HIV antibodies (Fig. 1A). In particular, the frequency of V_H replacement products in the 100 IgH genes encoding anti-HIV antibodies is 41% (Fig. 1A), which is significantly higher than that in the IgH sequences obtained from normal B cells (control, 5.7%) or patients with autoimmune diseases ($p=0.0001$, two-tailed Chi-square test). The frequency of V_H replacement products in IgH genes encoding anti-HIV antibodies is also significantly higher than those in IgH genes derived from HIV infected individuals ($p=0.0001$, two-tailed Chi-square test), suggesting that these V_H replacement products encoding anti-HIV antibodies were positively selected during anti-HIV immune response. This observation uncovered a previously unrealized contribution of V_H replacement to the generation of anti-HIV antibodies.

Based on the available information regarding the antigenic epitopes of these anti-HIV antibodies, we separated these IgH genes into following subgroups and further determined the distribution of V_H replacement products in each subgroup (Fig. 1B). In the 13 IgH genes encoding anti-gp41 antibodies (not including the IgH genes encoding the 2F5 and 4E10 antibodies), the frequency of V_H replacement products is about 15%. In the 42 IgH genes encoding anti-gp120 V3 loop antibodies, the frequency of V_H replacement products is 21%, which is significantly higher than that in the normal B cells (control, 5.7%) ($p=0.0006$, two-tailed Chi-square test). In the 17 IgH genes encoding anti-HIV gp120 antibodies, the frequency of V_H replacement products is 47%, which is significantly higher than that in the normal peripheral B cells ($p=0.0001$, two-tailed Chi-square test). In the 11 IgH genes encoding CD4i antibodies, the frequency of V_H replacement products is 73%, which is significantly higher than that in the normal peripheral naïve B cells ($p=0.0001$, two-tailed Chi-square test). For the 17 IgH genes encoding the PGT type of highly potent neutralizing antibodies, all of them (100%) are potential V_H replacement products. Such frequency is significantly higher than that in the control sequences ($p=0.0001$, two-tailed Chi-square test). These results showed that V_H replacement products are highly enriched in IgH genes encoding anti-HIV antibodies, especially CD4i and PGT antibodies.

3.2. Contribution of V_H replacement to the generation of anti-HIV antibodies

From a total of 100 IgH genes encoding different anti-HIV antibodies, we identified 41 potential V_H replacement products, which contain the pentameric V_H replacement footprint motifs with no mismatch or the hexameric V_H replacement footprint motifs with one mismatch (Table 1). Most of the identified V_H replacement products contain one V_H replacement “footprint” at the V-D junction. Strikingly, for the IgH genes encoding PGT antibodies, many of them have 3 or 4 V_H replacement footprints within their V-D junctions, which might be generated through 3 or 4 rounds of V_H replacement recombination,

respectively. Although our previous studies showed that serial V_H replacement could occur in the immortalized human EU12 B lineage cell line, IgH genes derived from multiple rounds of V_H replacement recombination are very rare in the normal IgH repertoire. In our previous analysis of 402 IgH genes derived from control donors, there is only one IgH sequence (or 0.3% of IgH genes) containing two potential V_H replacement “footprints” at its V-D junction. There is no IgH gene with 3 or 4 V_H replacement footprints in those 402 IgH sequences. Here, in the V-D junctions of the 17 IgH genes encoding anti-HIV antibodies, 7 of them (or 41%) have 4 V_H replacement footprints; 2 of them (or 12%) have 3 V_H replacement footprints; 4 of them (or 23.5%) have 2 V_H replacement footprints. The frequencies of IgH genes generated from multiple rounds of V_H replacement in IgH genes encoding PGT antibodies are significantly higher than that in the control group ($p = 0.0001$, two-tailed Chi-square test).

3.3. The accumulated V_H replacement products encoding anti-HIV antibodies are highly mutated

To further determine if the identified V_H replacement products had been actively involved in anti-HIV immune response, we analyzed the somatic mutations within the V_H coding regions. All these identified V_H replacement products have an overall mutation rate at 15% (Table 2). Moreover, the average replacement versus silent (R/S) mutation ratio in the V_H coding ranges is 4.59, suggesting that these V_H replacement products encoding anti-HIV antibodies have been positively selected during anti-HIV response.

