Innate and humoral immune responses to HIV-1

Kathryn Ann Kastor Finton

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Roland K. Strong, Chair

Joan M. Goverman

Ronald E. Stenkamp

Program authorized to offer degree:

Biochemistry

University of Washington

Abstract

Innate and humoral responses to HIV-1 infection

Kathryn Ann Kastor Finton

Chair of the Supervisory Committee:

Roland K. Strong, Ph.D.

Immunology

Understanding cellular and humoral immune responses to HIV-1 infection and designed HIV immunogens are imperative to the development of an effective anti-HIV vaccine. In this thesis we investigate aspects of both the cellular innate and adaptive humoral arms of the immune system in response to HIV-1 infection and vaccination with designed epitope-scaffold immunogens. First, we review the anti-viral contributions from natural killer (NK) cells, focusing on NK cell recognition of MHC class I proteins through a diverse array of activating and inhibitory receptors. We also seek to answer the question of the ligand specificity of an

activating NK cell receptor that is correlated with improved clinical outcomes in HIV-1 infection, KIR3DS1, through structural modeling and functional studies. We find that KIR3DS1 binds HLA -A, -B and -C alleles, but only in the absence of peptide, unusual for this class of receptor which shows fine specificity for allele type and preferences for bound peptide.

Second, we explore the humoral arm of the anti-HIV immune response through characterization of the broadly neutralizing anti-HIV antibody 4E10. 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. The previously proposed candidate 4E10 autoantigen was the mitochondrial lipid cardiolipin. 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. 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.

Lastly, we characterized 4E10 ontogeny through functional and structural studies of an ensemble of 4E10 germline encoded precursors (GEPs). GEPs showed detectable, but extremely weak, binding to soluble Env gp140s and extremely limited neutralization potency, though some reverse engineered epitope-scaffolds showed robust GEP affinities, well above the B cell activation threshold. 4E10 and GEP paratopes displayed a remarkable degree of structural conservation in the antigen-bound state, with little improvement in overall shape complementarity. Frame work region mutations had little discernable affect on global or local structure. Surprisingly, 4E10 thermostability was significantly worse than its GEPs; while 4E10 and GEPs displayed similarly constrained V_H/V_L interdomain movements upon binding, 4E10 maturation involved negligible combining site rigidification, with both 4E10 and GEPs sampling extensive HCDR conformer ensembles. The narrowing of polyspecificity assumed to concur with maturation was not observed with 4E10, as both 4E10 and its GEPs showed similar patterns of limited polyspecificity to phage-displayed human peptidomes (PhIP-Seq). While 4E10 is demonstrably autoreactive, GEPs exhibited a distinct profile of autoantigen recognition by PhIP-Seq, suggesting that autoreactivity was acquired during ontogeny. In many respects,

4E10 provides a divergent example of Ab ontogeny, broadening the known range of affinity maturation pathways and challenging the generality of the existing paradigm. Retained combining site flexibility, and discrepancies in GEP binding of engineered versus Env-derived antigens, suggest that higher order mechanisms of neutralization are in play and that conventional vaccination protocols are unlikely to generate 4E10-equivalent Abs.

Chapter 1: Introduction

In 1981, large numbers of young men in the Los Angeles and San Francisco areas started suffering from a number of opportunistic and often fatal infections due to immune incompetence. The term, acquired immunodeficiency syndrome (AIDS), was coined to describe the condition. At this time, only one year had passed since the discovery of the first human retrovirus, the human T-lymphotropic virus, by Dr. Robert Gallo of the National Cancer Institute. In just two more years, another retrovirus, the human immunodeficiency virus (HIV-1) would be isolated and found to be the causative agent of AIDS by Dr. Gallo and Dr. Luc Montagnier at the Pasteur Institute in Paris. Since the epidemic began, almost 70 million people have been infected with HIV and approximately 35 million people have died of AIDS. Currently, over 34 million people are infected with HIV-1.

