

# Autoreactivity and Exceptional CDR Plasticity (but Not Unusual Polyspecificity) Hinder Elicitation of the Anti-HIV Antibody 4E10

Kathryn A. K. Finton<sup>1</sup>, Kevin Larimore<sup>2</sup>, H. Benjamin Larman<sup>3</sup>, Della Friend<sup>1</sup>, Colin Correnti<sup>1</sup>, Peter B. Rupert<sup>1</sup>, Stephen J. Elledge<sup>3</sup>, Philip D. Greenberg<sup>2,4,5</sup>, Roland K. Strong<sup>1\*</sup>

**1** Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **2** Program in Immunology, Cancer Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **3** Department of Genetics, Harvard University Medical School, and Division of Genetics, Howard Hughes Medical Institute, Brigham and Women's Hospital, Boston, Massachusetts, United States of America, **4** Department of Immunology, University of Washington, Seattle, Washington, United States of America, **5** Department of Medicine, University of Washington, Seattle, Washington, United States of America

## Abstract

The broadly-neutralizing anti-HIV antibody 4E10 recognizes an epitope in the membrane-proximal external region of the HIV envelope protein gp41. Previous attempts to elicit 4E10 by vaccination with envelope-derived or reverse-engineered immunogens have failed. It was presumed that the ontogeny of 4E10-equivalent responses was blocked by inherent autoreactivity and exceptional polyreactivity. We generated 4E10 heavy-chain knock-in mice, which displayed significant B cell dysregulation, consistent with recognition of autoantigen/s by 4E10 and the presumption that tolerance mechanisms may hinder the elicitation of 4E10 or 4E10-equivalent responses. Previously proposed candidate 4E10 autoantigens include the mitochondrial lipid cardiolipin and a nuclear splicing factor, 3B3. However, using carefully-controlled assays, 4E10 bound only weakly to cardiolipin-containing liposomes, but also bound negatively-charged, non-cardiolipin-containing liposomes comparably poorly. 4E10/liposome binding was predominantly mediated by electrostatic interactions rather than presumed hydrophobic interactions. The crystal structure of 4E10 free of bound ligands showed a dramatic restructuring of the combining site, occluding the HIV epitope binding site and revealing profound flexibility, but creating an electropositive pocket consistent with non-specific binding of phospholipid headgroups. These results strongly suggested that antigens other than cardiolipin mediate 4E10 autoreactivity. Using a synthetic peptide library spanning the human proteome, we determined that 4E10 displays limited and focused, but unexceptional, polyspecificity. We also identified a novel autoepitope shared by three ER-resident inositol trisphosphate receptors, validated through binding studies and immunohistochemistry. Tissue staining with 4E10 demonstrated reactivity consistent with the type 1 inositol trisphosphate receptor as the most likely candidate autoantigen, but is inconsistent with splicing factor 3B3. These results demonstrate that 4E10 recognition of liposomes competes with MPER recognition and that HIV antigen and autoepitope recognition may be distinct enough to permit eliciting 4E10-like antibodies, evading autoimmunity through directed engineering. However, 4E10 combining site flexibility, exceptional for a highly-matured antibody, may preclude eliciting 4E10 by conventional immunization strategies.

**Citation:** Finton KAK, Larimore K, Larman HB, Friend D, Correnti C, et al. (2013) Autoreactivity and Exceptional CDR Plasticity (but Not Unusual Polyspecificity) Hinder Elicitation of the Anti-HIV Antibody 4E10. *PLoS Pathog* 9(9): e1003639. doi:10.1371/journal.ppat.1003639

**Editor:** Alexandra Trkola, University of Zurich, Switzerland

**Received:** April 12, 2013; **Accepted:** August 4, 2013; **Published:** September 26, 2013

**Copyright:** © 2013 Finton et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award number P01AI094419 (RKS) and the National Cancer Institute of the National Institutes of Health under award number R01CA033084 (PDG). This research was also conducted as part of the Collaboration for AIDS Vaccine Discovery with support to RKS and PDG from the Bill and Melinda Gates Foundation. KAKF was supported in part by Public Health Service National Research Service Award T32GM008268, from the National Institute of General Medical Sciences of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: rstrong@fhcrc.org

## Introduction

An effective prophylactic AIDS vaccine will need to generate anti-HIV neutralizing antibodies (Abs) that target the HIV envelope glycoprotein (Env) [1–3] and broadly neutralize as many HIV isolates as possible (bNAbs). The bNAb 4E10 [4–10] recognizes an epitope that is highly conserved across HIV-1, HIV-2, and SIV and displays one of the widest breadths of any anti-HIV bNAb, neutralizing 98% of HIV-1 strains [11,12]. These properties have made 4E10 an attractive vaccine target, but

previous attempts to elicit 4E10 or equivalent Abs through vaccination have failed.

The HIV envelope protein (Env) consists of gp120 surface subunits and gp41 membrane-anchoring subunits assembled as noncovalent trimers of gp120/gp41 heterodimers to form mature, functional ‘spikes’ on the virion surface. 4E10 recognizes a conserved linear epitope (consensus clade B sequence: <sup>671</sup>NWFDITNWLW<sup>680</sup>; core epitope: NWF<sup>D</sup>/<sub>N</sub>IT), immediately adjacent to the viral membrane in the gp41 membrane proximal external region (MPER) [4,10]. bNAbs targeting the MPER, such

## Author Summary

4E10 is an example of an anti-HIV, broadly neutralizing antibody that is uncommon in infected patients and has not been successfully elicited by any vaccine approach attempted. 4E10 has been proposed to neutralize HIV through a mechanism that requires broad recognition of other antigens, including membrane phospholipids. Such a mechanism would also block the generation of 4E10 during B cell development, confounding vaccination strategies. Analysis of B cell development in 4E10 heavy-chain knock-in mice confirmed that 4E10 does recognize self-antigens. However, a previously proposed autoantigen candidate, the mitochondrial lipid cardiolipin, was not consistent with binding studies which showed that while 4E10 does bind liposomes containing cardiolipin, it does so only weakly and nonspecifically, also binding liposomes without cardiolipin. Using a synthetic human peptidome, 4E10 was shown to be polyreactive, binding peptides from various proteins, but only in a limited manner. Three of the top five hits are from types 1, 2 and 3 inositol trisphosphate receptors, with high scoring peptides sharing a conserved sequence motif. Validation of the top hits was performed by binding analyses and staining of tissue sections, which combined to identify the type 1 inositol trisphosphate receptor as the most likely 4E10 physiological autoantigen.

as 4E10 and a second bNAb, 2F5, are uncommon in infected individuals, and Env-derived immunogens do not efficiently elicit these Abs in vaccinees [13,14]. 2F5 recognizes a linear epitope (consensus clade B sequence: <sup>662</sup>ELDKWA<sup>667</sup>) neighboring the 4E10 epitope in the MPER [15]. It has been proposed that these Abs are inherently polyspecific and autoreactive as a direct consequence of their epitope specificity and neutralization mechanism, which has been posited to require interactions with lipids and viral membrane components in addition to and outside of their peptide epitopes [16–20]. Therefore, B cell tolerance mechanisms would limit the natural production of MPER-specific bNAbs, like 2F5 and 4E10, in people infected with HIV and hinder the elicitation of equivalent bNAbs by vaccination. Structural elements of 4E10 that have been proposed to mediate lipid or liposome recognition include the long, hydrophobic heavy chain complementarity determining region 3 (HCDR3) sequence and, in particular, the side-chain of residue W100H, which protrudes from the tip of HCDR3 at a position predicted to penetrate the viral membrane when bound to Env, based on the 4E10/epitope peptide complex crystal structure [10] and mutagenesis studies [21,22].

Polyspecificity is the property of recognizing multiple distinct ligands (recognition degeneracy) combined with the specificity with which each of those ligands is recognized [23,24]. Polyspecificity is particularly important for the function of adaptive immunoreceptors (T cell receptors and B cell receptors (BCRs)), which are functionally enhanced by broadly recognizing foreign antigens. Autoreactivity is the potential downside of unchecked polyspecificity, although autoreactivity does play a role in B cell activation and regulation (reviewed in [25–27]). On the order of 50% of newly-rearranged BCRs are estimated to be autoreactive and some degree of self-reactivity is necessary to pass through positive selection during B cell development; self-recognition by the pre-BCR induces the expansion of precursor cells and broadens antibody diversity. Autoreactivity can be defined functionally as dysregulation of B cell development (*e. g.* the developmental arrest, loss of immature B cells to central tolerance

mechanisms and reduced numbers of residual splenic B cells with low surface IgM density observed in homozygous 2F5 V<sub>H</sub>DJ<sub>H</sub> knock-in mice [28]) or *in vitro* with binding assays or immunofluorescence (IF) staining [29].

*In vivo* experiments demonstrating functional 4E10 autoreactivity had not been reported when we started these studies. 2F5 and 4E10 were originally concluded to be polyspecific and autoreactive on the basis of binding assays against 11 purified lipidic and nuclear autoantigens [16,19]. 2F5 and 4E10 also both showed HEp-2 cell reactivity, exhibiting diffuse cytoplasmic and weaker nuclear staining patterns [19]. On the basis of these results, the 4E10 autoantigen was proposed to be the mitochondrial diphosphatidylglycerol lipid cardiolipin (CL) [30,31], though 4E10 also showed comparable cross-reactivity against every lipid tested, including phosphatidylserine (PS), phosphatidylcholine (PC), phosphatidylethanolamine (PE) and sphingomyelin (SM). However, detection of clinically-relevant anti-CL autoantibodies (ACLA) by simple binding assays with isolated antigens, as has been previously reported [16], is associated with high false positive rates, representing challenges in the laboratory characterization of sera and Abs and the clinical evaluation of autoantibody induced morbidity [32]. More quantitative assays have been contradictory, with an estimated equilibrium dissociation constant ( $K_D$ ) in the low micromolar range in one study [18] but weaker, unquantifiable binding in others [33,34]; 4E10 was reported to bind to model viral liposomes which did not contain CL with a sub-micromolar  $K_D$  [21], but showed no binding to 'bald', proteolytically-digested virus-like particles [35]. The disagreement of these results is due, in part, to employed methodologies, such as surface plasmon resonance (SPR), that can be difficult to reproduce and hard to evaluate in the absence of positive controls and when analyzing interactions with lipidic species [36].

Abs against phospholipids, including ACLA, are seen in primary antiphospholipid syndrome (APS) or in association with other autoimmune diseases, such systemic lupus erythematosus (SLE). ACLA in this context generally do not recognize CL in isolation, but as complexes with the phospholipid-binding serum protein  $\beta$ 2 glycoprotein I ( $\beta$ 2GPI) or other proteins with anti-coagulant activity. The presence of ACLA in these conditions can provoke serious complications, including arterial and venous thromboses and recurrent miscarriages. A second class of antiphospholipid antibodies can be generated after certain viral and bacterial infections, such as HIV, hepatitis B and C, syphilis and leprosy. ACLA are found in approximately 50% of people infected with HIV (*versus* ~2% of uninfected controls) and are strongly linked with the level of viral replication, the level of B cell activation and the level of MPER-specific Abs [37,38]; however, these Abs are usually not associated with APS manifestations and are not dependent on  $\beta$ 2GPI for binding [39]. If 4E10 CL autoreactivity is sufficiently strong to trigger B cell tolerance mechanisms, it might be expected that 4E10 would contribute to APS; however, 4E10 showed only weak activity in APS clinical assays [33] and passive infusion studies [40], similar to that of antiphospholipid antibodies elicited during infections rather than that of ACLA.

While polyreactive across phospholipids, subsequent experiments with larger antigen arrays (a microarray of 106 connective tissue disease-related autoantigens plus controls [41] or the UNiChip AV-400 panel of 400 bacterially-expressed human proteins [33]) concluded that anti-HIV bNAbs were either not exceptionally polyreactive, including 2F5, 4E10 and the anti-CD4 binding site antibody b12, or displayed only limited polyreactivity, including 2F5 and 4E10 [42,43]. The reactivity profile of 4E10 was, therefore, concluded to be fundamentally distinct from those

of pathogenic autoantibodies that would trigger B cell tolerance mechanisms [41]. However, a recent analysis of 4E10 binding to a commercially-available array (Invitrogen ProtoArray 5) of 9,000 intact human proteins, expressed in baculovirus as GST-fusion proteins, concluded that 4E10 displayed “exceptional polyreactivity” on the basis of very broad, but low-avidity binding [20].

In this report, we found that knocking-in the 4E10 heavy chain into murine B cells results in comparable levels of dysregulation in the B cell compartment as had been observed in 2F5 heavy chain knock-in mice, demonstrating significant inherent autoreactivity, which was confirmed by clear patterns of specific autoreactivity by immunohistochemistry (IHC) on murine tissue sections. We determined that 4E10 does not stain mitochondria in HEp-2 cells *in vitro* or in tissue sections and that 4E10 bound only relatively weakly, but comparably, to liposomes containing CL or virus-like compositions lacking CL in well-controlled SPR binding assays using improved methodologies [36]. Therefore, 4E10 phospholipid binding is relatively weak, not specific for CL compared to other phospholipids, and distinct from properties of true ACLA responses. In order to isolate the 4E10 structural elements that might mediate general phospholipid binding activity, binding studies were performed with engineered HCDR3-grafted ubiquitin (Ubq) fusion constructs, with 4E10 point mutants, and at varying ionic strengths, which showed that the weak phospholipid binding activity displayed by 4E10 is driven by electrostatic interactions, not hydrophobic ones. The overall positive charge of the 4E10 Fv cassette, a property of many anti-HIV bNAbs [44], does not account for this behavior, which is likely alternately driven by an electropositive pocket present in the dramatically restructured combining site of ligand-free 4E10 but absent in epitope bound-state structures. However, 4E10, in either conformational state (free or complexed), lacks structural features required for truly specific recognition of CL. We therefore conclude that CL is unlikely to be the physiologically-relevant self-antigen mediating 4E10 autoreactivity and deletion during B cell development.

Because 4E10 was clearly autoreactive in 4E10 knock-in mice, we sought to identify candidates for *bona fide* 4E10 autoantigens using a synthetic library of 413,611 36-mer peptides spanning the human proteome combined with phage immunoprecipitation sequencing (PhIP-Seq) [45]. The results identify a short list of protein candidates, validated by *in vitro* binding studies and consistent with IHC results. The PhIP-Seq results were also useful for more finely defining 4E10 polyspecificity beyond previous studies, leading to the conclusion that 4E10 displays limited, highly focused polyspecificity, unexceptional for this class of molecules. Our findings should inform future attempts to generate 4E10-equivalent functionalities by vaccination and determine the precise molecular mechanism of 4E10 neutralization of HIV.