3.4. The V_H replacement “footprints” in CD4i antibodies preferentially contribute negatively charged amino acids

The intrinsic feature of V_H replacement is that the V_H replacement footprint extends the IgH CDR3 region with preferential introduction of charged amino acids [20]. The dominant contribution of V_H replacement products in IgH genes encoding CD4i antibodies led us to further analyze the CDR3 length and amino acid usage of these V_H replacement products encoded CD4i antibodies. When compared with the IgH CDR3 length distribution of 4713 human IgH sequences [30], V_H replacement products encoded CD4i antibodies have an average CDR3 length of 20.8 AA, which is significantly longer than the average CDR3 length (14 AA) in the control group of 4713 human IgH sequences ($p = 0.0001$, two-tailed Chi-square test) (Fig. 2A). Moreover, the IgH CDR3 regions of CD4i antibodies contain many negatively charged Glu and Asp residues, but few positively charged Arg residue (Fig. 2B). Previous structural analysis of the CD4i antibodies indicates that the long and acidic IgH CDR3 regions are required for CD4i antibodies to reach the positively charged binding cavity on gp120 [9-11]. Here, this group of V_H replacement products exactly fulfilled the structural requirements of CD4i antibodies.

3.5. The negatively charged amino acids contributed by the V_H replacement footprints in CD4i antibodies are distributed at critical position for binding gp120

Detailed analysis of the amino acid sequences of the V_H replacement products encoding CD4i antibodies revealed that almost all of the V_H replacement “footprints” encoded negatively charged amino acid residues at the CDR3 regions (Fig. 3A). For antibodies 17b, 47e, SB1, and X5, the V_H replacement “footprints” contribute two negatively charged amino acids into the CDR3 region. Although V_H replacement footprint preferentially encodes charge amino acids, the residues can be either positively or negatively charged [20]. Here, in sharp contrast to the identified V_H replacement footprints encoded amino acids from 402 control IgH gene sequences, the V_H replacement footprints in these CD4i antibodies are dominantly encoding negatively charged residues (Fig. 3B). These results indicate that these V_H replacement products are specifically selected to encode CD4i antibodies.

Previous structural analyses of several CD4i antibodies showed that the long IgH CDR3 regions protrude into the conserved CCR5 binding cavity on gp120 [9, 10]. The negatively charged residues within the CDR3 of CD4i antibodies and tyrosine sulfation modifications are critical for contacting the positively charged CCR5 binding pocket [9]. To investigate the possible roles of these negatively charged residues encoded by the V_H replacement “footprints”, we searched the PDB database for solved structures with intact CDR3 regions [28] and mapped the locations of these negatively charged amino acids encoded by the V_H replacement “footprints”. For CD4i antibodies that don't have complete structure in the CDR3 regions, such as E51 and 47e, we applied homology modeling techniques to model their structures. Interestingly, all the negatively charged residues encoded by the V_H replacement “footprints” are located at or near the tip of the CDR3 loops, which are the predicted regions for contacting the binding pocket on gp120 (Fig. 3B). Thus, these negatively charged amino acid residues encoded by the V_H replacement ‘footprints’ are critical for binding gp120.

3.6. Multiple rounds of V_H replacement generate IgH genes with extremely long CDR3 to encode PGT antibodies

For the PGT type of antibodies, it has been shown that most of them have very long CDR3 regions, which are essential for penetration of the sugar glycan shield on gp120 [7, 8]. Our sequence analysis showed that 9 out of 17 IgH genes encoding PGT antibodies contain 3 or 4 V_H replacement footprints, which might be generated through 3 or 4 rounds of V_H replacement recombination, respectively. Multiple rounds of V_H replacement are very rare in normal B cells. Here, multiple rounds of V_H replacement may provide a unique mechanism to generate these IgH genes with extremely long CDR3. These identified V_H replacement footprints in the PGT antibodies encoded either positively or negatively charged residues (Fig. 4A and 4B). There is no special selection on the charged residues encoded by V_H replacement footprints in these PGT antibodies (Fig. 4B). Computer modeling of the CDR3 regions of these PGT antibodies indicated that the charged residues are distributed at the side of the CDR3 loops (Fig. 4C), which might be important for penetration of the negatively charged sugar glycan shield on gp120.