The three stages of HIV-1 infection are: (1) acute infection, (2) latent infection, and (3) AIDS. (1) Acute infection occurs immediately after contraction of the virus and can be associated with an influenza-like illness. The first cellular targets are Langerhans cells (dendritic cells (DCs) residing in the skin and mucosa). HIV-1 is carried by these antigen presenting cells to $CD4^+$ T cells; the infected $CD4^+$ T cells then traffic to the lymph nodes and disseminate within 4 to 11 days. Within the next couple months, the viral load increases exponentially and is accompanied by the death of many $CD4^+$ T cells. The elevated viral load is finally stabilized, predominately by the innate immune response (at one to two weeks) and, secondarily, by the adaptive immune response (at four to eight weeks); the viral load set point is established at a median time of two months. A high viral load set point and low $CD4^+$ T cell count are associated with a worse prognosis for disease progression[1]. (2) During the latent or chronic phase of infection, which

lasts eight years on average, the viral load is kept relatively constant due to the efforts of the host immune system and the rapid turnover of infected cells. $CD4^+$ T cell numbers decrease and virus sequence diversity increases. (3) The third phase is defined once the $CD4^+$ T cell count falls bellow 200/μL or with the occurrence of certain opportunistic infections or cancers and is characterized by dramatic increases in viral load and loss of $CD4^+$ T cells [2].

Host defenses

Following acute infection and in the absence of retroviral therapy, the disease course is observed to fall into one of four different classes: (1) *rapid progressors* develop AIDS in under 3 years; (2) *slow progressors* develop AIDS in 8-10 years; (3) *long-term survivors* do not develop AIDS, but do not maintain $CD4^+$ T cell levels; (4) *long-term nonprogressors,* or *elite controllers*, do not develop AIDS and do not show a decline in $CD4^+$ T cell levels. The fact that there are different outcomes to HIV infection, from rapid progression to elite control, illustrates the importance of host defenses in the course of HIV-1 infection.

Cellular innate immunity

The body's response to HIV in the first few weeks following infection plays a critical role in determining the viral set point and, consequently, disease progression. At this time, the innate immune system is responsible both for initial control of virus and also for inducing the adaptive immune response. Cells of the innate immune system include cytolytic cells (natural killer cells [NK], and natural killer T cells) capable of direct killing of pathogen infected cells, phagocytes (dendritic cells (DCs), macrophages, monocytes) capable of antigen clearance, professional

antigen-presenting cells (DCs, macrophages, B cells) able to capture and present foreign antigens to cells of the adaptive immune system, granulocytes (eosinophils, basophils, neutrophils), γδ T cells, and mast cells. These cells patrol the body and respond to foreign antigens through receptor engagement without the need for prior exposure to the antigen. Receptor activation results in the release of cytokines such as interferons (IFNs) that create an antiviral environment, recruit immune cells and shape the quality of the adaptive immune response. Among the innate immune cell types, NK cells in the mucosa are first in the line of defense against HIV-1 infected cells. As discussed in chapter two, epidemiologic data strongly suggest that NK cells, particularly those that express certain receptor alleles in conjunction with their ligands (major histocompatibility complex [MHC] class I molecules), are associated with a more favorable disease progression.

NK cells, as well as macrophages, neutrophils, and eosinophils, can also exert protective effects through antibody-dependent cell-mediated cytotoxicity (ADCC) and antibody-dependent cell-mediated virus inhibition (ADCVI). ADCC is a mechanism whereby innate effector cells bind to the Fc portion of an antibody (Ab) that is bound to virally infected cells and cause lysis of the infected cell through release of cytotoxic granules. ADCVI refers to viral inhibition as a result of the death of the infected cell. The importance of ADCC/ADCVI in HIV-1 infected individuals is demonstrated by the finding that disease progression is inversely correlated to high titers of gp120-specific ADCC Abs[3]. Furthermore, elite controllers show higher ADCC Ab titers than non-elite controllers[4]. Chapter two continues the discussion on NK cells and how they respond to viral infection in general and HIV-1 infection in particular.

Humoral immunity

Abs are the soluble form of the B cell receptor secreted by plasma or memory B cells and can

neutralize infection by blocking the interaction of virus with cellular receptors, inhibiting viral fusion with the host cell membrane, or inhibiting the transition of virus from endocytic vesicles into the cytoplasm. In addition to neutralization, Abs can bind to infected cells or virions and provide protection by recruiting effector cells or complement leading to the lysis and removal of infected cells or viral particles. The mechanisms of ADCC and ADCVI can be engaged by both neutralizing and non-neutralizing Abs. The term broadly neutralizing Ab (bnAb) is used to delineate Abs that are capable of neutralizing a broad range of viral isolates. It is important to keep in mind that epitope binding does not always equal neutralization. Furthermore, there is no evidence to suggest that neutralizing Abs are preferentially selected through the affinity maturation process as B cells are selected on the ability of their B cell receptor to bind antigen with sufficient affinity (not on the ability to neutralize infection).