## Results

### Impaired B cell development and function in 4E10 heavy chain (4E10H) knock-in mice

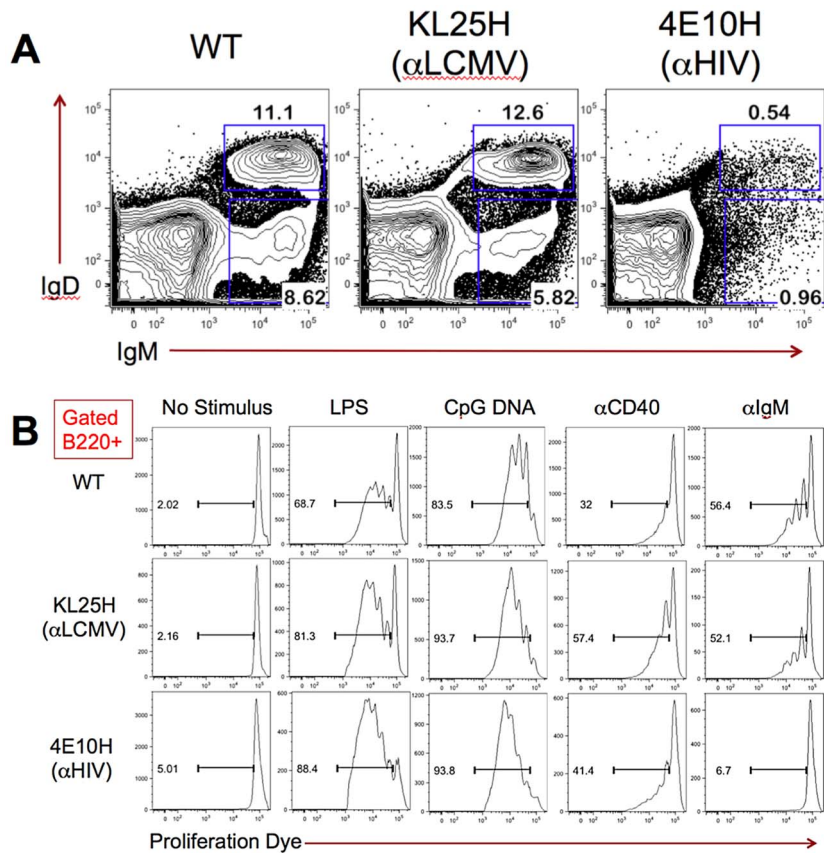
In order to evaluate the potential self-reactivity of 4E10, the rearranged human IgH VDJ antigen-binding domain of the 4E10 gene was targeted to the mouse immunoglobulin heavy chain (IgH) locus in B6 embryonic stem cells. B cell development in the derived 4E10H<sup>+/+</sup> knock-in mice was evaluated by flow cytometry and compared with both wild-type B6 and control KL25H<sup>+/+</sup> knock-in mice (Fig. 1A). KL25H<sup>+/+</sup> knock-in mice express from the mouse IgH locus the rearranged mouse IgH VDJ antigen-binding domain of the LCMV-neutralizing antibody [46]. Flow cytometric analysis of bone marrow-resident B220<sup>+</sup> cells from

4E10H<sup>+/+</sup> knock-in mice revealed a dramatic reduction of both B220<sup>+</sup>IgM<sup>+</sup>IgD<sup>-</sup> pre-B cells and B220<sup>+</sup>IgM<sup>+</sup>IgD<sup>+</sup> immature B cells when compared with B6 mice. In contrast, KL25H knock-in mice displayed normal populations of developing B cells. These results suggest that expression of the 4E10 Ig heavy chain led to deletion of the majority of developing B cells as a result of self-antigen recognition.

Although there was a profound defect in B cell development, low numbers of B220<sup>+</sup> B cells could still be found in the spleens and lymph nodes of 4E10H knock-in mice, suggesting that some 4E10H-expressing B cells had escaped deletion during development. To determine whether putative autoreactivity had impaired the function of 4E10H-expressing B cells, we tested the ability of these cells to respond to mitogenic and antigen receptor stimuli *in vitro*. Splenocytes were purified from 4E10H<sup>+/+</sup>, KL25H<sup>+/+</sup>, and WT B6 mice, loaded with cell proliferation dye eFlour670, and cultured overnight in the presence of LPS, CpG DNA, an activating anti-CD40 antibody, anti-IgM F(ab')<sub>2</sub>, or in culture media alone (no stimulus). The following day, B cell proliferation was assessed by flow cytometric measurement of dye dilution in B220<sup>+</sup> cells (Fig. 1B). B220<sup>+</sup> B cells from 4E10H, KL25H, and WT mice placed in culture media alone remained highly stained with proliferation dye, indicating they had not undergone cell division. 4E10H<sup>+/+</sup>, KL25H<sup>+/+</sup>, and WT B6 B cells diluted the dye after treatment with LPS, CpG DNA, or anti-CD40, indicating that these cells had undergone cell division. This suggested that 4E10H-expressing B cells were capable of responding to a variety of mitogenic signals and did not have a general cellular activation defect. However, while both KL25H<sup>+/+</sup> and B6 B cells proliferated in response to IgM cross-linking, 4E10H<sup>+/+</sup> B cells did not, suggesting that these cells were anergic to antigen receptor signaling, a phenotype consistent with self-antigen reactivity.

### 4E10 bound liposomes only weakly, but predominately through electrostatic interactions

Having confirmed that 4E10H was autoreactive in a murine background, we sought to characterize 4E10 interactions with CL in order to determine definitively whether the affinity and specificity for this phospholipid biochemically met criteria for autorecognition. We tested binding of 4E10 IgG to liposomes captured on SPR biosensors with three different compositions approximating (i) the external leaflet of cellular membranes (PC/cholesterol), (ii) membranes containing CL (PC/cholesterol/CL) or (iii) the HIV viral envelope (PC/SM/PE/PS/cholesterol [22]). We also incorporated improvements into these SPR protocols compared to previous analyses. Liposomes were prepared with biotinylated lipids (0.5% w/w) and captured on streptavidin-coated sensor chips (Biacore SA chips (GE)) [47] rather than adsorbing prepared liposomes to lipophilic groups covalently attached to carboxymethylated dextran-coated sensor chips (Biacore L1 chips (GE)). This approach reduces the need to block dextran lipid groups with either saturating amounts of liposomes or directly chip-coupled bovine serum albumin (BSA), improving signal-to-noise ratios, and has been determined to generate much more consistent and reliable results than L1 sensor chips [36] or ELISA-based methods [32,48]. BSA was added to all analyte buffers throughout analysis to act as a blocking reagent, preventing nonspecific binding. Annexin V, which binds to both CL and PS [37], was used as a positive control in order to provide comparative responses. Double subtraction [49] was also used to correct sensorgrams, which were run in duplicate, where both reference cell and buffer-matched blank responses are considered.



**Figure 1. Analysis of 4E10H knock-in mice.** (A) The percentages of IgM<sup>+</sup>IgD<sup>-</sup> pre-B and IgM<sup>+</sup>IgD<sup>+</sup> immature B cells within the B220<sup>+</sup> cell population in the bone marrow of B6 WT, B6-KL25H and B6-4E10H mice are shown. Marrow cells were isolated and stained for cell surface markers as described in the methods section. FACS plots were previously gated B220<sup>+</sup> and on live cell size based on forward and side scatter. Numbers represent the percentages of B220<sup>+</sup> bone marrow cells falling within each box gate. (B) The proliferation of B220<sup>+</sup> splenocytes from B6 WT, B6-KL25H, and B6-4E10H mice in response to overnight culture in the presence of B cell stimuli are shown. Cells were loaded with cell proliferation dye before overnight incubation, and proliferation was assessed by FACS analysis after cell surface staining to identify B220<sup>+</sup> B cells. Histograms were previously gated B220<sup>+</sup> and on live cell size based on forward and side scatter. Numbers represent the percentage of B220<sup>+</sup> cells falling within the bar gate on each plot, indicating dilution of the proliferation dye as a result of cell division. doi:10.1371/journal.ppat.1003639.g001

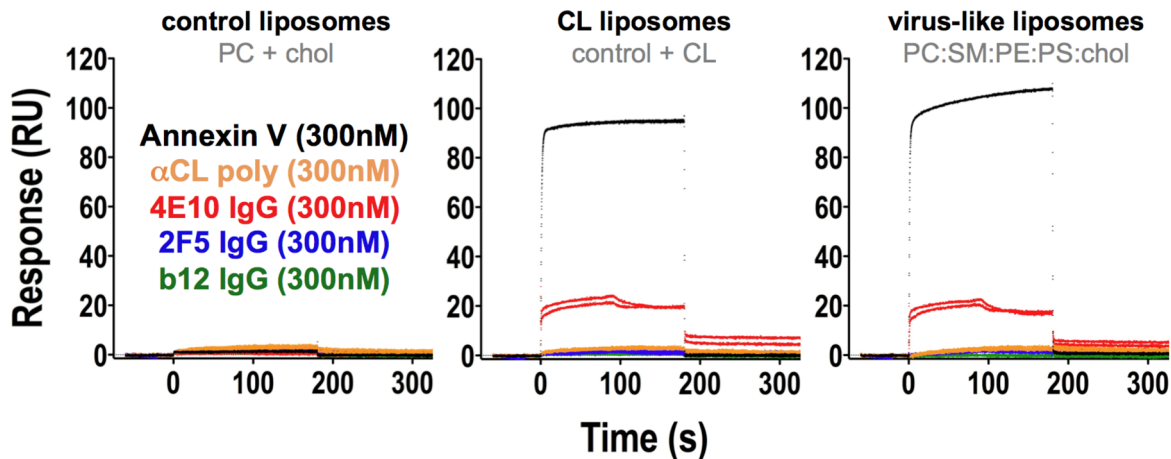
These combined improvements result in SPR sensorgrams that show cleanly negative results for the binding of Abs and Annexin V to control liposomes (PC/cholesterol) but clear, reproducible, positive results for Annexin V binding to liposomes incorporating CL or PS (Fig. 2). 2F5 and b12 monoclonal IgGs and a patient-derived ACLA polyclonal antiserum did not bind appreciably to liposomes of any composition tested. However, 4E10 IgG showed weak, but reproducible binding to CL liposomes containing PC, cholesterol and CL and virus-like liposomes containing PE, PC, PS, SM and cholesterol, but lacking CL. However, while detectable, the SPR binding responses, which reflect bivalent avidities, could not be quantified accurately, due to the weak responses and distinct, biphasic association profiles unfittable by any reasonable model. Nevertheless, 4E10 binding responses to CL and virus-like liposomes were qualitatively comparable, demonstrating a lack of specificity for CL over other negatively-charged phospholipids. 4E10 IgG, the bivalent form, was used in these experiments because univalent reagents, like scFvs [34], did not display detectable binding (*data not shown*), confirming the weakness of these interactions in absolute terms.

It had been hypothesized that the HCDR3 of 4E10 makes direct hydrophobic contacts with the viral membrane, facilitating binding through aromatic and apolar amino acid side-chains,

particularly W100H and W100bH [21]. In order to test whether or not 4E10 HCDR3 alone supports liposome binding, HCDR3-grafted Ubq fusions were designed (Fig. 3A). Ubq is a small, stable protein with an exposed loop (residues 46–47) bracketed by residues spaced by the same distance as the framework residues bracketing HCDR3 in Abs. The fusions were generated by replacing this two residue loop in Ubq with the HCDR3 sequences of 2F5, b12 or 4E10. The three Ubq fusion constructs, expressed in bacteria, were monodisperse in solution after purification, determined by size-exclusion chromatography (SEC; Fig. 3B), and showed detectable, but weak binding to chip-coupled HIV SF162 gp140 by SPR (Fig. 3C), confirming their utility for biochemical assays. However, all three HCDR3-grafted Ubq fusions failed to show any binding response by SPR to liposomes of any composition tested (Fig. 3D), even at relatively high analyte concentrations, showing that hydrophobic HCDR3 sequences alone, out of context of the rest of the folded Ab structure, are insufficient to confer even minimal binding to liposomes.

Biomolecular interactions are mediated by hydrophobic, electrostatic, van der Waals and hydrogen bonds. The widely-held presumption has been that 4E10 binds lipids, liposomes and membranes through predominately hydrophobic interactions mediated by HCDR3 and other combining site residues. In order

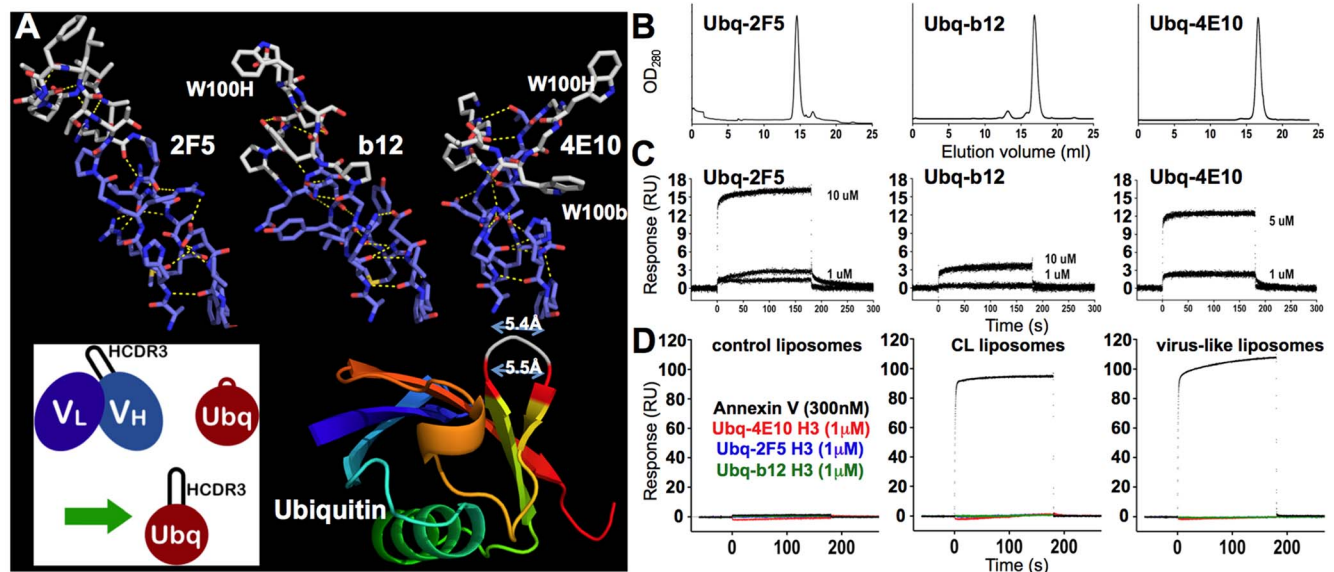




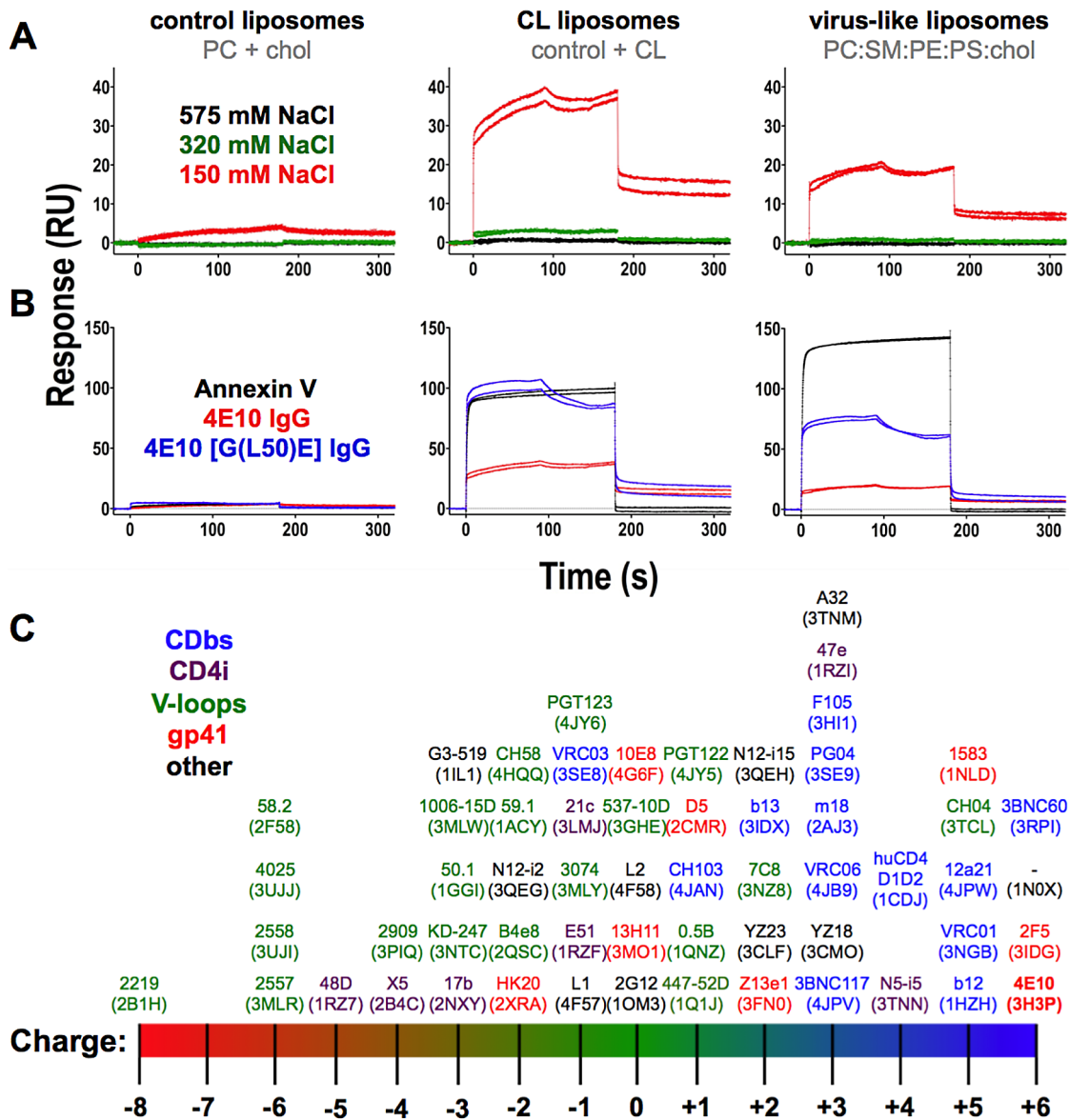
**Figure 2. SPR analyses of 4E10/liposome interactions.** Corrected SPR responses are shown (response units: RU; duplicate runs) of Annexin V (black),  $\alpha$ CL polyclonal Ab (orange), 4E10 IgG (red), 2F5 IgG (blue), and b12 IgG (green) analytes binding to liposomes incorporating biotinylated lipids captured on streptavidin-coated biosensor chips. All analyte concentrations were 300 nM. The Annexin V analyte buffer included 2.5 mM  $\text{CaCl}_2$ . Liposome compositions are indicated above each frame. doi:10.1371/journal.ppat.1003639.g002

to determine the overall character of 4E10/liposome binding, SPR analyses were conducted at a series of salt concentrations; increasing the salt concentration drives hydrophobic interactions

but ablates electrostatic ones, which is the fundamental principle behind ion exchange and hydrophobic interaction chromatography. SPR analyses showed that 4E10 IgG interactions with both