Discussion

V_H replacement was first observed in mouse pre B leukemia cells to generate functional IgH genes from non-functional IgH rearrangements [31, 32]. Later, the occurrence of V_H replacement has been demonstrated in different mouse models carrying artificially knocked-in IgH genes [21, 24, 26, 33]. Our previous studies demonstrated that V_H replacement occurs in human bone marrow immature B cells and V_H replacement products contribute to about 5% of the peripheral B cell repertoire in the control donors [20]. V_H replacement is considered as a receptor editing process to change non-functional IgH genes or IgH genes encoding self reactive antigen receptors and also contributes to antibody repertoire diversification [17, 18]. The biological significance of V_H replacement in human is still largely unknown. Here, the highly enrichment of V_H replacement products in IgH genes encoding different types of anti-HIV antibodies is a striking finding, which uncovers a potentially important function of V_H replacement products in B cell anti-HIV response. The intrinsic feature of V_H replacement is to generate IgH genes with long CDR3 regions and with additional charged amino acids contributed by the V_H replacement footprints. Such feature of V_H replacement products exactly fulfills the structural requirements for CD4i antibodies, long IgH CDR3 regions with negatively charged residues, to reach to the positively charged CCR5 binding pocket; multiple rounds of V_H replacement may also provide a unique mechanism to generate IgH genes with extremely long CDR3 regions to encode the PGT type of antibodies.

The highly enrichment of V_H replacement products in IgH genes encoding anti-HIV antibodies, especially CD4i and PGT antibodies suggested that V_H replacement provides a unique mechanism to generate anti-HIV antibodies. This idea is further supported by the accumulation of somatic mutations within the V_H coding regions of these identified V_H replacement products and the special selection of negatively charged amino acids encoded by the V_H replacement footprints in CD4i antibodies. Currently, it is not clear why and how V_H replacement products are enriched in IgH genes encoding anti-HIV antibodies. The frequencies of V_H replacement products are elevated in HIV infected individuals and in patients with RA and SLE. The further enrichment of V_H replacement products in IgH genes encoding anti-HIV antibodies could be due to both the induction of V_H replacement recombination in the immature B cells and the positive selection of V_H replacement products encoding anti-HIV antibodies in the mature B cells in HIV patients. Our previous studies showed that V_H replacement occurs in human bone marrow immature B cells. It is plausible that such recombination is enhanced during long term exposure to HIV viral antigens and chronic inflammatory stimulation. On the other hand, the selective usage of V_H replacement products with negatively charged CDR3 regions in the CD4i antibodies as well as the accumulation of somatic mutations in V_H replacement products encoding anti-HIV antibodies indicated that B cells expressing such IgH genes are positively selected during anti-HIV response. Given the current difficulties to elicit an effective anti-HIV neutralizing antibody response, understanding how these V_H replacement products are generated and selected during the course of HIV infection will provide novel information for future HIV vaccine design.

The enrichment of V_H replacement products in IgH genes encoding anti-HIV antibodies is not an isolated phenomenon. In our current studies, elevated frequencies of V_H replacement products were also found in IgH genes encoding other anti-viral antibodies, such as anti-HBV, anti-HCV antibodies, and in IgH genes derived from different autoimmune diseases, such as rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE). In general, the frequencies of V_H replacement products are elevated in anti-viral responses and autoimmune diseases. Interestingly, V_H replacement products seem to be further selected to generate anti-viral antibodies and antigen specific autoantibodies (Lange and Zhang, separate manuscripts). It is plausible that V_H replacement is initially induced by prolonged exposure to viral antigens to generate anti-viral antibodies. However, the left over antibodies after viral clearance may cross-react to self antigens and contribute to autoimmune diseases. In supporting this interesting notion, HIV infected individuals often develop multiple autoimmune symptoms at the later stage of diseases. Some of these symptoms are indistinguishable from SLE.