Early in infection, HIV-1 elicits many antibodies (Abs) that can bind the virus; however, they often will not neutralize the infection. bnAbs have the daunting task of finding conserved epitopes among the highly glycosylated (~50% total mass of gp120[5]) and sequence variable HIV-1 envelope trimer (*Env* displays up to 35% sequence diversity among clades, 20% within clades, and 10% in a single infected person[6,7]). Characteristics of anti-HIV bnAbs, thought to develop after persistent infection and exposure to virus, include extensive somatic mutations and nucleotide insertions or deletions (15- 44% change in amino acid substitutions versus 5-20% for non anti-HIV bnAbs[8]), a long heavy chain complementary determining region 3 (HCDR3) (up to 33 amino acids long versus an average of 13 for non anti-HIV bnAbs[9,10]). Another characteristic thought to be shared among bnAbs is polyspecificity, or the ability to bind multiple epitopes. It is a quality that, while on one hand, may allow bnAbs to bind changeable HIV epitopes, may also lead to autoreactivity. These features, along with Ab ontogeny, will be discussed in chapters three and four.

The potency of bnAbs in HIV-1 infection can be demonstrated by the protection against simian HIV (SHIV) afforded by passive transfer of bnAbs[11-16] and the association between maternal bnAbs and lower rates of HIV-1 transmission to the child[17]. However, in individuals newly infected with HIV, the bnAb response develops much too late (~2-4 years after infection) to provide protection against infection[18,19]. Furthermore, individuals making bnAbs do not show signs of increased control of viremia nor does the prevalence of bnAbs in long-term progressors versus nonprogressors differ; indicating that, once infected, bnAbs do little to help control infection. Ab based prophylactics and therapeutics, however, may hold value. Vaccination in humanized mice with an AAV vector containing the full-length VRC01 Ab gene resulted in sustained levels of Ab that were able to protect from challenge with HIV-1[20]. In another study, patients undergoing interruption of antiretroviral therapy treatment were given a mixture of bnAbs 2G12, 2F5, and 4E10. Initially, viral loads were reduced; however, the effect was transient due to the emergence of 2G12 escape mutants. More recently, in an effort to prevent viral escape mutants, a mixture of five Abs (NIH45-46^{G54W}, PG16, PGT128, 10-1074, and 3BC176) were selected and given to humanized HIV-1 infected mice. In this study, the Ab mixture effectively controlled HIV-1 replication in humanized mice, suppressing viral loads below the level of detection, and did not select for escape mutants whereas when a single Ab or a mixture of three Abs was used, escape mutants were found [21].

Vaccine design

Although neither cellular nor humoral immune responses are sufficient to clear an established HIV-1 infection, it is hoped that an immune system primed through vaccination may be suitably equipped to prevent or attenuate infection. The target of anti-HIV-1 bnAbs, and, therefore, for an

anti-HIV-1 vaccine, is the viral envelope protein. Since the 1980's, approximately 200 human vaccine trials have taken place. Due to safety concerns or lack of immunogenicity, only 17 have progressed to at least phase II trials, and, of those, only five have advanced to efficacy testing[22]. Vaccine strategies comprising those that have made it to efficacy testing include recombinant gp120 subunit vaccines, AIDSVAX B/B and AIDSVAX B/E[23,24], HIV-LIPO-5 (lipopeptides incorporating $CD8^+$ epitopes from Gag, Pol, and Nef)[25], and Adenovirus (Merck Ad5, rAd5)[26,27], modified vaccinia ankara (rMVA)[28] or canarypox (ALVAC)[29] vectors expressing Gag, Pol, and Nef. In these vaccine strategies, heterologous prime-boost immunizations were frequently used in order to strengthen and broaden immune responses. Nevertheless, none reduced infection rates or improved post infection secondary end points such as viral load or $CD4^+$ T cell count; however, many were able to elicit $CD8^+$ and $CD4^+$ responses and may be of value when incorporated into a prime-boost regime.