**Figure 3. Design and characterization of HCDR3-grafted Ubq fusion proteins.** (A) A ribbon representation of Ubq (1UBQ.pdb [106]) is shown at bottom right, colored prismatically from N- (blue) to C-terminus (red); the targeted loop for HCDR3 engraftment is highlighted in gray with red bounding residues. The structures of the three targeted HCDR3 segments in context of their parent Abs are shown above in a licorice-stick representation, colored by atom type (oxygen: red; nitrogen: blue; carbons of buried residues: marine blue; and carbons of exposed residues: gray). Residue exposure was determined with the program GetArea [107], with greater than 30% solvent accessibility considered “exposed”, based on the structural context in the corresponding crystal structures (2F5: 3IDG.pdb [108]; b12: 3RU8.pdb [104]; 4E10: 2FX7.pdb [9]). The three transferred sequences (2F5: **AHRRGPTTLFGVPIARGPVNAMDVW**; b12: **ARVGPYSWDDSPQNYNMDVW**; 4E10: **AREGTTGWGWLKPIGAFAHW**; exposed residues **bolded**), were characterized as “hydrophobic” with the Sigma-Aldrich PEPscreen Library Design Tool ( $\Phi_{2F5}=0.52$ ;  $\Phi_{b12}=0.51$ ;  $\Phi_{4E10}=0.56$ ). Key residues and  $\text{C}\alpha$ - $\text{C}\alpha$  distances ( $\text{\AA}$ ) are indicated. The inset shows a schematic representation of the design process. Molecular images were generated with MacPyMOL [109]. (B) SEC analyses of the solution properties of the purified recombinant HCDR3-grafted Ubq fusion proteins are shown, confirming monodispersity. (C) Corrected SPR responses (duplicate runs) of HCDR3-grafted Ubq fusion proteins to HIV SF162 gp140 protein amine-coupled to biosensor chips; analyte concentrations are indicated. While detectable, these responses are consistent with  $K_D$  values  $\gg 10 \mu\text{M}$  and are thus too weak to quantify reliably. (D) SPR responses (duplicate runs) of Annexin V and HCDR3-grafted Ubq fusion protein analytes to liposomes incorporating biotinylated lipids captured on streptavidin-coated biosensor chips are shown, with liposome compositions indicated above each frame. HCDR3-grafted Ubq fusion protein responses are colored as indicated, but none showed detectable binding on any liposome composition. doi:10.1371/journal.ppat.1003639.g003



**Figure 4. Analysis of the effects of salt concentration, mutation and overall cassette charge on 4E10/liposome interactions.** Corrected SPR responses are shown for Annexin V or 4E10 IgG analytes (300 nM; duplicate runs) to liposomes incorporating biotinylated lipids captured on streptavidin-coated biosensor chips; liposome compositions are indicated above each frame. **(A)** SPR responses of wild-type 4E10 IgG analytes at different salt concentrations are plotted. **(B)** SPR responses of Annexin V (300 nM), wild-type 4E10 IgG (300 nM) or the 4E10 [G(L50)E] mutant IgG (400 nM) are plotted. A higher concentration of 4E10 mutant IgG was used with the expectation that binding might be significantly reduced. Since this did not occur, mutant IgG responses appear elevated due to the concentration differential. **(C)** The net charge at neutral pH of the Fv cassettes of the anti-HIV Abs with structures currently available through the PDB [103] was calculated with the structure-based algorithm PDB2PQ [51–54]. Two Abs, PG16 and NIH45-46, were excluded because their structures included modified amino acids that could not be accommodated by PDB2PQ. Fvs are plotted with their names, with assigned PDB accession codes in parentheses. Ab labels are colored by the locale of their epitopes on Env, as indicated; 4E10 is also bolded.  
doi:10.1371/journal.ppat.1003639.g004

CL-containing and virus-like liposomes were abolished by increasing the salt concentration (**Fig. 4A**), demonstrating that 4E10 interacts with liposomes containing phospholipids with negatively-charged headgroups (*e.g.* CL) predominately through electrostatic interactions, not hydrophobic ones.

It had previously been suggested that 4E10 contains a phosphate-binding subsite [50] and that a structurally-reasonable phosphate binding site is present on the light chain side of the combining site, bracketed by the side-chains of K32L and K100eH [34]. In the latter study, a 4E10 [G(L50)E] IgG mutant

was constructed that places a glutamic acid side-chain directly between the K32L and K100eH side-chains, designed to disrupt phosphate binding. In order to identify features on 4E10 that could mediate phospholipid interactions outside of hydrophobic combining site residues, the [G(L50)E] mutant was tested for liposome binding by SPR (**Fig. 4B**). However, 4E10 [G(L50)E] IgG retained binding to liposomes containing CL and virus-like liposomes, demonstrating that mutational ablation of the K32L/K100eH site does not ablate phospholipid binding. Outside of the K32L/K100eH site, inspection of available 4E10 structures,

which were all determined as complexes with HIV epitope-related ligands, failed to reveal any obvious electropositive pocket, crevice or groove that might be reasonably inferred to confer CL binding, so no further targeted mutagenesis studies were performed. Structure-based calculations of the net charges at neutral pH of the Fv cassettes of all available anti-HIV antibody structures using PDB2PQ [51–54], showed that 4E10 and 2F5 have exceptional net positive charges (**Fig. 4C**). However, since this level of overall electropositivity of 2F5 is insufficient to confer liposome binding, it is also unlikely to account for 4E10 liposome binding.

### The crystal structure of unbound 4E10 revealed a completely restructured combining site

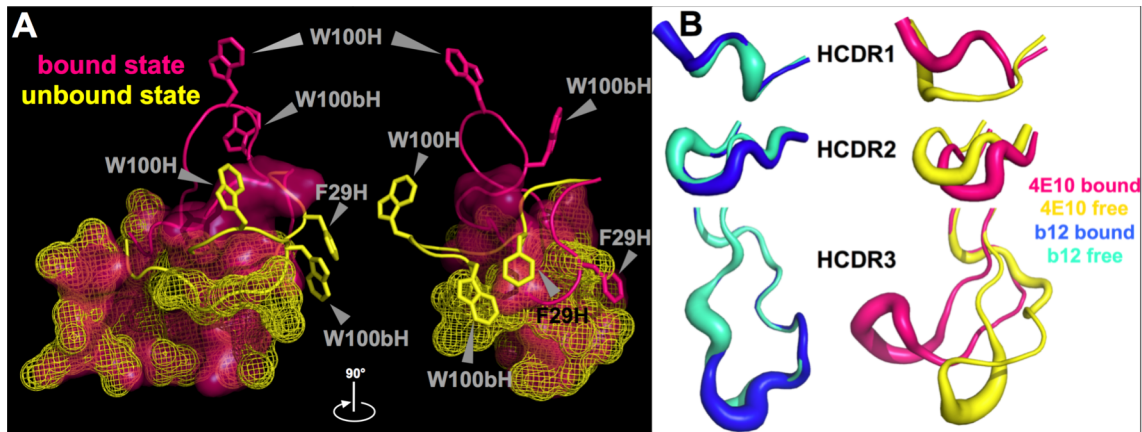
The species actually binding lipids in our studies was 4E10 free of bound ligand, where all previous 4E10 structures were determined as complexes with MPER epitope-related ligands. While highly affinity-matured Abs tend to have rigidified combining sites [55–67], we sought to crystallize free 4E10 to identify potential structural flexibility that might explain these discrepancies in the absence of alternative hypotheses. The 4E10 Fv in the unbound state was crystallized from  $\text{Li}_2\text{SO}_4$

( $d_{\text{min}} = 2.4 \text{ \AA}$ ), preliminary phases were determined by molecular replacement, and the structure was rebuilt and refined with good agreement statistics and geometry (**Table 1**). The asymmetric unit contains four copies of 4E10, grouped as two pairs related by a single NCS two-fold axis, arranged in a pinwheel-like pattern with the HCDR3s forming the central spoke.  $V_L$  domains of Fvs from one member of each dyad-related pair are relatively poorly ordered, likely due to sparse crystal contacts. Overall, the Fv domains show a high degree of structural conservation to each other, with the main differences arising from alternate HCDR1 and HCDR3 conformations between dyad-related pairs. The HCDR3 conformation of one dyad-related pair is extended into a solvent channel and unconstrained by crystal packing, with large accompanying crystallographic B-factors, likely closely recapitulating the free, solution-state structure (**Fig. 5A**). The HCDR3 conformation from the second dyad-related pair is alternatively crunched by extensive crystal packing and hydrophobic interactions across the plane bisecting the dyad axis, resulting in low B-factors stemming from these steric constraints. Comparison of the presumed solution-state, unbound 4E10 conformer, using the better-ordered dyad mate, with a reference epitope-peptide 4E10

**Table 1.** Crystallographic data collection and refinement statistics.

	Synchrotron	In-house
<b>Data collection*</b>		
Space group	C222 <sub>1</sub>	C222 <sub>1</sub>
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	88.89, 161.6, 163.3	88.89, 161.6, 163.3
Resolution (Å)	46.1–2.39 (2.44–2.39)	50.0–2.35 (2.43–2.35)
<i>R</i> <sub>merge</sub>	0.072 (0.382)	0.066 (0.438)
<i>I</i> / <i>σ</i> <i>I</i>	28.45 (5.6)	22.97 (2.66)
Completeness (%)	100.0 (99.8)	99.9 (99.9)
Redundancy	8.3 (3.0)	4.1 (2.5)
<b>Refinement</b>		
No. reflections (unique)	46,684 (2,305)	49,263 (4,863)
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.216/0.245	–
No. non-hydrogen atoms		
Protein	6,452	–
Ligand/ion	81	–
Water	296	–
B-factors (Å <sup>2</sup> )		
Protein	42.48	–
Ligand/ion	40.99	–
Water	26.12	–
R.m.s. deviations		
Bond lengths (Å)	0.006	–
Bond angles (°)	0.899	–
Ramachandran plot statistics (MolProbity)		
Residues in most favored regions (%)	97.6%	–
Residues in disallowed regions (%)	0.0%	–
MolProbity score	1.37	–
Est. coordinate error (max. likelihood ESU; Å)	0.161	–
Accession code	4LLV	–

\*Values in parentheses are for the highest-resolution shell.  
doi:10.1371/journal.ppat.1003639.t001



**Figure 5. Structure of 4E10 free of bound antigen.** (A) Superposition of the three HCDRs from bound and unbound b12 (3RU8.pdb [104], 1HZH.pdb [110] and 4E10 (2FX7.pdb [9], 4LLV.pdb) are shown in a B-factor putty representation (b12, blue: bound; cyan: unbound. 4E10, pink: bound, yellow: unbound). (B) Superposition of residues from the 4E10 epitope binding site and HCDR1 and 3 from bound (semi-transparent molecular surface in pink; 2FX7.pdb [9]) and unbound (mesh molecular surface in yellow; 4LLV.pdb) 4E10 Fv are shown; HCDR1 and 3 are shown in cartoon representations with side-chains of key residues shown in licorice-stick representations and labeled. The molecule is oriented so that the  $V_L$  domains are to the left and the  $V_H$  domains to the right. Molecular images were generated with MacPyMOL [109].  
doi:10.1371/journal.ppat.1003639.g005

Fab complex structure (2FX7.pdb [9]), showed that the largest structural differences are also between HCDR1 and 3, with movements of  $\sim 12$  Å between the C $\alpha$  positions of W100H in HCDR3 and  $\sim 6$  Å between the C $\alpha$  positions of F29H in HCDR1 (Fig. 5A and B; pair-wise superpositions resulted in a root mean square deviations (RMSD) on all common C $\alpha$ s of 0.76 Å for the heavy chain, with HCDR1 and 3 deleted, and 0.32 Å for the light chain). The movements of HCDR1 and 3 reposition the side-chains of F29H and W100bH, which are both buried in the  $V_H$  domain in epitope complexes, to form hydrophobic interactions partially covering the side-chains of L53H, L54H and I56H in HCDR2. These three side-chains are exposed in the epitope complex structures, making contacts with a hydrophobic face of the MPER epitope. In the complex structures, HCDR3 forms the back wall of the epitope binding pocket. In the unbound structure, however, HCDR displacements remodel the 4E10 combining site, completely occluding epitope binding but creating a shallow concavity on the backside of HCDR3, a surface typically buried in Ab structures. One of the ordered sulfates modeled into the structure sits near the middle of this concavity, coordinated by the guanidinium group of R94H and main-chain and side-chain contacts from S28H.

#### 4E10 did not stain HEp-2 cells by IF microscopy

4E10 had been reported to show diffuse cytoplasmic and nuclear staining of HEp-2 cells [19,20], suggesting the presence of a cross-reacting autoantigen. Higher resolution confocal microscopy of HEp-2 cells stained with 4E10 IgG was used to expand the prior result by narrowing the subcellular localization and begin the process of identifying the antigen/s. Fixed and permeabilized HEp-2 cells were incubated with 4E10 or, as comparative controls, b12 or an anti-protein disulfide isomerase ( $\alpha$ PDI) IgG (Fig. 6). b12 has been reported to stain HEp-2 cells in a similar pattern, though with stronger nucleolar staining in the nucleus, but this Ab has not been reported to be autoreactive [16], and  $\alpha$ PDI marks the endoplasmic reticulum (ER). If 4E10 can bind CL with a physiologically relevant affinity, 4E10 would be predicted to stain cells with a mitochondrial pattern. The  $\alpha$ PDI IgG stains HEp-2 cells in the expected ER pattern (Fig. 6A), but 4E10 and b12 bind

irreproducibly in replicate experiments and failed to show any specific staining (Figs. 6B and C), negating this approach.

#### Analysis of 4E10 autoreactivity and polyspecificity using a synthetic human peptidome

A phage-displayed library of 413,611 overlapping 36-mer peptides spanning the complete human proteome [45] was used to identify candidate autoantigens and assess the polyspecificity of 4E10; b12 was also analyzed for comparison (Figs. 7 and 8). In this library, neighboring peptides in contiguous open reading frames overlap by seven residues, the approximate size of a linear Ab epitope. Both Abs were analyzed in duplicate, with results plotted as replicate #1 versus replicate #2  $-\log_{10}$   $P$ -values [45]. Plotted in this way, highly discordant replicate pairs lie near the axes and were discarded from further consideration; 241 36-mers were also culled because of nonspecific binding. Based on sequence searches, no sequence motifs associated with the core epitope of 4E10 (NWF<sup>D</sup>/N<sup>IT</sup>) or alternate proposed HIV epitopes [68] are present in the human proteome, so positive binding results in these assays reflect truly polyspecific recognition of distinct epitopes. b12 showed results consistent with a non-autoreactive Ab of limited polyspecificity (Fig. 7A), with the highest scoring peptide having a replicate average  $-\log_{10}$   $P$ -value equal to 14.3 and with approximately two dozen peptides rising above the main cluster (Figs. 7A and 8). Detailed inspection of the 15 top-scoring b12 36-mer peptides (Fig. 8) showed a consistent pattern of highly positively-charged peptides (average predicted net charge at neutral pH of +9) with blocks of low sequence complexity in 12 of 15 peptides.