Collectively, the results presented in this study revealed a remarkable, unanticipated contribution of V_H replacement products to the generation of anti-HIV antibodies, especially CD4i and PGT antibodies. Further studies are underway focused on the biological significance of V_H replacement products in the anti-HIV immune response and the potential regulation of V_H replacement in human immature B cells by HIV antigens. Understanding how such V_H replacement products are generated and enriched during the course of HIV infection will provide novel information of anti-HIV antibody response and will help us to design better HIV vaccines and immunization protocols to elicit effective neutralizing antibody response.

Acknowledgments

We thank Drs. Peter Kwong and Chih-Chin Huang for sharing the IgH gene sequences of the CD4i antibodies and helpful discussions; Drs. Max D. Cooper, George Shaw, Beatrice Hahn, and Peter D. Burrows for critical discussions and reviewing of the manuscript. This work was supported in part by a UAB CFAR pilot grant and by

NIH grants (K01AR048592, R21AI073174, and R56 AI098576-01A1) to ZZ, DFG grant (SFB/TR22 TPA17) to MZ.

References

- Burton DR, Desrosiers RC, Doms RW, Koff WC, Kwong PD, Moore JP, Nabel GJ, Sodroski J, Wilson IA, Wyatt RT. HIV vaccine design and the neutralizing antibody problem. *Nat Immunol*. 2004; 5:233–36. [PubMed: 14985706]
- Burton DR, Stanfield RL, Wilson IA. Antibody vs. HIV in a clash of evolutionary titans. *Proc Natl Acad Sci U S A*. 2005; 102:14943–48. [PubMed: 16219699]
- Nabel GJ. Immunology. Close to the edge: neutralizing the HIV-1 envelope. *Science*. 2005; 308:1878–79. [PubMed: 15976295]
- Decker JM, Bibollet-Ruche F, Wei X, Wang S, Levy DN, Wang W, Delaporte E, Peeters M, Derdeyn CA, Allen S, Hunter E, Saag MS, Hoxie JA, Hahn BH, Kwong PD, Robinson JE, Shaw GM. Antigenic conservation and immunogenicity of the HIV coreceptor binding site. *J Exp Med*. 2005; 201:1407–19. [PubMed: 15867093]
- Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, Salazar-Gonzalez JF, Salazar MG, Kilby JM, Saag MS, Komarova NL, Nowak MA, Hahn BH, Kwong PD, Shaw GM. Antibody neutralization and escape by HIV-1. *Nature*. 2003; 422:307–12. [PubMed: 12646921]
- Haynes BF, Fleming J, St Clair EW, Katinger H, Stiegler G, Kunert R, Robinson J, Scearce RM, Plonk K, Staats HF, Ortel TL, Liao HX, Alam SM. Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science*. 2005; 308:1906–08. [PubMed: 15860590]
- Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien JP, Wang SK, Ramos A, Chan-Hui PY, Moyle M, Mitcham JL, Hammond PW, Olsen OA, Phung P, Fling S, Wong CH, Phogat S, Wrin T, Simek MD, Koff WC, Wilson IA, Burton DR, Poignard P. Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature*. 2011; 477:466–70. [PubMed: 21849977]
- Pejchal R, Doores KJ, Walker LM, Khayat R, Huang PS, Wang SK, Stanfield RL, Julien JP, Ramos A, Crispin M, Depetris R, Katpally U, Marozsan A, Cupo A, Malveste S, Liu Y, McBride R, Ito Y, Sanders RW, Ogohara C, Paulson JC, Feizi T, Scanlan CN, Wong CH, Moore JP, Olson WC, Ward AB, Poignard P, Schief WR, Burton DR, Wilson IA. A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. *Science*. 2011; 334:1097–103. [PubMed: 21998254]
- Huang CC, Venturi M, Majeed S, Moore MJ, Phogat S, Zhang MY, Dimitrov DS, Hendrickson WA, Robinson J, Sodroski J, Wyatt R, Choe H, Farzan M, Kwong PD. Structural basis of tyrosine sulfation and VH-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proc Natl Acad Sci U S A*. 2004; 101:2706–11. [PubMed: 14981267]
- Huang CC, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, Wyatt R, Kwong PD. Structure of a V3-containing HIV-1 gp120 core. *Science*. 2005; 310:1025–28. [PubMed: 16284180]
- Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*. 1998; 393:648–59. [PubMed: 9641677]
- Rajewsky K. Clonal selection and learning in the antibody system. *Nature*. 1996; 381:751–58. [PubMed: 8657279]
- Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. *Cell*. 2002; 109:S45–S55. [PubMed: 11983152]
- Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. *Annu Rev Genet*. 2011; 45:167–202. [PubMed: 21854230]
- Nussenzweig MC. Immune receptor editing: revise and select. *Cell*. 1998; 95:875–78. [PubMed: 9875841]
- Nemazee D, Weigert M. Revising B cell receptors. *J Exp Med*. 2000; 191:1813–17. [PubMed: 10839798]
- Zhang Z, Burrows PD, Cooper MD. The molecular basis and biological significance of V_H replacement. *Immunol Reviews*. 2004; 197:231–42.