In the midst of all these failures, the RV144 trial has been the first to raise some hope that HIV-1 prevention through vaccination is an achievable goal[30]. The vaccine used in the phase III RV144 trial consisted of ALVAC at 0, 1, 2, and 6 months followed by boosting with AIDSVAX B/E at 3 and 6 months. The vaccine showed an efficacy of 31% at 42 months; however, the correlates of protection are still not clear. A major limitation of the RV144 trial is that peak Abs levels quickly dropped with half-lives in the range of a few weeks. In fact, efficacy in the first year was 60% which suggests that, were the Abs longer lived, the vaccine would have been much more successful. Apart from sustaining Ab levels, an additional problem encountered in the development of vaccines is that Abs elicited through immunizations are only able to neutralize tier 1 viruses (viruses highly sensitive to Ab-mediated neutralization), and not neutralization-resistant tier 2 and 3 viruses, which account for the majority of clinical isolates.

A number of bnAbs have been isolated and characterized in an effort to explicate potential epitope targets for immunogen design. These Abs include several that recognize conserved regions on gp120 (b12, PG9, PG16, VRC01, 2G12[31-34]) as well as others that recognize linear epitopes on the membrane proximal external region (MPER) of gp41 (4E10, 2F5, Z13e1[35-37]). The vaccine design approach of "reverse engineering" seeks to stabilize an anti-HIV bnAb epitope in its antibody-bound conformation (determined by X-ray crystallography) by computationally grafting the epitope onto a non-immunogenic scaffold protein. "Epitope-scaffolds" (ESs), in theory, are superior to whole protein or peptide immunogens as they focus the immune response on a single conserved epitope in the hopes of re-eliciting the bnAb. This is in contrast to whole protein immunogens that display a broad spectrum of non-conserved epitopes, or flexible peptide immunogens that present non-native conformers representing irrelevant epitopes. ESs have been generated for linear epitopes 2F5[38] and 4E10[39,40] and the discontinuous b12[41] epitope. In each case, ESs bind the respective bnAbs with nanomolar to picomolar affinity and also generate ES-binding sera through vaccination. However, Abs raised against the 2F5 and 4E10 ESs lacked the capability to neutralize virus (the b12 ES has not been tested). In another study[42], a V3 ES "boost" was used in combination with a gp120 DNA "prime" to effectively elicit cross-clade bnAbs. Although V3 ESs were not used alone in vaccinations, another study using a constrained V3 peptide alone was able to generate neutralizing Abs in rabbits[43]. In contrast, a prime boost immunization strategy using heterologous 2F5 ESs, some incorporating T cell epitopes, resulted in the generation of 2F5-like binding Abs that could not neutralize HIV[44]. The failure to elicit bnAbs from MPER derived 2F5 and 4E10 epitope-scaffolds may be due to the possibility that MPER epitopes recognized on an intact virion consist not only of the core epitope included in the scaffold design, but also include additional viral protein or membrane contacts. Despite the general failure for MPER ESs to generate bnAbs, they have provided a useful tool for characterizing binding interactions with germline encoded precursors (GEPs) of mature bnAbs that typically display very weak (or

nonexistent) binding to intact gp120/gp41 or peptide epitopes (covered in chapter four).

Consistently, epitope-scaffold and gp120-based vaccines are able to generate sera that bind to the desired protein or epitope[29,38,40,44-47], but fail to elicit bnAbs. Possible explanations include: (1) the epitope eliciting a bnAb response is only a small, and often less accessible, fraction of the total possible epitopes present, (2) characteristics inherent to bnAb (e.g. polyspecificity and long and hydrophobic or charged HCDR3s) cause autoreactivity and subsequent deletion through tolerance mechanisms (discussed in chapter 3), (3) germline VH and VL genes in non-human vaccinees were not compatible with the development of anti-HIV bnAbs, and, in extension, humans may also not commonly possess the needed germline genes (discussed in chapter four), and (4) the vaccine protocol itself was insufficient or flawed (e.g. immunogen proteins may not maintain native conformations in the designed construct or in the adjuvant (alum) used; a series of different immunogens may need to engage, step-wise, B cell receptors in order to guide the affinity maturation process from germline to mature bnAbs (discussed in chapter four)). It is also important to note that the structural and mechanistic basis for neutralizing versus binding antibodies is not clear and may not be resolved until the atomic level resolution structure of the entire envelope trimer both alone and in complex with bnAbs has been determined. One caveat is that bnAb binding and neutralization may be a dynamic or transient process; for example, binding and neutralization may take place through a multi-step process requiring multiple Ab conformations[48]. Therefore, a single snapshot may be insufficient to explain neutralization mechanisms.