The 4E10 PhIP-Seq results were distinct from the b12 results (Figs. 7B and 8). Using the highest scoring b12 peptide as a benchmark under the assumption that b12 is not autoreactive, 17 36-mer peptides showed significantly higher scores (replicate average  $-\log_{10}$   $P$ -values of 18.4 to 255). The chemical character of these 17 4E10 peptides was distinct, as well: overall moderately hydrophobic or hydrophobic (though none were very hydrophobic), with a range of predicted net charges at neutral pH from  $-5.7$  to  $+1.9$  and greater sequence complexity. None of the 17 high-scoring peptides had sequences related to HIV, as expected,





**Figure 6. IF microscopy of HEp-2 cells.** Fixed and permeabilized HEp-2 cells were stained with (A)  $\alpha$ PDI IgG (an ER marker), (B) 4E10 IgG, or (C) b12 IgG, appropriate fluorescently-labeled secondary Abs, and visualized microscopically. Image parameters were carefully matched across experiments.

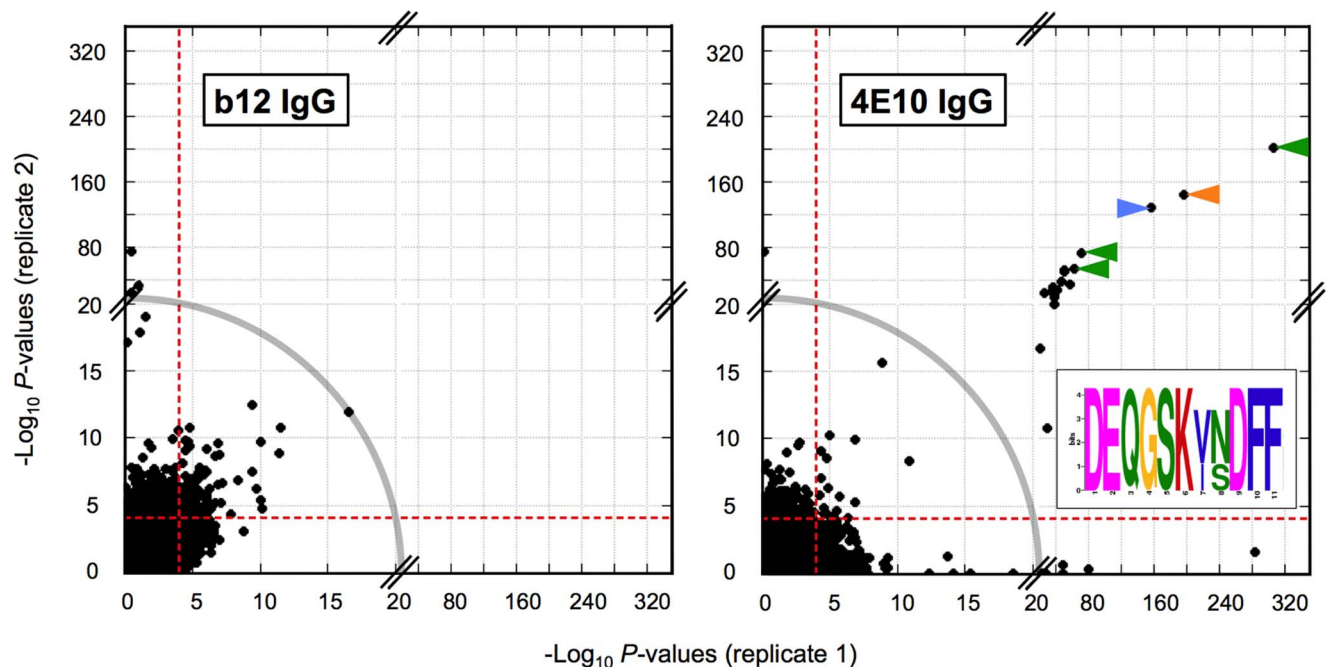
doi:10.1371/journal.ppat.1003639.g006

but we were able to identify a conserved motif. Three of the five top-scoring peptides, including the highest scoring peptide, are derived from an ER-resident inositol trisphosphate receptor (IP<sub>3</sub>R; NP\_002214 [69,70] [71]. MEME motif analysis [72] revealed a highly conserved core motif (DEQGSK<sup>V</sup>/<sub>1</sub><sup>N</sup>/<sub>5</sub>DFF) within these three peptides quite unrelated to the 4E10 MPER epitope (Fig. 7B inset). While no homologous crystal structure is available, Phyre<sup>2</sup> threading [73] predicts that this sequence forms part of an extended helix in IP<sub>3</sub>R. The other two top-five scoring 36-mers were from cytoplasmic dynein 1 intermediate chain 2 (NP\_001369 [69,70]) and the complement receptor type 1 isoform S precursor (NP\_000642 [69,70]). None of the peptides from the top two scoring proteins from the previously reported ProtoArray 5

analysis [20] (retinoid X receptor  $\beta$  and interleukin-1 family member 6) or the proposed autoantigen, splicing factor 3B3 (SF3B3; ProtoArray 5 rank not reported), had average  $-\text{Log}_{10} P$ -values above 1.54 in the PhIP-Seq analysis (4E10 library average = 0.27 ( $\sigma = 0.82$ ); b12 library average = 0.33 ( $\sigma = 0.45$ )). The close concordance in library scoring behavior also argues against the presence of very broad, low-avidity binding by 4E10 relative to b12.

#### Validation of PhIP-Seq top-scoring peptides by SPR

The five top scoring peptides identified by PhIP-Seq (Fig. 8) were further validated by testing for binding to 4E10 directly by SPR. Peptides were expressed in a eukaryotic secretion-pathway



**Figure 7. PhIP-Seq analyses of 4E10 and b12.** PhIP-Seq results are plotted as  $-\text{Log}_{10} P$  values, one replicate on the abscissa, the other on the ordinate; note the discontinuity in axis scales. The top five scoring 4E10 peptides are highlighted with arrows; the three peptides derived from IP<sub>3</sub>R isoform sequences are highlighted with green arrows, with the consensus core sequence shown as a MEME logo [72,111] in the inset. The orange arrow corresponds to a peptide derived from cytoplasmic dynein 1 and the blue arrow to a peptide derived from complement receptor type 1. The gray arc corresponds to the  $-\text{Log}_{10} P$  values observed for the highest-scoring b12 peptide, so defines a threshold for potential autoreactive binding events. Proximity to the diagonal indicates good replicate concordance; peptides with highly discordant replicate values, falling along the axes, outside the bounding dashed red lines were discarded from the analysis.

doi:10.1371/journal.ppat.1003639.g007

Rank	Avg. -logP	Peptide sequence	Parent protein	$\Phi$	Charge
<b>b12 IgG</b>			<i>average:</i>	<b>0.1 ± 0.7</b>	<b>+9 ± 4</b>
1	14.3	YKVDV <b>TWTRARGASRGWRSR</b> HQLKGRNGWRNSRVYK	TP53-target gene 5 protein	<b>0.12</b>	<b>+10.1</b>
2	11.1	<b>KKAKVAKPKKAAKSAAKAVKPKAAKPKVVKPKKAA</b> P	histone H1.2	<b>-0.08</b>	<b>+15</b>
3	11.0	<b>KSPQTWPRRTRPRSPPEAAPSGVGRG</b> STWTRRRDTPR	metaxin-1 isoform 1	<b>0.02</b>	<b>+8</b>
4	10.2	NEEGKEN <b>RDRDRDYSRRRGGPPRRGRGASRGREFRG</b>	ubiquitin associated protein 2-like	<b>-0.34</b>	<b>+6</b>
5	9.89	AVLKYENNVMNIRQFNCSPPHYWLPNFMDFVFTWVSLP	serine/threonine-protein phosphatase 2B catalytic subunit $\alpha$ isoform 1	<b>0.81</b>	<b>-0 -</b>
6	8.41	<b>SSSSQSRSRSSSRERSRSGSKSRSSSRSHRGSSS</b>	zinc finger Ran-binding domain-containing protein 2 isoform 2	<b>-0.32</b>	<b>+9.1</b>
7	8.26	AMAGRIYISGMAPRPSLAKKQRFHRNRKGYRSQRG	vitronectin precursor	<b>0.09</b>	<b>+11.1</b>
8	7.96	RTATQRKIFQCNKHM <b>KVFKYSNRNKVRHTKKKTFK</b>	zinc finger protein 729	<b>0.10</b>	<b>+13.2</b>
9	7.84	PK <b>GEPTRRGRGGTFRRGGDRDPGGRPSR</b> PSTLRRPAY	protein PRRC2C	<b>-0.08</b>	<b>+9</b>
10	7.80	<b>HRACSTYRTIYRTAYRRSPGLAPARPRYACCP</b> GWKR	epidermal growth factor-like protein 7 precursor	<b>0.29</b>	<b>+9</b>
11	7.68	QHPFYNRPFAPYLYTPRYYPGGSQLISRPSVKTS	kinase D-interacting substrate of 220 kDa	<b>0.50</b>	<b>+4.2</b>
12	7.68	AKYN <b>PRKHRAKRHPRRPPRSPGMEPPFSHRCFP</b> RYI	telomerase protein component 1	<b>0.16</b>	<b>+10.2</b>
13	7.63	<b>AKKVKTQPQKAAKSPAKAKAPKPKAAKPKSGKPKV</b>	histone H1.3	<b>-0.11</b>	<b>+14</b>
14	7.61	RRKAPTCKSPLSGVSQEGKQV <b>QRNAANARE</b> RARMRV	transcription factor 21	<b>-0.06</b>	<b>+9</b>
15	7.49	<b>SPPKSLFSQRPQRMRSRFSRHRSCSRSPYSR</b> SR	peroxisome proliferator-activated receptor gamma coactivator 1 $\alpha$	<b>0</b>	<b>+11</b>
<b>4E10 IgG</b>			<i>average:</i>	<b>0.4 ± 0.1</b>	<b>-1 ± 2</b>
1*	255	SKCRVFNTTERDEQGSKVNDFFQQTEDLYNEMKWQK	<b>inositol 1,4,5-trisphosphate receptor type 2</b>	<b>0.17</b>	<b>-1</b>
2*	170	PPHELTEEEKQQLHSEEFSLFFDHSTRIVERALSE	cytoplasmic dynein 1 intermediate chain 2	<b>0.28</b>	<b>-5.7</b>
3*	143	PPRCISTNKCTAPEVENAIRVPGNRSFFSLTEIVR	complement receptor type 1 isoform S precursor	<b>0.48</b>	<b>+1.9</b>
4*	72.5	TERDEQGSKINDFFLRSEDLFNEMNWQKLRAPVL	<b>inositol 1,4,5-trisphosphate receptor type 1 isoform 2</b>	<b>0.28</b>	<b>-1</b>
5*	58.0	LTEETKHRLFTTTEQDEQGSKVSDFFDQSSFLHNEM	<b>inositol 1,4,5-trisphosphate receptor type 3</b>	<b>0.21</b>	<b>-4.8</b>
6	51.4	TFGHLGRLREL <b>LSLRNNALSALS</b> GDIFAASPALYRLD	TLR4 interactor with leucine rich repeats precursor	<b>0.48</b>	<b>+1.1</b>
7	50.5	ASSRHQRIQDFNYTDHTLGRIILNAMNETNFFGVTG	$\gamma$ -aminobutyric acid type B receptor subunit 2 precursor	<b>0.45</b>	<b>+0.2</b>
8*	46.0	GAQPPFDAQSPLDSPQPSGQPWNFHASTSWYWRQS	nuclear fragile X mental retardation-interacting protein 1	<b>0.42</b>	<b>-0.9</b>
9	42.3	IITVNACSQSDQLNPEPGENSISEEEYSKNWFTVSK	transient receptor potential cation channel subfamily M member 6 isoform a	<b>0.39</b>	<b>-4</b>
10*	35.3	ESSQHTLNYWKEQSLNVSQDLDIRSNINFFEFK	nesprin-2 isoform 1	<b>0.41</b>	<b>-1.9</b>
11	34.1	RLLCSAATDGLAFWDLTTMLDHDSTVLEPPVDPG	WD repeat-containing protein 6	<b>0.55</b>	<b>-5</b>
12	33.9	NGHINFKKFWEISRQIHEFMTWTQVECPFEKDKKIQ	ras-GEF domain-containing family member 1A	<b>0.43</b>	<b>+1.1</b>
13	31.8	VYEYKYGPCANHVEPSENCLKGYVNNLSFFDLSEL	anoctamin-3	<b>0.55</b>	<b>-3</b>
14	29.3	PGYENDSVEDLKEVTSISSRKRGRRYFWEYSEQLT	UPF0474 protein C5orf41 isoform 1	<b>0.12</b>	<b>-0 -</b>
15	25.4	AAEPVRAKLDRQRRVFQPSPLASQFELPGDFFNLTA	UBX domain-containing protein 6 isoform 1	<b>0.36</b>	<b>+1</b>
16	20.2	QFNTLYLRVRGDGRPVMVNIKEDTDFQRTNQMYSY	complex I intermediate-associated protein 30, mitochondrial precursor	<b>0.42</b>	<b>+1</b>
17	18.4	TTGEYKSTFNGNRPSSDRHLIPVAFVSEKWFEISC	adherens junction-associated protein 1	<b>0.41</b>	<b>-0 -</b>

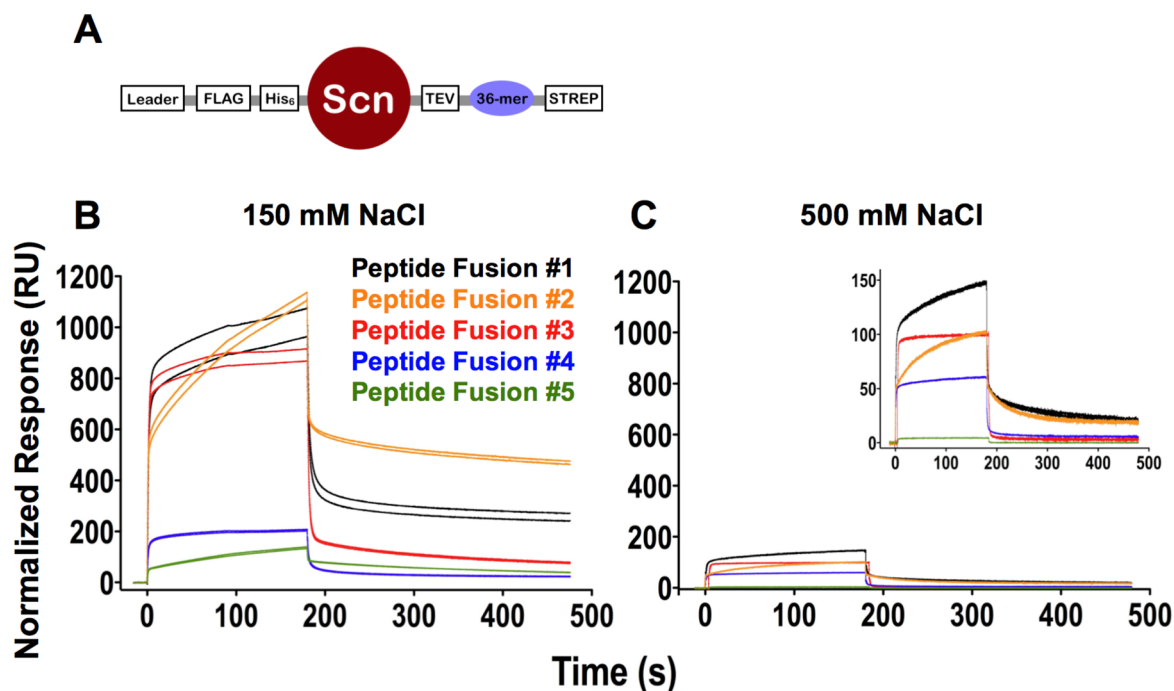
**Figure 8. Top scoring 4E10 and b12 PhIP-Seq 36-mer peptides.** The top fifteen scoring 36-mer peptides from the b12 IgG PhIP-Seq analysis (above) and the seventeen top scoring 36-mer peptides from the 4E10 IgG analysis exceeding the threshold set by the highest-scoring b12 peptide (below) are shown, compiled by their relative rank order, replicate-averaged PhIP-Seq score, sequence, identity of the parent protein, hydrophobicity ( $\Phi$ ) and predicted charge at neutral pH. Hydrophobicities were determined with the Sigma-Aldrich PEPscreen Library Design Tool (green: hydrophilic; orange: moderately hydrophobic; red: hydrophobic; no peptides were predicted to be “very hydrophobic”); overall charge was determined with the Innovagen Peptide Property Calculator (red: negative; green: neutral; blue: positive). Segments of low sequence complexity (defined by the program SEG [112]) are highlighted in red; the core motif in the IP<sub>3</sub>R isoform peptides (labeled in bold), identified by MEME [72], is underlined. Asterisks indicate peptide sequences selected for expression as fusion constructs for subsequent validation.  
doi:10.1371/journal.ppat.1003639.g008

expression system as fusion constructs with the human protein siderocalin (Scn; **Fig. 9A**) [74]; expression in bacterial systems resulted in unacceptable levels of proteolytic degradation (*data not shown*). The constructs incorporated various purification and epitope-tags, including a C-terminal Strep-tag [75]. Western blots (*data not shown*) and SPR analyses with an anti-Strep-tag Ab confirmed the presence of full length peptides, with specific activities assayed by SPR of 12% to 34% as coupled, which allowed for normalizing results between SPR experiments (**Figs. 9B and C**). All five peptides showed definitively positive binding results (**Fig. 9B**), whereas no binding was seen to the siderocalin construct minus peptide, confirming specific interactions. Binding was reduced by 75% to 85%, but not ablated, in the presence of 500 mM NaCl (**Fig. 9C**), indicating that the interactions were not exclusively electrostatic in nature and drawing a distinction with 4E10 recognition of liposomes. The qualitative, normalized SPR responses monotonically decreased from peptide #1 to #5 to fairly weak values, so no further peptide fusions from hits farther down the PhIP-Seq list were evaluated.