18. Zhang Z. V_H replacement in mice and humans. *Trends Immunol.* 2007; 28:132–37. [PubMed: 17258935]
19. Zhang Z, Wang YH, Zemlin M, Findley HW, Bridges SL, Burrows PD, Cooper MD. Molecular mechanism of serial V_H gene replacement. *Ann N Y Acad Sci.* 2003; 987:270–73. [PubMed: 12727651]
20. Zhang Z, Zemlin M, Wang YH, Munfus D, Huye LE, Findley HW, Bridges SL Jr, Roth DB. Contribution of V_H gene replacement to the primary B cell repertoire. *Immunity.* 2003; 19:21–31. [PubMed: 12871636]
21. Chen C, Nagy Z, Prak EL, Weigert M. Immunoglobulin heavy chain gene replacement: a mechanism of receptor editing. *Immunity.* 1995; 3:747–55. [PubMed: 8777720]
22. Chen C, Nagy Z, Radic MZ, Hardy RR, Huszar D, Camper SA, Weigert M. The site and stage of anti-DNA B-cell deletion. *Nature.* 1995; 373:252–55. [PubMed: 7816141]
23. Chen C, Prak EL, Weigert M. Editing disease-associated autoantibodies. *Immunity.* 1997; 6:97–105. [PubMed: 9052841]
24. Cascalho M, Wong J, Wabl M. V_H gene replacement in hyperselected B cells of the quasimonoclonal mouse. *J Immunol.* 1997; 159:5795–801. [PubMed: 9550375]
25. Lopez-Macias C, Kalinke U, Cascalho M, Wabl M, Hengartner H, Zinkernagel RM, Lamarre A. Secondary rearrangements and hypermutation generate sufficient B cell diversity to mount protective antiviral immunoglobulin responses. *J Exp Med.* 1999; 189:1791–98. [PubMed: 10359583]
26. Koralov SB, Novobrantseva TI, Konigsmann J, Ehlich A, Rajewsky K. Antibody repertoires generated by V_H replacement and direct V_H to J_H joining. *Immunity.* 2006; 25:43–53. [PubMed: 16860756]
27. Kirkham PM, Mortari F, Newton JA, Schroeder HW. Immunoglobulin V_H clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO J.* 1992; 11:603–09. [PubMed: 1537339]
28. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res.* 2000; 28:235–42. [PubMed: 10592235]
29. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993; 234:779–815. [PubMed: 8254673]
30. Zemlin M, Klinger M, Link J, Zemlin C, Bauer K, Engler JA, Schroeder HW Jr, Kirkham PM. Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures. *J Mol Biol.* 2003; 334:733–49. [PubMed: 14636599]
31. Kleinfield R, Hardy RR, Tarlinton D, Dangi J, Herzenberg LA, Weigert M. Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1+ B-cell lymphoma. *Nature.* 1986; 322:843–46. [PubMed: 3092106]
32. Reth M, Gehrmann P, Petrac E, Wiese P. A novel V_H to V_HDJ_H joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production. *Nature.* 1986; 322:840–42. [PubMed: 3092105]
33. Lutz J, Muller W, Jack HM. V_H replacement rescues progenitor B cells with two nonproductive VDJ alleles. *J Immunol.* 2006; 177:7007–14. [PubMed: 17082616]