Development of a vaccine remains a formidable undertaking. The maturation pathway from germline to mature anti-HIV-1 bnAbs appears to be the result of an elaborate and long process of selection against rapidly mutation virus that, in the end, has the potential to generate self reactive Abs. In addition to the humoral response, a successful vaccine will most likely have to

engage the early responding innate arm of the immune system through ADCC and/or effector cells of the adaptive immune system, such as $CD8^+$ and $CD4^+$ T cells. Difficulties inherent in dealing with HIV biology are compounded by the lack of clear immune correlates of protection against infection. The challenge now is to gain a more in-depth understanding of both the innate and adaptive immune response in the context, not only of infection, but also vaccination, and to translate this knowledge into the development of new immunogens and vaccine strategies that will elicit potent and lasting immunity to HIV infection. In an effort to meet this challenge, this body of work, beginning in chapter two, first describes the innate immune response of NK cells to viral infection with an emphasis on HIV-1 and their role in disease progression. Chapter three addresses why 4E10, an important anti-HIV-1 bnAb, is not readily elicited through natural infection or immunization. The discussion focuses on 4E10's apparent autoreactivity, polyspecificity, and uncommon flexibility: characteristics that may be common to many hard-to-elicit bnAbs. In chapter four, 4E10 bnAb ontogeny is explored including discussions on structure, binding kinetics, and neutralization profiles of GEPs versus mature Ab. These findings contradict the current understanding of Ab ontogeny and illustrate how difficult elicitation of a 4E10-like bnAb may be. Finally, chapter five includes future directions for all three lines of research including ongoing experimental work describing the specificity of two NK cell receptors for MHC class I ligands in the context of HIV-1 infection, structural verification of the 4E10 autoantigen, and characterization of 4E10 ontogeny. Chapters two, three, and four and the references and figures therein, are written and arranged in the style mandated by the journal in which they were published (or submitted). Therefore, references and figures are designated by the chapter in which they are contained. Coauthors are listed in the acknowledgements at the end of each relevant chapter.

References

1. Richey, L. E. & Halperin, J. Acute human immunodeficiency virus infection. *Am. J. Med. Sci.* **345,** 136–142 (2013).
2. Alizon, S. & Magnus, C. Modelling the Course of an HIV Infection: Insights from Ecology and Evolution. *Viruses* **4,** 1984–2013 (2012).
3. Baum, L. L. *et al.* HIV-1 gp120-specific antibody-dependent cell-mediated cytotoxicity correlates with rate of disease progression. *J. Immunol.* **157,** 2168–2173 (1996).
4. Lambotte, O. *et al.* Heterogeneous neutralizing antibody and antibody-dependent cell cytotoxicity responses in HIV-1 elite controllers. *AIDS* **23,** 897–906 (2009).
5. Leonard, C. K. *et al.* Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J Biol Chem* **265,** 10373–10382 (1990).
6. Buonaguro, L., Tornesello, M. L. & Buonaguro, F. M. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *Journal of Virology* **81,** 10209–10219 (2007).
7. Taylor, B. S. & Hammer, S. M. The challenge of HIV-1 subtype diversity. *N. Engl. J. Med.* **359,** 1965–1966 (2008).
8. Corti, D. & Lanzavecchia, A. Broadly Neutralizing Antiviral Antibodies. *Annu. Rev. Immunol.* **31,** 705–742 (2013).
9. Johnson, G. & Wu, T. T. Kabat database and its applications: 30 years after the first variability plot. *Nucleic Acids Research* **28,** 214–218 (2000).
10. Johnson, G. & Wu, T. T. Preferred CDRH3 lengths for antibodies with defined specificities. *International Immunology* **10,** 1801–1805 (1998).

11. Mascola, J. R. *et al.* Protection of Macaques against pathogenic simian/human immunodeficiency virus 89.6PD by passive transfer of neutralizing antibodies. *Journal of Virology* **73,** 4009–4018 (1999).