#### IHC showed specific patterns of 4E10 staining

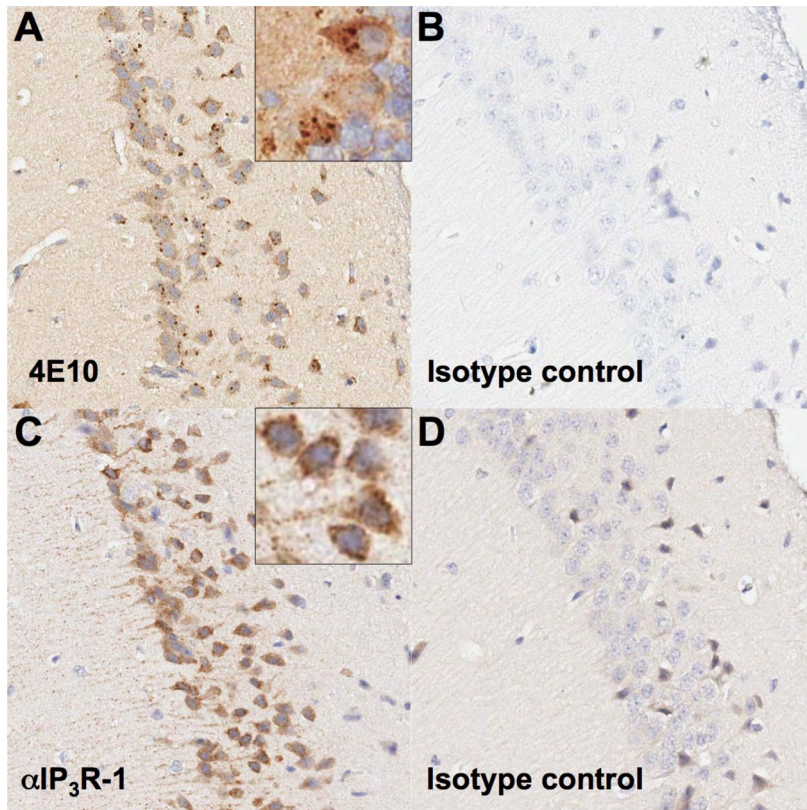
Previous IHC studies of the three isoforms of IP<sub>3</sub>R [76] show that IP<sub>3</sub>R1 is markedly enriched in Purkinje neurons in the

cerebellum, IP<sub>3</sub>R2 is overwhelmingly found in glial cells and IP<sub>3</sub>R3 is predominantly found in neurons, but not glia, and is enriched in neuropil, especially in neuronal terminals in limbic and basal forebrain regions. To further validate IP<sub>3</sub>R as the candidate 4E10 autoantigen, serial sections from fixed, paraffin-embedded mouse cerebellum were deparaffinized, rehydrated and treated to retrieve antigens, followed by staining with either 4E10 or an IHC-validated anti-IP<sub>3</sub>R1 rabbit polyclonal Ab (LifeSpan Biosciences LS-B2627) and appropriate secondary Abs. The commercial anti-IP<sub>3</sub>R1 polyclonal was raised against a synthetic peptide corresponding to C-terminal residues of human IP<sub>3</sub>R1, cross-reacts with mouse and rat orthologs and strongly stains Purkinje cells per the manufacturer’s specifications (LifeSpan Biosciences). 4E10 strongly stained specific nodular substructures in a subset of neurons, identified as Purkinje cells on the basis of morphology, with minimal background staining of other cell types or substructures in these sections (**Fig. 10A and B**). The anti-IP<sub>3</sub>R1 polyclonal strongly stained a corresponding subset of neurons, also identified as Purkinje cells, and stained a combination of similar nodular structures as well as exhibiting more diffuse cytoplasmic and cellular process staining (**Fig. 10C and D**). SF3B3, together with splicing factor 3a and a 12S RNA, forms the U2 small nuclear ribonucleoprotein complex (U2 snRNP), which



**Figure 9. SPR analyses of Scn-peptide fusions.** (A) The generalized siderocalin fusion construct used for PhIP-Seq peptide expression is shown. (B) Corrected and normalized SPR responses for 4E10 IgG binding to the top five peptides immobilized on a CM5 biosensor chip are plotted. (C) Corrected and normalized SPR responses of 4E10 IgG (300 nM; duplicate runs) binding to peptides (colored as in (B)) at 500 mM NaCl are plotted; inset shows magnified view. Normalized SPR responses are plotted in arbitrary units.  
doi:10.1371/journal.ppat.1003639.g009





**Figure 10. IHC of serial mouse cerebellar sections.** (A) Immunohistochemical staining with 4E10 IgG in mouse cerebellum sections reveals perinuclear, nodular immunoreactivity in Purkinje neurons; inset shows a magnified view. (B) Human IgG Isotype control staining of mouse cerebellum is shown. (C) Staining for IP<sub>3</sub>R1 with an anti-IP<sub>3</sub>R1 IgG shows a perinuclear, nodular and diffuse cytoplasmic staining pattern in Purkinje neurons, mirroring that seen in (A); inset shows a magnified view. (D) Rabbit IgG isotype control staining of mouse cerebellum is shown. doi:10.1371/journal.ppat.1003639.g010

is localized to the nucleus [77]. The staining patterns of both 4E10 and the anti-IP<sub>3</sub>R1 polyclonal were inconsistent with recognition of antigens widely-expressed across cell types, like CL or SF3B3, or exclusively nuclear antigens, like SF3B3.

## Discussion

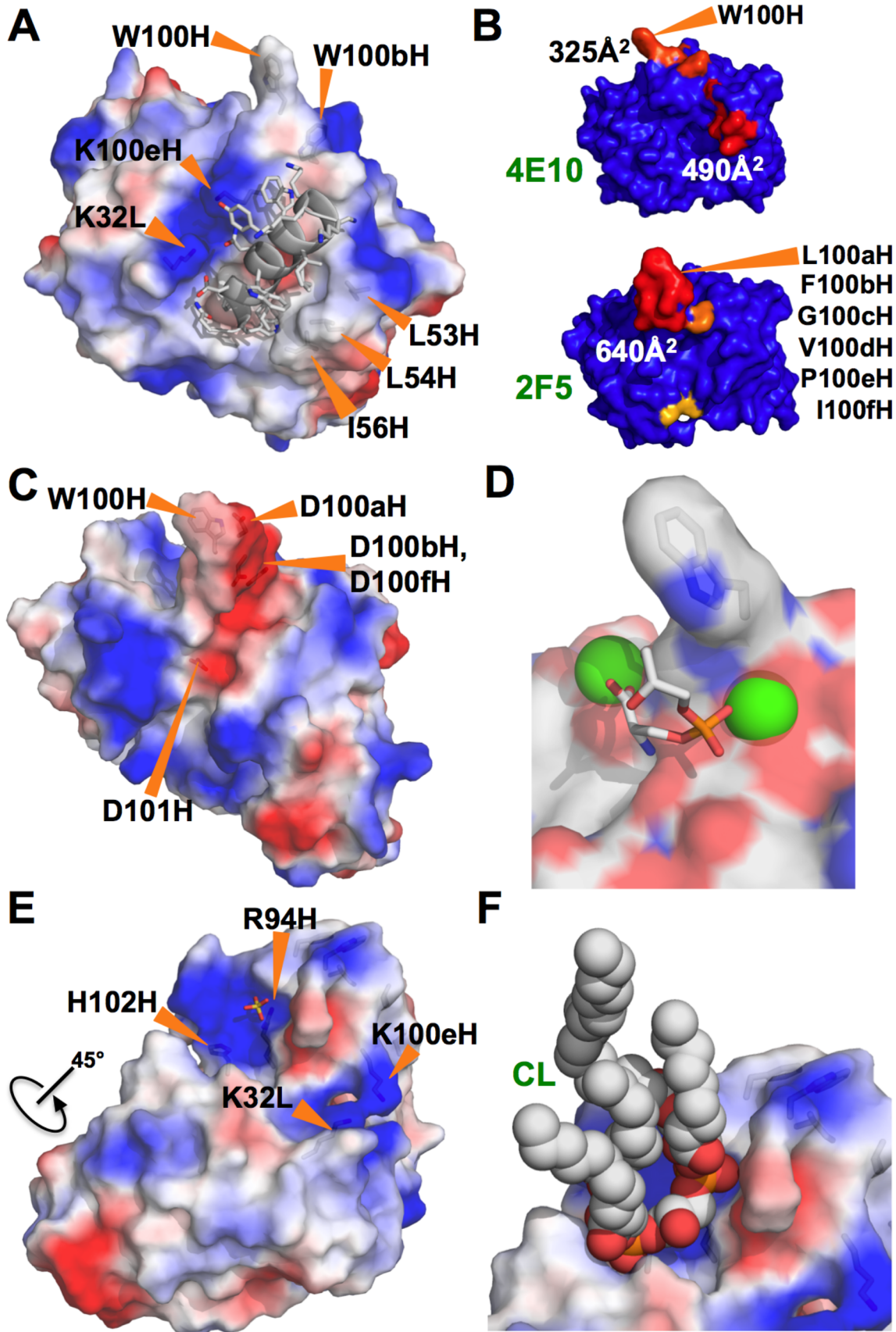
When applying “reverse vaccinology” [78] to the design of an HIV vaccine, it is important to determine whether or not the targeted bNAbs, *e.g.* 4E10, are autoreactive, and, if so, whether autoantigen recognition is separable from antigen recognition and neutralization. Despite usage of a breadth of murine light chains, 4E10H knock-in mice exhibited a profound blockade during B cell development consistent with the induction of central tolerance mechanisms, in which immature B cells expressing an autoreactive BCR are deleted (Fig. 1). In addition, the small residual population of B cells exported to the periphery, while able to respond to a variety of mitogenic signals, was unable to proliferate in response to IgM cross-linking. This phenotype substantiates the hypothesis that 4E10 is autoreactive. The questions then arise as to what autoantigen/s are mediating deletion and anergy and whether the mechanisms of 4E10 recognition of HIV and autoantigen epitopes are interdependent or separable, permitting the design of immunogens that generate Abs functionally equivalent to 4E10 with focused specificities that evade central and peripheral tolerance mechanisms.

Refined SPR analyses of Ab/liposome interactions (Fig. 2) confirmed that 4E10 binds to liposomes with compositions that

include negatively-charged lipids (CL, PS), but not net neutrally-charged liposomes that contain only PC and cholesterol; by contrast, 2F5, b12 and ACLA show no binding to liposomes of any composition tested. The failure of ACLA to bind to CL-containing liposomes was likely accounted for by the absence, in these experiments, of required protein binding partners, like  $\beta$ 2GPI, and draws a clear distinction between 4E10 binding behavior and autoimmune ACLA responses. The bivalent avidity of 4E10 for negatively-charged liposomes was weak compared to the univalent affinity of Annexin V binding. Crystal structures of 4E10/epitope peptide complexes [8–10,34] limited how 4E10 might interact with lipids, liposomes or membranes (Fig. 11A). The MPER epitope peptide, structured as a short  $\alpha$ -helix with a turn at the N-terminus, fits into a deep pocket in the Ab combining site, the back side of which consists of HCDR3. Two small hydrophobic patches are found centered on the side-chains of L53H, L54H and I56H, packing against a complementary hydrophobic face on the structured epitope peptide in complex structures, and W100H, which extends outwards from the tip of HCDR3. However, neither patch is particularly large or exceptional for Ab combining sites, as can be seen in comparison with 2F5 (Fig. 11B). Since 2F5 did not display liposome binding (Fig. 2), hydrophobic patches of this size order are not likely sufficient to account for 4E10 lipid binding.

The exposed side-chain of W100H on an extended HCDR3 structure, a feature which is shared with b12 (Fig. 11C), is positioned to potentially interact with membranes, but the side-chain of W100bH is partially buried behind the heavy chain N-





**Figure 11. Structures of 4E10, b12, 2F5, the CL-binding yeast cytochrome bc1 complex and Annexin V.** (A) The structure of the Fv domain of 4E10 (from 2FX7.pdb [9]) is shown in a semi-transparent molecular surface representation colored by electrostatic potential (blue: positive; red: negative). Side-chains of key residues are shown in licorice-stick representations and labeled. In this orientation, the V<sub>L</sub> domain is in the upper left and the V<sub>H</sub> domain is in the lower right. The backbone of the epitope peptide co-crystallized in this structure is shown in a cartoon representation, with side-chains shown in licorice-stick representations colored by atom-type (carbon: gray; oxygen: red; nitrogen: blue). (B) The molecular surfaces of the Fv domains of 4E10 (top) and 2F5 (bottom) are shown, oriented with V<sub>L</sub> domains at left and the V<sub>H</sub> domains at right. The surface is colored to show hydrophobic patches, defined by the program HotPatch [113]; patches are colored in descending order of total area (red, orange, yellow ...). Key residues and hydrophobic patch areas are indicated. (C) The structure of the Fv domain of b12 (from 3RU8.pdb [104]) is shown as in (A). (D) Annexin V (1A8A.pdb [114]) is shown as a semi-transparent molecular surface colored by underlying atom type, highlighting the exposed tryptophan side-chain (shown in licorice-stick representation). Calcium ions, shown as green spheres, coordinate the phosphate moiety from PS (shown in a licorice-stick representation, colored by atom-type as in (A) plus phosphorus in orange). (E) The structure of the unbound form of 4E10 (4LLV.pdb) is shown as a semi-transparent molecular surface representation colored as in (A). Side-chains of key residues are shown in licorice-stick representations and labeled. The view of the molecule has been rotated roughly 45° from the orientation of 4E10 shown in (A). The coordinated SO<sub>4</sub> ion is shown in a licorice-stick representation colored by atom type as in (A) plus sulfur in yellow. (F) CL is shown docked onto the unbound structure of 4E10 oriented and colored as in (E). Molecular images were generated with MacPyMOL [109]. doi:10.1371/journal.ppat.1003639.g011

terminus and the backbone of HCDR3 (**Fig. 11A**), rendering it inaccessible for docking to an encountered lipid bilayer. However, a sole exposed tryptophan side-chain on an extended HCDR3 is not sufficient on its own to confer liposome binding, otherwise b12 would also have displayed binding (**Fig. 2**), as it shares this feature with 4E10. We note that Annexin V also has a sole exposed tryptophan side-chain (**Fig. 11D**) which also does not support liposome binding in the absence of CL, PS or calcium (**Fig. 2**). Tryptophan side-chains can partition into lipid phases, but typically influence the precise positioning of integral membrane proteins rather than confer membrane binding [79–81], consistent with our results. The sequence of the 4E10 HCDR3 segment is quite hydrophobic overall, but is mostly buried in the core of the protein (**Fig. 3A**). HCDR3 peptides from 2F5, b12 or 4E10, even when fully exposed as grafted Ubq-fusion constructs, did not measurably bind to liposomes of any composition tested, showing that hydrophobic, tryptophan-rich HCDR3 sequences are also insufficient to confer lipid binding out of context of an intact Ab.