Abbreviations used in this paper

HIV	human immunodeficiency virus
Ig	immunoglobulin
IgH	immunoglobulin heavy chain
RAG	recombination activating gene products
RSS	recombination signal sequence
cRSS	cryptic recombination signal sequence

CD4i	CD4 binding induced
CDR	complementarity determine region

Highlight

1. New results of the frequencies of VH replacement products in IgH genes derived HIV infected individuals and autoimmune patients (RA and SLE) (Figure 1A).
2. New results of the highly enrichment of VH replacement products in IgH genes encoding the PGT type of highly potent neutralizing antibodies (Figure 1B, Figure 4, and part of Table 1 and Table 2).

Figure 1. Elevated frequencies of V_H replacement products in IgH genes encoding anti-HIV antibodies

A. Frequencies of V_H replacement products in normal IgH genes (control, $n=402$), IgH genes derived from patients with RA ($n=772$), SLE ($n=145$), and HIV ($n=422$), and IgH genes encoding anti-HIV antibodies ($n=100$). **B.** Detailed comparison of the frequencies of V_H replacement products in IgH genes encoding anti-gp41 antibodies ($n=13$), anti-gp120 V3 loop antibodies ($n=42$), anti-gp120 antibodies ($n=17$), CD4i antibodies ($n=11$), and PGT antibodies ($n=17$). Statistical significance against the frequency of V_H replacement products in the normal peripheral B cells (control, 5.7%) were determined by two tailed Chi-square test with Yates' correction. Significant difference ($P < 0.05$) is indicated by (*) and extremely significant difference ($P < 0.0001$) is indicated by (**).

Figure 2. CD4i antibodies encoded by V_H replacement products have long CDR3 regions enriched with negatively charged amino acid residues

A. V_H replacement products encoded CD4i antibodies have long CDR3 compared with those from 4713 normal IgH sequences [30]. Histogram shows the distribution of IgH genes according to their CDR3 lengths. B. V_H replacement products encoded CD4i antibodies have more negatively charged amino acid residues within their CDR3 regions. The usage of different amino acids in the CDR3 regions of V_H replacement products encoding CD4i antibodies and in 4713 normal IgH sequences.

Figure 3. The negatively charged amino acids contributed by the V_H replacement “footprints” in CD4i antibodies are located at the interaction interface for binding gp120

A. Amino acid sequence analysis of the CDR3 region of CD4i antibodies encoded by V_H replacement products. The amino acids contributed by the V_H replacement “footprints” are underlined. Negatively charged residues are highlighted in red. B. Frequencies of none charged, positively charged, and negatively charged amino acids encoded by the identified V_H replacement products from control IgH gene sequences or IgH genes encoding CD4i antibodies. The differences of negatively charged residues in these two groups of V_H replacement footprints encoded amino acids were determined by two tailed Chi-square test with Yates' correction ($p = 0.0011$). C. Structural analysis of the IgH CDR3 of CD4i antibodies. The complex structure (PDB ID: 2B4C) shows the relative positions of each peptide chain, gp120 (orange), CD4 (yellow), X5 heavy chain (blue), and X5 light chain (gray). Solved or predicted structures of the IgH CDR3 regions in CD4i antibodies X5, 17b, 47e, and E51 are shown. The yellow arrows represent β -strand conformation, which are located at the ends of the CDR3 regions. Amino acids encoded by the V_H replacement “footprints” (highlighted in red) within X5, 17b, 47e, and E51 antibodies are located at or near the tips of the CDR3 loops, which are the predicted interaction interface for binding gp120.