12. Mascola, J. R. *et al.* Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nature Medicine* **6,** 207–210 (2000).

13. Parren, P. W. *et al.* Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro. *Journal of Virology* **75,** 8340–8347 (2001).

14. Baba, T. W. *et al.* Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. *Nature Medicine* **6,** 200–206 (2000).

15. Ferrantelli, F. *et al.* Complete protection of neonatal rhesus macaques against oral exposure to pathogenic simian-human immunodeficiency virus by human anti-HIV monoclonal antibodies. *J INFECT DIS* **189,** 2167–2173 (2004).

16. Hofmann-Lehmann, R. *et al.* Postnatal passive immunization of neonatal macaques with a triple combination of human monoclonal antibodies against oral simian-human immunodeficiency virus challenge. *Journal of Virology* **75,** 7470–7480 (2001).

17. Barin, F. *et al.* Revisiting the role of neutralizing antibodies in mother-to-child transmission of HIV-1. *J INFECT DIS* **193,** 1504–1511 (2006).

18. Gray, E. S. *et al.* The neutralization breadth of HIV-1 develops incrementally over four years and is associated with CD4+ T cell decline and high viral load during acute infection. *Journal of Virology* **85,** 4828–4840 (2011).

19. Mikell, I. *et al.* Characteristics of the Earliest Cross-Neutralizing Antibody Response to HIV-1. *PLoS Pathog* **7,** e1001251 (2011).

20. Balazs, A. B. *et al.* Antibody-based protection against HIV infection by vectored

immunoprophylaxis. *Nature* **481,** 81–84 (2012).

21. Klein, F. *et al.* HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* **492,** 118–122 (2012).

22. O'Connell, R. J., Kim, J. H., Corey, L. & Michael, N. L. Human Immunodeficiency Virus Vaccine Trials. *Cold Spring Harbor Perspectives in Medicine* **2,** a007351–a007351 (2012).

23. Harro, C. D. *et al.* Recruitment and baseline epidemiologic profile of participants in the first phase 3 HIV vaccine efficacy trial. *J. Acquir. Immune Defic. Syndr.* **37,** 1385–1392 (2004).

24. Flynn, N. M. *et al.* Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. *J INFECT DIS* **191,** 654–665 (2005).

25. Salmon-Ceron, D. *et al.* Immunogenicity and safety of an HIV-1 lipopeptide vaccine in healthy adults: a phase 2 placebo-controlled ANRS trial. *AIDS* **24,** 2211–2223 (2010).

26. Buchbinder, S. P. *et al.* Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet* **372,** 1881–1893 (2008).

27. Kibuuka, H. *et al.* A phase 1/2 study of a multiclade HIV-1 DNA plasmid prime and recombinant adenovirus serotype 5 boost vaccine in HIV-Uninfected East Africans (RV 172). *J INFECT DIS* **201,** 600–607 (2010).

28. Mwau, M. *et al.* A human immunodeficiency virus 1 (HIV-1) clade A vaccine in clinical trials: stimulation of HIV-specific T-cell responses by DNA and recombinant modified vaccinia virus Ankara (MVA) vaccines in humans. *J Gen Virol* **85,** 911–919 (2004).

29. Gilbert, P. *et al.* Magnitude and breadth of a nonprotective neutralizing antibody response in an efficacy trial of a candidate HIV-1 gp120 vaccine. *J INFECT DIS* **202,** 595–605 (2010).

30. Rerks-Ngarm, S. *et al.* Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection

in Thailand. *N. Engl. J. Med.* **361,** 2209–2220 (2009).

31. Trkola, A. *et al.* Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *Journal of Virology* **70,** 1100–1108 (1996).

32. Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326,** 285–289 (2009).

33. Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329,** 811–817 (2010).

34. Zhou, T. *et al.* Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* **445,** 732–737 (2007).

35. Cardoso, R. M. F. *et al.* Broadly Neutralizing Anti-HIV Antibody 4E10 Recognizes a Helical Conformation of a Highly Conserved Fusion-Associated Motif in gp41. *Immunity* **22,** 163–173 (2005).

36. Muster, T. *et al.* Cross-neutralizing activity against divergent human immunodeficiency virus type 1 isolates induced by the gp41 sequence ELDKWAS. *Journal of Virology* **68,** 4031–4034 (1994).