Many examples of crystal structures showing specific CL recognition are available (*e. g.* cytochrome bc1 complexes, like 1P84.pdb [82]). In these structures, the acyl chains of CL are typically found wedged in hydrophobic grooves or packed against extensive hydrophobic faces, while the phosphoglycerol-phosphate head group of CL makes polar contacts with positively-charged residues on the protein, with CL bound in a deeply invaginated, highly shape-complementary site. However, SPR analyses of 4E10 binding to liposomes in response to increasing salt concentrations showed that the binding was dominated by electrostatic interactions (**Fig. 4A**), drawing a clear distinction with examples of specific CL recognition in protein crystal structures, which are dominated by apolar and van der Waals interactions within a subset of specific, polar bonds. The previously available 4E10 crystal structures, determined as complexes with epitope ligands, did reveal a potential phosphate binding site, between the side-chains of K32L and K100eH, albeit in the absence of a crevice or groove complementary to a full CL headgroup. However, mutational ablation of this site did not reduce binding to negatively-charged liposomes (**Fig. 4B**). No other candidate phosphate or CL binding sites were apparent from inspection of the complex crystal structures. Annexin V binds PS via a pair of metal coordination sites interacting with negatively-charged phosphate and carboxylate groups (**Fig. 11D**), hence membrane binding is calcium-dependent, but 4E10 liposome binding is not (**Fig. 2**), suggesting a different recognition mechanism; 4E10 is also not known to bind metal ions and does not possess recognizable metal binding sites.

The substantial reordering of the 4E10 combining site in the unbound state, however, uncovers an electropositive pocket large enough to accommodate various phospholipid headgroups,

including the diphosphoglycerol headgroup of CL (**Fig. 11E**). An ordered sulfate ion occupies one end of this pocket in the crystal structure, designating a possible phosphate-binding subsite (SO<sub>4</sub><sup>2-</sup> and PO<sub>4</sub><sup>2-</sup> ions are chemically and structurally similar and can interact interchangeably within a protein binding site [83,84]). CL coordinates were extracted from 1P84.pdb and modeled into this site by rigidly docking one of the CL phosphate moieties onto the position of the ordered sulfate ion and manually adjusting for optimal fit, which was easily achieved (**Fig. 11F**). This pocket readily accommodates the CL headgroup without clashes, with the second phosphate moiety making reasonable interactions with the side-chain of H102H, approximately 7 Å away from the sulfate site, and the main-chain nitrogens of Q1H and V2H. Docking CL in this manner orients the lipid tails outwards from the Ab, sterically accommodating docking to a negatively-charged liposome surface. While fitting CL well, the pocket does not display close shape complementarity, permitting other phospholipid headgroups to bind, consistent with the observed recognition degeneracy and weak absolute affinities. Because the residues involved in inferred CL contacts in this site all make important contributions to the overall structure of 4E10 in the complex structures, cleanly interpretable mutagenesis studies could not be performed.

We conclude that 4E10 is not specific for CL; rather, 4E10 weakly and non-specifically interacts with phospholipid headgroups or liposomes carrying a net negative charge primarily through electrostatic interactions associated with an electropositive pocket on the unbound conformer of 4E10. As displayed by the absence of binding to control liposomes, 4E10 does not show a general proclivity to indiscriminately bind liposomes, consistent with this conclusion. Such weak, non-specific, electrostatic binding of phospholipids is unlikely to be relevant to B cell tolerance induction nor is likely a requirement for neutralization, consistent with prior studies that showed reducing liposome affinity did not ablate neutralization [34], but may contribute generally to targeting of viral surfaces. However, since CL recognition is likely mediated by a pocket on the 4E10 combining site present only in the unbound form of 4E10, CL binding would directly compete with MPER epitope binding. In other words, 4E10 must release weakly bound CL before or concurrent with adopting the CDR conformation required to bind its MPER epitope and, therefore, cannot use membrane associations as a guide to direct epitope binding. The requirement to undergo a large conformational shift is consistent with the relatively slow association kinetics and weak overall affinity ( $K_D = 135$  nM) observed when binding to Env proteins [34]; sampling an ensemble of CDR conformers may also contribute to the unusual biphasic association kinetics observed by SPR when binding liposomes.

Discarding CL from consideration, we utilized an unbiased proteomic approach to identify novel autoantigen candidates and to more fully characterize the polyreactivity of 4E10. A synthetic representation of the complete human proteome, composed of overlapping 36 amino acid peptides, was displayed on T7 phage and coupled with parallel DNA sequencing (PhIP-Seq) [45]. The PhIP-Seq method offers many improvements to prior methods, including the breadth of coverage and its quantitative nature [45]. 36-mers are long enough to retain substantial secondary and residual tertiary structure from the intact, folded, parent protein, potentially allowing identification of conformational epitopes. However, 36-mers are unlikely to represent stably folded species, and these peptides likely sample large conformer ensembles, including many non-native (non-self) states, increasing their utility for assaying polyspecificity. Since the peptides span a wide range of chemical properties (hydrophobic *vs.* hydrophilic, aromatic *vs.* charged *vs.* zwitterionic), the total conformer space also likely encompasses many peptide mimetics of non-protein species. A caveat with this approach is the difficulty in identifying discontinuous epitopes spanning more than one peptide, though linear segments of discontinuous epitopes may retain detectable binding. However, epitopes inaccessible in folded proteins, but which are exposed by proteolytic cleavage or other antigen processing events during B cell selection, can potentially be identified.

241 peptides were culled from the analysis because of uniform binding across a larger array of distinct Abs, but additional peptides retained in the analysis showed strong binding to more than one Ab in broader analyses not reported here, so binding results do not necessarily reflect just the specificity or polyspecificity of the Ab tested, but may also reflect the polyspecificity of a peptide. b12, an Ab that is presumably not autoreactive or unusually polyreactive, was analyzed to provide a baseline for PhIP-Seq results. The dozen top-scoring b12 peptides rising above the bulk response (**Fig. 8**) showed a common pattern of high positive charge and low sequence complexity, consistent with describing these peptides as poorly-ordered globs of positivity. The structure of b12 (**Fig. 11C**) is consistent with this observed, weak polyspecificity. Four aspartic acids adjacent to W100H, clustered on one face of the extended HCDR3 hairpin, form a large negatively-charged patch ideal for binding such peptides, at least weakly, as observed, and reveals a previously unrecognized aspect of b12 binding. The absence of extreme enrichments is consistent with the presumption that b12 is not self-reactive.

4E10, on the other hand, showed a strikingly different result (**Figs. 7B and 8**), with highly-significant enrichment for 17 peptides well beyond the  $10^{-15}$  *P*-value threshold set by b12, the top five with *P*-values ranging from  $10^{-58}$  to  $10^{-255}$ , consistent with functional autoreactivity and limited, focused polyspecificity quite distinct from b12. There is no evidence that 4E10 displays an exceptional, low-avidity polyreactivity from these results, especially in comparison with b12. The 17 top-scoring 4E10 peptides, constituting novel candidate autoantigens, were moderately hydrophobic to hydrophobic in nature overall, with a range of net charge, consistent with ordered parts of folded proteins. The top-scoring 4E10 peptide is derived from IP<sub>3</sub>R2, with additional hits from two related IP<sub>3</sub>R types (IP<sub>3</sub>R1 and IP<sub>3</sub>R3) within the top-scoring five peptides (**Fig. 8**), providing evidence that one, two or all three IP<sub>3</sub>R types constitute candidate 4E10 autoantigens. None of the peptides from the top two scoring proteins from the ProtoArray 5 analysis [20] ranked in the top 232,000 peptides; the highest scoring peptide from the alternate proposed 4E10 autoantigen SF3B3 had a replicate-averaged  $-\text{Log}_{10}$  *P*-value of 1.54, ranking it at 143,927. We also note that of the top 17 4E10 PhIP-Seq peptides, only four, numbers 11 (WD repeat-containing

protein 6), 12 (ras-GEF domain-containing family member 1A), 15 (UBX domain-containing protein 6 isoform 1) and 16 (complex I intermediate-associated protein 30, mitochondrial precursor), are from proteins represented on the ProtoArray 5, potentially accounting for the discordance in results between these two approaches. Scoring in PhIP-Seq also reflects internally-controlled peptide binding and not binding relative to a reference antibody, as in [20], which is necessarily affected by the binding profile of the reference antibody.

Three types of IP<sub>3</sub>R were identified as top-scoring PhIP-Seq hits and validated by SPR. In neuronal cell lines, IP<sub>3</sub>R1 and IP<sub>3</sub>R3 are expressed predominantly in the ER and perinuclear regions, while IP<sub>3</sub>R2 is expressed primarily in the nucleus [85]. IP<sub>3</sub>R3 shows an evenly diffuse cytoplasmic staining pattern, but the IP<sub>3</sub>R1 cytoplasmic pattern is more punctate, with strongly-staining nodular structures apparent [85]. However, failure to demonstrate binding of 4E10 to HEp-2 cells (which are not known to highly express IP<sub>3</sub>R) prevented comparative immunofluorescence analyses. We therefore sought to determine if the PhIP-Seq results identifying IP<sub>3</sub>R1, IP<sub>3</sub>R2 and IP<sub>3</sub>R3 as top candidate autoantigens were consistent with 4E10 recognition in a tissue-specific context by IHC. 4E10 staining of mouse cerebellar sections showed perinuclear staining of nodular structures specifically in Purkinje neurons, mirroring the staining of an anti-IP<sub>3</sub>R1 control polyclonal antibody. The 4E10 staining pattern observed was most consistent in terms of cell specificity with IP<sub>3</sub>R1 IHC patterns reported in previous studies [76], strongly arguing that IP<sub>3</sub>R1 is the lead candidate for a 4E10 autoantigen. The 4E10 staining pattern observed was also inconsistent with recognition of SF3B3. SF3B3, together with splicing factor 3a and a 12S RNA, forms the U2 small nuclear ribonucleoprotein complex (U2 snRNP) which is exclusively localized to the nucleus and widely-expressed across cell types [77].

Confirmation that 4E10 is autoreactive helps to identify the reasons vaccination strategies have failed, as tolerance mechanisms would limit the generation of anti-HIV Abs equivalent to 4E10. Our findings have important implications for future strategies to induce MPER-specific bNAbs. While lipid cross-reactivity mediated by HCDR3 had been suggested to play a role both in inducing B cell tolerance and in increasing neutralization potency through interactions with viral membranes, mutual functionalities precluding an engineered solution, our results alternately suggest that viral membrane binding is so weak and nonspecific, and competes with HIV epitope binding, that it is unlikely to be involved in the induction of tolerance mechanisms or to play a key role in neutralization. Weak but detectable binding of HCDR3 ubiquitin fusions to gp140s suggests that HCDR3 may make interactions to Env, rather than membrane, outside of the core 4E10 epitope. Neutralizing anti-influenza Abs (*e.g.* Abs CR6261, F10, FI6 and C179) often interact with a hydrophobic groove in the membrane proximal stalk region of hemagglutinin (HA) delineated by helix A and HA2 through hydrophobic HCDR residues [86–90]. The hydrophobic HCDR3 of 2F5, which recognizes an MPER motif neighboring the 4E10 epitope, has been argued to make similar contacts in the MPER region of Env as part of its neutralization mechanism [91]. We speculate that 4E10 may make analogous interactions to those seen in anti-HA Abs and 2F5, where 4E10 HCDR flexibility and hydrophobicity enable interactions with the presumed helical bundle of HIV gp41. However, the structure of the unbound form of 4E10 suggests that the long HCDR3 of 4E10 may also serve a role to shield hydrophobic patches in the combining site important for MPER recognition, prior to Env binding.

The fundamental concern in eliciting 4E10-like bNAbs was that viral epitope recognition would not be separable from autoantigen recognition. However, we showed that the viral epitope is not mimicking a self-antigen, because 4E10-selected sequences in the PhIP-Seq array analysis completely diverged from the 4E10 core epitope. The alternate core motif epitope from the likeliest candidate autoantigens, the three types of IP<sub>3</sub>R, was completely distinct from the 4E10 core epitope. Therefore, because the molecular mechanisms employed by 4E10 to recognize such divergent sequences are likely also distinct, even if they involve partially overlapping sites, it may be possible to ablate recognition of self-reactive epitopes while preserving interactions with neutralizing MPER epitopes. Future studies should concentrate on determining the complex crystal structure of 4E10 bound to the IP<sub>3</sub>R epitope in order to guide efforts towards engineering an immunogen that will elicit a non-autoreactive 4E10-like bNAb. However, if functional ontogeny of 4E10-equivalent responses requires engineering the ability to sample multiple, distinct conformers, in order to shield hydrophobic surfaces necessary for functional MPER recognition or to enable other molecular mechanisms necessary for recognition and/or neutralization, and if conventional Ab maturation typically proceeds by structural stabilization of the lone conformer optimal for epitope binding (*e.g.* [67]), it may not be possible to recapitulate 4E10-equivalent activities by any conventional immunization strategy.

## Materials and Methods

### Generation and analysis of 4E10H knock-in mice

4E10H targeted transgenic (knock-in) mice were generated by targeted insertion of a DNA construct encoding a mouse immunoglobulin heavy chain (IgH) promoter upstream of the rearranged Ab 4E10 IgH VDJ antigen-binding domain plus an mRNA splice donor sequence. The transgene was targeted to the mouse IgH genomic locus upstream of the immunoglobulin constant domains utilizing homologous recombination in C57BL/6 (B6) strain mouse embryonic stem (ES) cells. Transgene targeting was confirmed in ES cell clones by Southern blot (*data not shown*). 4E10H knock-in B6 ES cells were microinjected into FVB strain blastocysts, which were implanted into pseudo-pregnant female mice. Chimeric knock-in pups were identified by coat color and PCR genotyping, and crossed with B6 mice to generate pure B6 strain 4E10H knock-in mice. Germline transmission of the targeted 4E10H allele to progeny was confirmed by PCR genotyping (*data not shown*). Knock-in mice were inter-bred to establish homozygosity of the knock-in allele as assessed by PCR genotyping (*data not shown*). B6 background TgH(KL25) knock-in mice (referred to here as KL25H mice), expressing the mouse LCMV-neutralizing antibody KL25 rearranged IgH VDJ antigen-binding domain from the mouse IgH locus [46], were generously provided by Daniel Pinschewer (University of Geneva).

In order to analyze the composition of the bone marrow in transgenic animals, age-matched female wild type B6, KL25H knock-in, and 4E10H knock-in mice 6 to 8 weeks of age were euthanized and leg bones were isolated by dissection. Bone marrow was flushed from the leg bones with cold B cell media (RPMI/10% FBS/Penicillin/Streptomycin/2-mercaptoethanol) using a syringe and 18 g needle. Red blood cells were lysed with ACK lysing buffer (Gibco) and live cells were purified by density gradient centrifugation using Histopaque 1119 (Sigma-Aldrich).  $5 \times 10^6$  live cells from each mouse were stained with anti-mouse B220-Alexa647, IgM-PE, IgD-V450, and CD4-FITC (BD Biosciences), and analyzed on a FACS Canto II flow cytometer (BD Biosciences; **Fig. 1A**).