Figure 4. Multiple rounds of V_H replacement generate IgH genes with extremely long CDR3 to encode PGT antibodies

A. Amino acid sequence analysis of the CDR3 regions of PGT antibodies encoded by V_H replacement products. The amino acids contributed by the V_H replacement “footprints” are underlined. Positively charged residues are highlighted in *red*; negatively charged residues are highlighted in *blue*. B. Frequencies of none charged, positively charged, and negatively charged amino acids encoded by the identified V_H replacement products from control IgH gene sequences or IgH genes encoding PGT antibodies. The differences of charged residues in these two groups of V_H replacement footprints encoded amino acids were compared and showed no significant difference. C. Structural analysis of IgH CDR3 of PGT antibodies. The complex structure (PDB ID: 3TYG) shows the relative positions of gp120 (gray) and PGT128 antibody (blue and orange). Solved or predicted structures of the IgH CDR3 regions in antibodies PGT121, 122, 123, 127, 128, 130, and 141 are shown. The yellow arrows represent β -strand conformation, which are located at both ends of the CDR3 regions. Positively charged residues are shown in *red* and negatively charged residues are shown in *blue*.

Table 1
Identification of VH replacement products in IgH genes encoding anti-HIV antibodies.

[illegible]

Sequence ID	V _H	V _H	P	N1	P	D _H	D _H
PGT	PGT121H	IGHV4-59	tgccgcgaga	acactgcacggagagaaattatgggaacgtgtgcccacat		gagtggtt	IGHD3-10
	PGT122H	IGHV4-61	tgccgcgga	caacaaaacacgcggggagattatggcgcgttgccctcaagagtggtt		cacctat	IGHD3-16
	PGT123H	IGHV4-59	tgccgcg	cg	agccacttcacgcggggagattatggggatggttgcccctcgga	gagtggtt	IGHD3-16
	PGT125H	IGHV4-b	tgctgcgcgga	ttcgacggcgaaag		tcctgggct	IGHD4-17
	PGT126H	IGHV4-b	tgctgcgcgga	ttcgacggcgaaag		ttttggggt	IGHD3-22
	PGT127H	IGHV4-61	tgctgcgcgga	ttcggcgcgcggaag		ttctagtggt	IGHD3-10
	PGT128H	IGHV4-b	tgctgcgcgga	ttcggcgcgcggaag		ttttacgct	IGHD4-17
	PGT130H	IGHV4-39	tg	cgtacgataccggcgcgacactta		tattatattg	IGHD1-14
	PGT131H	IGHV4-39	tg	cgttcgataccggcgcgacattta		tactatattg	IGHD2-8
	PGT135H	IGHV4-34	tgctgcgaga	cacggacatacatgatatacatggttttcaatgggtggtccctattgcg		ggcggtt	IGHD3-16
	PGT136H	IGHV4-39	tgctgcggagaca	t	aaataatcat	gataattt	IGHD2-15
	PGT137H	IGHV4-39	tgctgtga	agcacaataatacatgacattgtca		tggttggt	IGHD1-14
	PGT141H	IGHV1-8	tgctacggagaggg	ctcaaaacatcgttttgctgagacatacgttctctacgatactacggcttaataataatatacaagagtgga		tgactaccttga	IGHD4-17
	PGT142H	IGHV1-8	tgctacggagaggg	ctcaaaacatcgtcttgctgagacatacgttctctacgatactacggcttaataataatacaagagtgga		tgactaccttga	IGHD4-17
	PGT143H	IGHV1-8	tgctacggagaggg	ttcaaaacatcgtcttgctgagacatacgttctctacgatactacggcttaataataatacaagagtgga		tgactaccttga	IGHD4-17
	PGT144H	IGHV1-8	tgctacggagaggg	ctcaaaacatcgtcttgctgagacatacgttctctacgatactacggcttaataataatacaagagtgga		tgactaccttga	IGHD4-17
	PGT145H	IGHV1-8	tgcttgaccgg	ctcaaaacatcgtcttgctgagattattttctgatacatgataatggcccccaagagtggtg		gggtgactac	IGHD4-17

^aPotential V_H replacement footprints are highlighted in *red* (no mismatch) or *purple* (with one mismatch).