37. Nelson, J. D. *et al.* An affinity-enhanced neutralizing antibody against the membrane-proximal external region of human immunodeficiency virus type 1 gp41 recognizes an epitope between those of 2F5 and 4E10. *Journal of Virology* **81,** 4033–4043 (2007).

38. Ofek, G. *et al.* Elicitation of structure-specific antibodies by epitope-scaffolds. *Proceedings of the National Academy of Sciences* **107,** 17880–17887 (2010).

39. Correia, B. E. *et al.* Computational Protein Design Using Flexible Backbone Remodeling and Resurfacing: Case Studies in Structure-Based Antigen Design. *Journal of Molecular Biology* **405,** 284–297 (2011).

40. Correia, B. E. *et al.* Computational Design of Epitope-Scaffolds Allows Induction of Antibodies Specific for a Poorly Immunogenic HIV Vaccine Epitope. *Structure* **18,** 1116–

1126 (2010).

41. Azoitei, M. L. *et al.* Computation-Guided Backbone Grafting of a Discontinuous Motif onto a Protein Scaffold. *Science* **334,** 373–376 (2011).

42. Zolla-Pazner, S. *et al.* Cross-Clade HIV-1 Neutralizing Antibodies Induced with V3-Scaffold Protein Immunogens following Priming with gp120 DNA. *Journal of Virology* **85,** 9887–9898 (2011).

43. Moseri, A. *et al.* An optimally constrained V3 peptide is a better immunogen than its linear homolog or HIV-1 gp120. *Virology* **401,** 293–304 (2010).

44. Guenaga, J. *et al.* Heterologous Epitope-Scaffold Prime□Boosting Immuno-Focuses B Cell Responses to the HIV-1 gp41 2F5 Neutralization Determinant. *PLoS ONE* **6,** e16074 (2011).

45. Li, Y. *et al.* Broad HIV-1 neutralization mediated by CD4-binding site antibodies. *Nature Medicine* **13,** 1032–1034 (2007).

46. Pitisuttithum, P. *et al.* Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. *J INFECT DIS* **194,** 1661–1671 (2006).

47. Gilbert, P. B. *et al.* Correlation between immunologic responses to a recombinant glycoprotein 120 vaccine and incidence of HIV-1 infection in a phase 3 HIV-1 preventive vaccine trial. *J INFECT DIS* **191,** 666–677 (2005).

48. Kim, M. *et al.* Antibody mechanics on a membrane-bound HIV segment essential for GP41-targeted viral neutralization. *Nat Struct Mol Biol* **18,** 1235–1243 (2011).

Chapter 2: Structural insights into activation of antiviral NK cell responses

Introduction

Natural killer (NK) cells play key roles in combating infections with many viruses, including human immunodeficiency virus (HIV), influenza virus (IV), hepatitis viruses, poxviruses, and herpesviruses, by *i*) directly lysing infected cells and *ii*) by promoting antiviral adaptive immune responses through interactions with dendritic cells (DCs) and through the release of cytokines (1,2,3). NK cells are large granular lymphocytes with cytotoxic activity. In humans, NK cells can be divided into subsets, based on expression levels of two cell-surface markers, CD56 and CD16 ($CD56^{dim}CD16^{+}$ *versus* $CD56^{bright}CD16^{-}$ NK cells), which differ in their effector functions and homing properties (reviewed in 4). Around 90% of NK cells found in the peripheral blood and spleen are $CD56^{dim}CD16^{+}$ and likely develop from $CD56^{bright}CD16^{-}$ precursors (5). $CD56^{dim}CD16^{+}$ NK cells are cytotoxic and express high levels of perforin and the low-affinity Fcγ receptor CD16. $CD56^{bright}CD16^{-}$ NK cells, in contrast, predominate in the secondary lymphoid tissues (lymph nodes and mucosa-associated lymphoid tissues) and are copious producers of cytokines, but are only weakly cytotoxic.