In order to analyze the ability of knock-in B cells to proliferate, splenocytes were purified from B6, KL25H, and 4E10H mice by density gradient centrifugation using Histopaque 1083 (Sigma-Aldrich). Purified cells were loaded with cell proliferation dye eFluor670 (eBioscience) according to the manufacturer's instructions.  $1 \times 10^6$  cells per well were incubated overnight in round-bottom 96-well plates with or without B cell stimuli (LPS, CpG DNA, an activatory anti-CD40 antibody, or anti-IgM F(ab')<sub>2</sub>) in B cell media, in a 37°C CO<sub>2</sub> cell culture incubator. The following day, cells were stained for cell surface markers (B220 and Thy1.2) and B cell proliferation was assessed by flow cytometry (**Fig. 1B**).

### Ethics statement

The activities involving the use of vertebrate animals described in herein have been reviewed and approved by the Fred Hutchinson Cancer Research Center (FHCR) Institutional Animal Care and Use Committee (IACUC), file #1695. FHCR has an approved Animal Welfare Assurance on file with the Office of Laboratory Animal Welfare (OLAW), Assurance number A3226-01.

### Liposome preparation

Lipids were purchased from Avanti Polar Lipids: bovine heart CL, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (PE), egg L- $\alpha$ -phosphatidylcholine or 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (PC), 1-stearoyl-2-oleoyl-sn-glycero-3-phosphoserine (PS), N-palmitoyl-D-erythro-sphingosylphosphorylcholine (SM), ovine wool cholesterol, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-biotinyl (biotinylated PE). Control liposomes (4:1 w/w PC:cholesterol), CL liposomes (1:7:2 w/w CL:PC:cholesterol), and viral mimic liposomes (3.2:5.4:5.1:4.8:10.2 w/w cholesterol:PC:SM:PE:PS) [22], each containing 0.5% w/w biotinylated PE, were prepared by mixing each lipid component in chloroform, evaporating the solvent, and then resuspending at a final concentration of 1 mg/mL in phosphate buffered saline (PBS). In control and CL liposomes, PC was included as egg L- $\alpha$ -phosphatidylcholine, while in viral mimic liposomes, PC was included as 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine. Lipid solutions were incubated at 65°C for 5 min and then subjected to five freeze/thaw cycles alternating between a dry ice ethanol bath for 3 min and a 65°C water bath for 3 min with intermittent vortexing. Solutions were then passed through an extruder (Avanti) with a 0.1  $\mu$ m polycarbonate membrane (Avanti) 21 times.

### Protein expression and purification

HCDR3-grafted Ubq fusion constructs (**Fig. 3**) with C-terminal His<sub>6</sub>-purification tags were engineered, cytoplasmically expressed in *E. coli* BL21(DE3) RIL cells (Invitrogen), and purified employing the methodology described in [92]. 4E10 [G(L50E)] IgG was produced as described in [34]. Seven of the top ten scoring 4E10 binding peptides identified in the PhIP-Seq analyses were cloned as fusion constructs with the human protein Scn [74] (**Fig. 9A**) using direct synthesis of human codon-optimized DNA (IDT) and Gibson cloning [93,94] (GeneArt Seamless Cloning kit, Invitrogen) and expressed in HEK293 Freestyle cells (Invitrogen) using the Daedalus lentiviral transduction system [74]. Scn-peptide fusion constructs were purified by SEC from concentrated culture supernatants and validated by PAGE and Western analyses as described in [74] (*data not shown*).

### SPR interaction analyses

All SPR experiments were performed at 25°C on a Biacore T100 instrument with the T200 sensitivity enhancement (GE



Healthcare). Liposome experiments (**Figs. 2, 3D** and **4**) were run in 10 mM HEPES, pH 7.4, 150 mM NaCl and 0.5 mg/mL BSA (HBS-N-BSA) on SA sensor chips (GE Healthcare). Liposome preparations containing 0.5% biotinylated lipids were injected at 50 to 100  $\mu\text{g/mL}$ , 5  $\mu\text{L/min}$ , until approximately 1000 SPR response units (RUs) were captured, then allowed to equilibrate for four hours at 10  $\mu\text{L/min}$ . Flow cell 1 was left blank as a reference surface. Single concentrations of 4E10 IgG (Polymun Scientific), 2F5 IgG (Polymun Scientific), b12 IgG (the kind gift of Pamela Bjorkman, Caltech), human ACLA polyclonal serum (LifeSpan BioSciences), and Annexin V (BD Biosciences)  $\pm 2.5$  mM  $\text{CaCl}_2$ , all at 300 nM (**Figs. 2** and **4**), or 4E10[G(L50)E] IgG [34] at 400 nM (**Fig. 4B**) and the three HCDR3-grafted Ubq-fusions at 1  $\mu\text{M}$  (**Fig. 3D**), were injected in duplicate analyses at 50  $\mu\text{L/min}$  for three minutes and allowed to dissociate for five minutes. In the salt series experiment (**Fig. 4A**), 4E10 IgG was diluted to 300 nM in HBS-N-BSA with added NaCl to 575, 320, or 150 mM total NaCl and analyzed in matched running buffers. HIV gp140 SF162 trimers [34] at 30  $\mu\text{g/mL}$  in 10 mM sodium acetate (pH = 5.0) were immobilized at a density of 3250 RUs on a CM5 sensor chip (GE Healthcare) using standard amine coupling chemistry (**Fig. 3C**). An activated/deactivated flow cell was used as a reference surface. HCDR3-grafted Ubq-fusions were injected in duplicate experiments in 10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% (v/v) surfactant P20 (HBS-EP+) at 1  $\mu\text{M}$  and 5 or 10  $\mu\text{M}$  (depending on maximum responses) at 50  $\mu\text{L/min}$  for three minutes and allowed to dissociate for six minutes (**Fig. 3C**). Approximately 3000 RUs of each Scn-peptide fusion construct or a Strep-tagged Scn control [74] were immobilized on CM5 chips (GE Healthcare) using standard amine coupling chemistry and HBS-EP+ as running buffer (**Fig. 9**); optimal immobilization conditions were determined for each fusion by pH scouting. DTT was included in the immobilization buffer for fusions 1 and 3 to prevent oxidation of cysteines in the 36-mer peptide sequences. Duplicate 4E10 IgG samples at 300 nM were injected at 50  $\mu\text{L/min}$  for three minutes and allowed to dissociate for 15 minutes in HBS-EP+ or HBS-EP+ at 500 mM NaCl. In order to determine the specific activity of the fusion-coupled surfaces, each Scn-peptide fusion surface was saturated with StrepMAB-Immo IgG (IBA BioTAGnology) by making three five minute injections of 100 nM IgG at 30  $\mu\text{L/min}$ . Final SPR responses of 4E10 IgG binding to Scn-peptide fusions (**Fig. 9**) were normalized in Scrubber 2.0b (Biologic) software by comparing the observed specific activities (fusion #1: 13%; fusion #2: 14%; fusion #3: 29%; fusion #4: 12%; fusion #5: 34%; Scn controls: 51% or 40%). Specific activities are reduced from 100% by either steric occlusion through coupling or proteolytic degradation of the Scn-peptide fusions. All SPR data were double-referenced [49] using BiaEval version 2.0.3 (GE Healthcare) and plotted with Prism 5.0d for Mac OSX software.

### Crystallography

Crystals of ligand-free 4E10 Fv [34], expressed without purification tags using the Daedalus system [74], were grown at 25°C by the hanging-drop vapor-diffusion method with a well solution of 0.8 M  $\text{Li}_2\text{SO}_4$  buffered with 0.1 M NaOAc (pH = 4.6). Crystals grew within one week and were cryopreserved in a mother liquor of well solution plus 15% v/v glycerol. Diffraction data were collected with both  $\text{CuK}\alpha$  radiation on an in-house R-AXIS IV++ image plate detector with HR optics (Rigaku) at  $-170^\circ\text{C}$  and at the Advanced Light Source beamline 5.0.2 (ALS; Lawrence Berkeley National Laboratory, Berkeley, CA) and reduced using HKL-2000 [95]. Initial phase information was determined by molecular replacement with Phaser [96], as

implemented in the CCP4i program suite [97], a previous bound 4E10 structure (3H3P.pdb [34]), with CDRs deleted, as a search model. Phases were improved by subsequent rounds of model building and refinement using COOT [98] and REFMAC [99] with the ALS data. Eight TLS groups, one for each heavy and light chain, were used during refinement [100,101]. Structure validation was carried out with the MolProbity server [102], and the RCSB ADIT validation server. An anomalous difference Fourier synthesis, using in-house native data reduced without merging Bijvoet pairs, was calculated to verify  $\text{SO}_4$  placement and the initial correctness of the molecular replacement solution. Crystallographic statistics are reported in **Table 1** and the final coordinates have been deposited in the PDB [103], accession code 4LLV.

### PhIP-Seq analysis

For each of the two Abs analyzed, 3 mg of magnetic beads (Invitrogen Dynabeads M-270 Epoxy) were resuspended in 60  $\mu\text{L}$  0.1 M  $\text{NaPO}_4$ , pH 7.4. Beads were rocked at ambient for 24 hrs with 60  $\mu\text{g}$  of either 4E10 or b12 IgG in 1 M  $(\text{NH}_4)_2\text{SO}_4$  and then washed with 10 mM glycine in PBS to cap unreacted epoxy groups. Activity of 4E10 and b12 beads was confirmed by epitope-scaffold [92,104] binding prior to PhIP-Seq analyses. PhIP-Seq analyses were performed as described in [45,105]. Briefly, 4E10 and b12 beads were incubated with the T7-Pep library. DNA from immunoprecipitated phage was recovered and sequenced. Two replicate experiments were performed for both 4E10 and b12 and the enrichment significance for each clone was calculated as  $-\text{Log}_{10}$  *P*-values.

### IHC

Four micron sections of formalin-fixed and paraffin-embedded tissues were cut and baked for 1 hr at 60°C, deparaffinized with xylene and rehydrated in a series of graded ethanol solutions (100%, 95%, and 80%) followed by distilled water. Antigen retrieval in pH 6.0 Citrate Buffer (Dako Target Retrieval Solution S1699) was then performed for 5 min in a pressure cooker. All of the following incubations were performed at room temperature on a Dako Autostainer Plus. Wash buffer (0.05M Tris, 0.3M NaCl, and 0.1% Tween-20, pH 7.2–7.6) was used between all subsequent steps except between the protein block step and application of the primary Ab. A 3% v/v hydrogen peroxide block step was done for 8 min. IP<sub>3</sub>R1 staining: a protein block step consisting of 0.05M Tris, 0.15M NaCl, 0.25% Casein, 0.1% Tween 20, pH 7.6 (TCT Buffer) was performed for 10 minutes. Tissues were then stained with polyclonal rabbit anti-human IP<sub>3</sub>R1 IgG (LifeSpan Biosciences) at a dilution of 1:200 (2.5  $\mu\text{g/mL}$ ) or rabbit IgG matched to the protein concentration of the primary Ab for 60 minutes. Mach 2 Rabbit HRP-polymer (Biocare Medical RHRP520L) was applied for 30 minutes. 4E10 staining: avidin block (Biocare Medical AB972M-A) was applied for 10 minutes, followed by biotin block (Biocare Medical AB972M-B) for 10 minutes. Next, a protein block step (TCT Buffer plus 15% goat serum) was performed for 10 minutes. Tissues were then stained with 4E10 IgG (Polymun AB004) at a dilution of 1:100 (10  $\mu\text{g/mL}$ ) or normal human serum matched to the protein concentration of the primary Ab for 60 minutes. The Fab of a biotinylated goat anti-human secondary Ab (Jackson Laboratory 109-067-993) was applied for 30 minutes followed by incubation with streptavidin-HRP (Jackson Laboratory 016-030-084) at a dilution of 1:2000 for 30 minutes. Finally, a DAB+ substrate-chromogen incubation was done (Dako K3468) for two applications of 4 min each. The slides were counterstained with

Automation Hematoxylin (Dako S3301) for 2 minutes and then coverslipped.

## Acknowledgments

We would like to thank Chris Carrico for suggesting Ubq as an HCDR3 scaffold and Bonnie Cuthbert, and Madeline LaPatra for assistance with fusion protein production; Dave McDonald and Julio Vazquez from the FHCRC Scientific Imaging shared resource for assistance with microscopy and image analysis; and Kimberly Melton and Julie Randolph-Habecker

from the FHCRC Experimental Histopathology shared resource for assistance with IHC.

## Author Contributions

Conceived and designed the experiments: KAKF KL HBL DF CC SJE PDG RKS. Performed the experiments: KAKF KL HBL DF CC PBR. Analyzed the data: KAKF KL HBL DF PBR SJE PDG RKS. Contributed reagents/materials/analysis tools: KAKF KL HBL DF CC PBR SJE PDG RKS. Wrote the paper: KAKF KL HBL DF CC PBR SJE PDG RKS.