Table 2

The enriched V_H replacement products encoding anti-HIV antibodies are highly mutated.

	Sequence ID	V _H Name	VH (nt)	Mutations (nt)	Mutation Rate (%)	Replacement (nt)	Silent (nt)	R/S ratio
Anti-gp41	EU794402	IGHV1-69	294	49	16.67	33	16	2.06
	EU794399	IGHV1-69	295	44	14.92	33	11	3.00
Anti-V3 loop	EU794433	IGHV3-7	293	35	11.95	27	8	3.38
	EU794439	IGHV4-59	288	32	11.11	19	13	1.46
	EU794437	IGHV5-51	294	20	6.80	19	1	19.00
	EU794436	IGHV1-18	295	36	12.20	28	8	3.50
	EU794422	IGHV4-59	291	24	8.25	19	5	3.80
	S77136	IGHV3-11	295	29	9.83	22	7	3.14
	EU794425	IGHV2-5	297	20	6.73	12	8	1.50
	EU794427	IGHV5-51	293	10	3.41	8	2	4.00
	EU794414	IGHV5-51	294	20	6.80	17	3	5.67
Anti-gp120	Z35118	IGHV2-26	299	11	3.68	9	2	4.50
	Z35126	IGHV4-31	298	11	3.69	4	7	0.57
	U82950	IGHV1-69	290	32	11.03	26	6	4.33
	AF086903	IGHV3-30	294	16	5.44	13	3	4.33
	U82767	IGHV1-69	292	45	15.41	32	13	2.46
	S67982	IGHV1-69	291	29	9.97	24	5	4.80
	S67984	IGHV2-5	292	8	2.74	4	4	1.00
	U82942	IGHV1-69	294	32	10.88	24	8	3.00
CD4i	17b	IGHV1-69	295	34	11.53	28	6	4.67
	c12	IGHV1-69	296	27	9.12	20	7	2.86
	23e	IGHV1-69	295	21	7.12	13	8	1.63
	M16	IGHV1-69	292	44	15.07	32	12	2.67
	SB1	IGHV1-69	292	24	8.22	16	8	2.00
	X5	IGHV1-69	296	27	9.12	21	6	3.50
	47e	IGHV1-69	295	9	3.05	7	2	3.50
	E51	IGHV1-69	293	27	9.22	20	7	2.86

	Sequence ID	V _H Name	VH (nt)	Mutations (nt)	Mutation Rate (%)	Replacement (nt)	Silent (nt)	R/S ratio
PGT	PGT121H	IGHV4-59	291	56	19.24	38	18	2.11
	PGT122H	IGHV4-61	295	56	18.98	37	19	1.95
	PGT123H	IGHV4-59	288	64	22.22	42	22	1.91
	PGT125H	IGHV4-b	277	106	38.27	97	9	10.78
	PGT126H	IGHV4-b	277	105	37.91	97	8	12.13
	PGT127H	IGHV4-61	280	99	35.36	91	8	11.38
	PGT128H	IGHV4-b	277	112	40.43	107	5	21.40
	PGT130H	IGHV4-39	297	63	21.21	49	14	3.50
	PGT131H	IGHV4-39	297	65	21.89	49	16	3.06
	PGT135H	IGHV4-34	219	59	26.94	47	12	3.92
	PGT136H	IGHV4-39	196	46	23.47	40	6	6.67
	PGT137H	IGHV4-39	196	56	28.57	47	9	5.22
	PGT141H	IGHV1-8	296	47	15.88	37	10	3.70
	PGT142H	IGHV1-8	296	48	16.22	38	10	3.80
	PGT143H	IGHV1-8	296	47	15.88	36	11	3.27
	PGT144H	IGHV1-8	296	51	17.23	42	9	4.67
	PGT145H	IGHV1-8	296	52	17.57	40	12	3.33