## References

- Wyatt R, Sodroski J (1998) The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* 280: 1884–1888.
- Montefiori D, Sattentau Q, Flores J, Esparza J, Mascola J (2007) Antibody-based HIV-1 vaccines: recent developments and future directions. *PLoS Med* 4: e348.
- Schief WR, Ban YE, Stamatatos L (2009) Challenges for structure-based HIV vaccine design. *Curr Opin HIV AIDS* 4: 431–440.
- Zwick MB, Labrijn AF, Wang M, Spelshauer C, Saphire EO, et al. (2001) Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. *J Virol* 75: 10892–10905.
- D'Souza MP, Geyer SJ, Hanson CV, Hendry RM, Milman G (1994) Evaluation of monoclonal antibodies to HIV-1 envelope by neutralization and binding assays: an international collaboration. *AIDS* 8: 169–181.
- Stiegler G, Kunert R, Purtscher M, Wolbank S, Voglauer R, et al. (2001) A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1. *AIDS Res Hum Retroviruses* 17: 1757–65.
- Binley JM, Wrinn T, Korber B, Zwick MB, Wang M, et al. (2004) Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 78: 13232–13252.
- Brunel FM, Zwick MB, Cardoso RM, Nelson JD, Wilson IA, et al. (2006) Structure-function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody. *J Virol* 80: 1680–1687.
- Cardoso RM, Brunel FM, Ferguson S, Zwick M, Burton DR, et al. (2007) Structural basis of enhanced binding of extended and helically constrained peptide epitopes of the broadly neutralizing HIV-1 antibody 4E10. *J Mol Biol* 365: 1533–1544.
- Cardoso RM, Zwick MB, Stanfield RL, Kunert R, Binley JM, et al. (2005) Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity* 22: 163–173.
- Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, et al. (2011) Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* 477: 466–470.
- Walker LM, Phogat SK, Chan-Hui PY, Wagner D, Phung P, et al. (2009) Broad and Potent Neutralizing Antibodies from an African Donor Reveal a New HIV-1 Vaccine Target. *Science* 326: 285–289.
- Li Y, Svehla K, Louder MK, Wycuff D, Phogat S, et al. (2009) Analysis of neutralization specificities in polyclonal sera derived from human immunodeficiency virus type 1-infected individuals. *J Virol* 83: 1045–1059.
- Sather DN, Armann J, Ching LK, Mavrantoni A, Sellhorn G, et al. (2009) Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection. *J Virol* 83: 757–769.
- Muster T, Steindl F, Purtscher M, Trkola A, Klima A, et al. (1993) A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *J Virol* 67: 6642–6647.
- Haynes BF, Moody MA, Verkoczy L, Kelsoe G, Alam SM (2005) Antibody polyspecificity and neutralization of HIV-1: a hypothesis. *Hum Antibodies* 14: 59–67.
- Verkoczy L, Kelsoe G, Moody MA, Haynes BF (2011) Role of immune mechanisms in induction of HIV-1 broadly neutralizing antibodies. *Curr Opin Immunol* 23: 383–390.
- Alam SM, McAdams M, Boren D, Rak M, Searce RM, et al. (2007) The role of antibody polyspecificity and lipid reactivity in binding of broadly neutralizing anti-HIV-1 envelope human monoclonal antibodies 2F5 and 4E10 to glycoprotein 41 membrane proximal envelope epitopes. *Journal of Immunology* 178: 4424–4435.
- Haynes BF, Fleming J, St Clair EW, Katinger H, Stiegler G, et al. (2005) Cardiophilic polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science* 308: 1906–1908.
- Yang G, Holl TM, Liu Y, Li Y, Lu X, et al. (2013) Identification of autoantigens recognized by the 2F5 and 4E10 broadly neutralizing HIV-1 antibodies. *J Exp Med* 210: 241–256.
- Alam SM, Morelli M, Dennison SM, Liao HX, Zhang R, et al. (2009) Role of HIV membrane in neutralization by two broadly neutralizing antibodies. *Proc Natl Acad Sci U S A* 106: 20234–20239.
- Scherer EM, Leaman DP, Zwick MB, McMichael AJ, Burton DR (2010) Aromatic residues at the edge of the antibody combining site facilitate viral glycoprotein recognition through membrane interactions. *Proc Natl Acad Sci U S A* 107: 1529–1534.
- Cohn M (2008) An in depth analysis of the concept of “polyspecificity” assumed to characterize TCR/BCR recognition. *Immunol Res* 40: 128–147.
- Wucherpfennig KW, Allen PM, Celada F, Cohen IR, De Boer R, et al. (2007) Polyspecificity of T cell and B cell receptor recognition. *Semin Immunol* 19: 216–224.
- Herzog S, Jumaa H (2012) Self-recognition and clonal selection: autoreactivity drives the generation of B cells. *Curr Opin Immunol* 24: 166–172.
- Shlomchik MJ (2008) Sites and stages of autoreactive B cell activation and regulation. *Immunity* 28: 18–28.
- Vettermann C, Jack HM (2010) The pre-B cell receptor: turning autoreactivity into self-defense. *Trends Immunol* 31: 176–183.
- Verkoczy L, Diaz M, Holl TM, Ouyang YB, Bouton-Verville H, et al. (2010) Autoreactivity in an HIV-1 broadly reactive neutralizing antibody variable region heavy chain induces immunologic tolerance. *Proc Natl Acad Sci U S A* 107: 181–186.
- Kallenberg CG, van der Meulen J, Pastoor GW, Snijder JA, Feltkamp TE, et al. (1983) Human fibroblasts, a convenient nuclear substrate for detection of anti-nuclear antibodies including anti-centromere antibodies. *Scand J Rheumatol* 12: 193–200.
- Houtkooper RH, Vaz FM (2008) Cardiophilin, the heart of mitochondrial metabolism. *Cell Mol Life Sci* 65: 2493–2506.
- Osman C, Voelker DR, Langer T (2011) Making heads or tails of phospholipids in mitochondria. *J Cell Biol* 192: 7–16.
- Suh-Lailam BB, Cromar A, Davis KW, Tebo AE (2012) APhL antibody ELISA as an alternative to anticardiolipin test for the diagnosis of antiphospholipid syndrome. *Int J Clin Exp Pathol* 5: 210–215.
- Scherer EM, Zwick MB, Teyton L, Burton DR (2007) Difficulties in eliciting broadly neutralizing anti-HIV antibodies are not explained by cardiophilin autoreactivity. *AIDS* 21: 2131–2139.
- Xu H, Song L, Kim M, Holmes MA, Kraft Z, et al. (2010) Interactions between lipids and human anti-HIV antibody 4E10 can be reduced without ablating neutralizing activity. *J Virol* 84: 1076–1088.
- Tong T, Crooks ET, Osawa K, Binley JM (2012) HIV-1 virus-like particles bearing pure env trimers expose neutralizing epitopes but occlude nonneutralizing epitopes. *J Virol* 86: 3574–3587.
- Ferguson CG, James RD, Bigman CS, Shepard DA, Abdiche Y, et al. (2005) Phosphoinositide-containing polymerized liposomes: stable membrane-mimetic vesicles for protein-lipid binding analysis. *Bioconjug Chem* 16: 1475–1483.
- Martinez V, Diemert MC, Braibant M, Potard V, Charuel JL, et al. (2009) Anticardiolipin antibodies in HIV infection are independently associated with antibodies to the membrane proximal external region of gp41 and with cell-associated HIV DNA and immune activation. *Clin Infect Dis* 48: 123–132.
- Sene D, Piette JC, Cacoub P (2009) [Antiphospholipid antibodies, antiphospholipid syndrome and viral infections]. *Rev Med Interne* 30: 135–141.
- Galrao L, Brites C, Atta ML, Lima A, et al. (2007) Antiphospholipid antibodies in HIV-positive patients. *Clin Rheumatol* 26: 1825–1830.
- Vcelar B, Stiegler G, Wolf HM, Muntean W, Leschnik B, et al. (2007) Reassessment of autoreactivity of the broadly neutralizing HIV antibodies 4E10 and 2F5 and retrospective analysis of clinical safety data. *AIDS* 21: 2161–2170.
- Singh H, Henry KA, Wu SS, Chruscinski A, Utz PJ, et al. (2011) Reactivity profiles of broadly neutralizing anti-HIV-1 antibodies are distinct from those of pathogenic autoantibodies. *AIDS* 25: 1247–1257.
- Burton DR, Barbas 3rd CF, Persson MA, Koenig S, Chanock RM, et al. (1991) A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. *Proc Natl Acad Sci U S A* 88: 10134–10137.
- Roben P, Moore JP, Thali M, Sodroski J, Barbas 3rd CF, et al. (1994) Recognition properties of a panel of human recombinant Fab fragments to the CD4 binding site of gp120 that show differing abilities to neutralize human immunodeficiency virus type 1. *J Virol* 68: 4821–4828.
- Sajadi MM, Lewis GK, Seaman MS, Guan Y, Redfield RR, et al. (2012) Signature biochemical properties of broadly cross-reactive HIV-1 neutralizing antibodies in human plasma. *J Virol* 86: 5014–5025.

45. Larman HB, Zhao Z, Laserson U, Li MZ, Ciccio A, et al. (2011) Autoantigen discovery with a synthetic human peptidome. *Nat Biotechnol* 29: 535–541.
46. Hangartner L, Senn BM, Ledermann B, Kalinke U, Seiler P, et al. (2003) Antiviral immune responses in gene-targeted mice expressing the immunoglobulin heavy chain of virus-neutralizing antibodies. *Proc Natl Acad Sci U S A* 100: 12883–12888.
47. Hodnik V, Anderluh G (2010) Capture of intact liposomes on biacore sensor chips for protein-membrane interaction studies. *Methods Mol Biol* 627: 201–211.
48. Harris EN, Pierangeli SS (2002) Revisiting the anticardiolipin test and its standardization. *Lupus* 11: 269–275.
49. Myszka DG (1999) Improving biosensor analysis. *J Mol Recognit* 12: 279–284.
50. Brown BK, Karasavvas N, Beck Z, Matyas GR, Bix DL, et al. (2007) Monoclonal antibodies to phosphatidylinositol phosphate neutralize human immunodeficiency virus type 1: role of phosphate-binding subsites. *J Virol* 81: 2087–2091.
51. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, et al. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* 35: W522–W525.
52. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 32: W665–W667.
53. Li H, Robertson AD, Jensen JH (2005) Very fast empirical prediction and rationalization of protein pKa values. *Proteins* 61: 704–721.
54. Unni S, Huang Y, Hanson RM, Tobias M, Krishnan S, et al. (2011) Web servers and services for electrostatics calculations with APBS and PDB2PQR. *J Comput Chem* 32: 1488–1491.
55. Patten PA, Gray NS, Yang PL, Marks CB, Wedemayer GJ, et al. (1996) The immunological evolution of catalysis. *Science* 271: 1086–1091.
56. Wedemayer GJ, Patten PA, Wang LH, Schultz PG, Stevens RC (1997) Structural insights into the evolution of an antibody combining site. *Science* 276: 1665–1669.
57. Romesberg FE, Spiller B, Schultz PG, Stevens RC (1998) Immunological origins of binding and catalysis in a Diels-Alderase antibody. *Science* 279: 1929–1933.
58. Li Y, Li H, Yang F, Smith-Gill SJ, Mariuzza RA (2003) X-ray snapshots of the maturation of an antibody response to a protein antigen. *Nat Struct Biol* 10: 482–488.
59. Terzian S, Ramsland PA, Voss Jr EW, Herron JN, Edmundson AB (2004) Three-dimensional structures of idiotypically related Fabs with intermediate and high affinity for fluorescein. *J Mol Biol* 339: 1141–1151.
60. Manivel V, Sahoo NC, Salunke DM, Rao KV (2000) Maturation of an antibody response is governed by modulations in flexibility of the antigen-combining site. *Immunity* 13: 611–620.
61. Sagawa T, Oda M, Ishimura M, Furukawa K, Azuma T (2003) Thermodynamic and kinetic aspects of antibody evolution during the immune response to hapten. *Mol Immunol* 39: 801–808.
62. Li Y, Lipschultz CA, Mohan S, Smith-Gill SJ (2001) Mutations of an epitope hot-spot residue alter rate limiting steps of antigen-antibody protein-protein associations. *Biochemistry* 40: 2011–2022.
63. Mohan S, Sinha N, Smith-Gill SJ (2003) Modeling the binding sites of anti-hen egg white lysozyme antibodies HyHEL-8 and HyHEL-26: an insight into the molecular basis of antibody cross-reactivity and specificity. *Biophys J* 85: 3221–3236.
64. Acierno JP, Braden BC, Klinke S, Goldbaum FA, Cauerhff A (2007) Affinity maturation increases the stability and plasticity of the Fv domain of anti-protein antibodies. *J Mol Biol* 374: 130–146.
65. Zimmermann J, Romesberg FE, Brooks 3rd CL, Thorpe IF (2010) Molecular description of flexibility in an antibody combining site. *J Phys Chem B* 114: 7359–7370.
66. Kallewaard NL, McKinney BA, Gu Y, Chen A, Prasad BV, et al. (2008) Functional maturation of the human antibody response to rotavirus. *J Immunol* 180: 3980–3989.
67. Schmidt AG, Xu H, Khan AR, O'Donnell T, Khurana S, et al. (2013) Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. *Proc Natl Acad Sci U S A* 110: 264–269.
68. Hager-Braun C, Katinger H, Tomer KB (2006) The HIV-neutralizing monoclonal antibody 4E10 recognizes N-terminal sequences on the native antigen. *J Immunol* 176: 7471–7481.
69. Pruitt KD, Tatusova T, Brown GR, Maglott DR (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res* 40: D130–D135.
70. Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 35: D61–D65.
71. Mikoshiba K (2002) [IP3 receptor, a Ca<sup>2+</sup> oscillator—role of IP3 receptor in development and neural plasticity]. *Nihon Yakurigaku Zasshi* 120: 6P–10P.
72. Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2: 28–36.
73. Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4: 363–371.
74. Bandaranayake AD, Correnti C, Ryu BY, Brault M, Strong RK, et al. (2011) Daedalus: a robust, turnkey platform for rapid production of decigram quantities of active recombinant proteins in human cell lines using novel lentiviral vectors. *Nucleic Acids Res* 39: e143.
75. Schmidt TG, Skerra A (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* 2: 1528–1535.
76. Sharp AH, Nucifora Jr FC, Blondel O, Sheppard CA, Zhang C, et al. (1999) Differential cellular expression of isoforms of inositol 1,4,5-triphosphate receptors in neurons and glia in brain. *J Comp Neurol* 406: 207–220.
77. Das BK, Xia L, Palandjian L, Gozani O, Chyung Y, et al. (1999) Characterization of a protein complex containing spliceosomal proteins SAPs 49, 130, 145, and 155. *Mol Cell Biol* 19: 6796–6802.
78. Burton DR (2002) Antibodies, viruses and vaccines. *Nat Rev Immunol* 2: 706–713.
79. Killian JA, von Heijne G (2000) How proteins adapt to a membrane-water interface. *Trends Biochem Sci* 25: 429–434.
80. de Planque MR, Kruijtz JA, Liskamp RM, Marsh D, Greathouse DV, et al. (1999) Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane alpha-helical peptides. *Journal of Biological Chemistry* 274: 20839–20846.
81. Yau WM, Wimley WC, Gawrisch K, White SH (1998) The preference of tryptophan for membrane interfaces. *Biochemistry* 37: 14713–14718.
82. Palsdottir H, Lojero CG, Trumppower BL, Hunte C (2003) Structure of the yeast cytochrome bc1 complex with a hydroxyquinone anion Q<sub>o</sub> site inhibitor bound. *J Biol Chem* 278: 31303–31311.
83. Chakrabarti P (1993) Anion binding sites in protein structures. *J Mol Biol* 234: 463–482.
84. Copley RR, Barton GJ (1994) A structural analysis of phosphate and sulphate binding sites in proteins. Estimation of propensities for binding and conservation of phosphate binding sites. *J Mol Biol* 242: 321–329.
85. Duncan RS, Hwang SY, Koulen P (2007) Differential inositol 1,4,5-triphosphate receptor signaling in a neuronal cell line. *Int J Biochem Cell Biol* 39: 1852–1862.
86. Corti D, Voss J, Gamblin SJ, Codoni G, Macagno A, et al. (2011) A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. *Science* 333: 850–856.
87. Dreyfus C, Ekiert DC, Wilson IA (2013) Structure of a classical broadly neutralizing stem antibody in complex with a pandemic h2 influenza virus hemagglutinin. *J Virol* 87: 7149–7154.
88. Dreyfus C, Laursen NS, Kwaks T, Zuidgeest D, Khayat R, et al. (2012) Highly conserved protective epitopes on influenza B viruses. *Science* 337: 1343–1348.
89. Ekiert DC, Bhabha G, Elsliger MA, Friesen RH, Jongeneelen M, et al. (2009) Antibody recognition of a highly conserved influenza virus epitope. *Science* 324: 246–251.
90. Sui J, Hwang WC, Perez S, Wei G, Aird D, et al. (2009) Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat Struct Mol Biol* 16: 265–273.
91. Guenaga J, Wyatt RT (2012) Structure-guided alterations of the gp41-directed HIV-1 broadly neutralizing antibody 2F5 reveal new properties regarding its neutralizing function. *PLoS Pathog* 8: e1002806.
92. Correia BE, Ban YE, Holmes MA, Xu H, Ellington K, et al. (2010) Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope. *Structure* 18:1116–1126.
93. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison 3rd CA, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6: 343–345.
94. Gibson DG, Smith HO, Hutchison 3rd CA, Venter JC, Merryman C (2010) Chemical synthesis of the mouse mitochondrial genome. *Nat Methods* 7: 901–903.
95. Otwinowski Z, Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Meth Enzymol* 276: 307–326.
96. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. *Journal of Applied Crystallography* 40: 658–674.
97. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* 59: 1131–1137.
98. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126–2132.
99. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53: 240–255.
100. Painter J, Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *Journal of Applied Crystallography* 39: 109–111.
101. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62: 439–450.
102. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35 W375–W383.
103. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. *Nucleic Acids Research* 28: 235–242.

104. Azoitei ML, Correia BE, Ban YE, Carrico C, Kalyuzhnyi O, et al. (2011) Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science* 334: 373–376.
105. Larman HB, Laserson U, Querol L, Verhaeghen K, Solimini NL, et al. (2013) PhIP-Seq characterization of autoantibodies from patients with multiple sclerosis, type 1 diabetes and rheumatoid arthritis. *J Autoimmun* 3: 1–9.
106. Vijay-Kumar S, Bugg CE, Cook WJ (1987) Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* 194: 531–544.
107. Fraczekiewicz R, Braun W (1998) Exact and Efficient Analytical Calculation of the Accessible Surface Areas and Their Gradients for Macromolecules. *J Comp Chem* 19: 319–333.
108. Bryson S, Julien JP, Hynes RC, Pai EF (2009) Crystallographic definition of the epitope promiscuity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5: vaccine design implications. *J Virol* 83: 11862–11875.
109. DeLano WL (2002) The PyMOL Molecular Graphics System. DeLano Scientific: San Carlos, CA, USA.
110. Saphire EO, Parren PW, Pantophlet R, Zwick MB, Morris GM, et al. (2001) Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science* 293: 1155–1159.
111. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 18: 6097–6100.
112. Wootton JC (1994) Non-globular domains in protein sequences: automated segmentation using complexity measures. *Comput Chem* 18: 269–285.
113. Pettit FK, Bare E, Tsai A, Bowie JU (2007) HotPatch: a statistical approach to finding biologically relevant features on protein surfaces. *J Mol Biol* 369: 863–879.
114. Swairjo MA, Concha NO, Kaetzel MA, Dedman JR, Seaton BA (1995) Ca<sup>2+</sup>-bridging mechanism and phospholipid head group recognition in the membrane-binding protein annexin V. *Nat Struct Biol* 2: 968–974.