NK cell anti-viral functions are governed by the integration of potentially opposing signals received through a repertoire of germ-line encoded, activating or inhibitory, cell surface receptors that recognize and respond to the presence or absence of ligands on virally infected cells and tumor cells (Table 1). With the exception of CD16, which recognizes the Fc portion of an antibody and mediates antibody-dependent cell-mediated cytotoxicity (ADCC), full activation

of NK cell effector functions requires stimulation through at least two receptors, or through one receptor plus cytokine stimulation (6,7). Activating and inhibitory receptors are thought to signal based on the missing-self and non-self/altered-self hypothesis. In the missing-self hypothesis, downregulation of human leukocyte antigen (HLA) proteins, encoded at three loci within the major histocompatibility complex (MHC), HLA-A, -B & -C, constituting the 'classical' MHC class I proteins, or the 'nonclassical' MHC class I protein HLA-E, leads to the loss of inhibitory receptor interactions and subsequent disinhibition of cytotoxic activity (8-11). MHC class I proteins are heterodimers composed of the membrane-spanning and highly polymorphic heavy chain (α-chain) associated with the non-polymorphic light-chain β_2-microglobulin (β_2m) (Figure 1). They present peptide fragments, generally derived from endogenously expressed proteins, within a groove formed by the α1 and α2 sequence domains of the heavy chain (together forming the α1/α2 'platform' structural domain); this allows the adaptive immune system to survey the proteome of any given cell through interactions with T cell receptors (TCRs). Viruses and tumors will downregulate class I proteins in order to evade T cell-mediated responses; this action, however, leaves them susceptible to NK cell attack through missing-self recognition. In the non-self or altered-self hypothesis, activating receptors recognize MHC class I proteins presenting viral or stress associated peptides or directly detect intact, virally-encoded proteins (12-16).

A historical perspective

We are fortunate to be able to take for granted our understanding of the molecular underpinnings of MHC restriction and antigen presentation to TCRs. However, the process of achieving that understanding illustrates the power and efficiency of structural biology to inform us about the details of protein function recalcitrant to alternate experimental approaches. One

of the authors (RKS) is able to provide a perspective, as a side-line observer to that process (a junior graduate student in the mid-1980's), that is particularly apropos. The first MHC class I crystal structure (of HLA-A2) was determined by Pamela Bjorkman in 1987 in Don Wiley's group at Harvard University (17) (see 18 for the details of the story). The contemporaneous first-year, graduate-level immunology course at Harvard was taught by the eminent John Kimball using his own, recently published textbook (19). The class covered the consensus understanding of antigen presentation at that time (for example, reviewed in 20), which included receptor, receptor-and-a-half and two receptor models, with antigen, in an as yet undefined form, somehow associated with 'restricting elements' (HLA proteins) for presentation at the cell surface. Models of the process proposed that restricting elements bound intact or minimally-processed antigens, raising the very significant question of how the limited repertoire of HLA proteins in a given host could bind to the huge array of potential antigens in a way consistent with the rules of protein recognition, even as understood at that time (the suggestion that 'antigen fragmentation' may be both necessary and sufficient for presentation by restricting elements was first forwarded in 1983 (21)). The conflict between proposed models of antigen presentation to TCRs and established rules of protein recognition was palpable, as was the tenuous nature of communication between the disparate disciplines. The initial details of the HLA-A2 structure, including the now famous observation that '[a] large groove between the α-helices provides a binding site for processed foreign antigens', revealed the logical and elegant mechanism of peptide presentation that we now understand and appreciate. The HLA-A2 structure (22) remains perhaps one of, if not the best, example of how a single, timely protein structure can not only resolve a scientific conundrum, but rewrite entire fields. The impact of the HLA-A2 structure also cemented the ties between the structural and immunological communities that have since resulted in many significant advances. In this review, we endeavor to identify systems where structural approaches have been particularly fruitful and questions that may yet also be resolvable by the application of modern structural molecular immunology,

particularly in regard to ligand recognition by NK receptors and the strategies employed by viruses to evade them.

MHC class I specific receptors

TCRs bind complexes of polymorphic MHC class I proteins and antigenic peptides with the capability of recognizing even small changes in the peptide or HLA protein sequences. The huge array of specificities needed is generated by blending combinatorial gene segment rearrangement and recombination (to generate a large repertoire of unique receptors) with plastic binding sites engineered to sample a wide range of conformers (to incorporate functional polyspecificity (23) into each receptor). Multiple lines of evidence also argue that TCRs and MHC class I proteins have co-evolved as interaction partners (24-28). However, NK cell receptors are germline encoded and, with few exceptions, display conventional rigidities. Therefore, fundamental questions to be answered by structural studies of NK receptor interactions with MHC class I proteins include understanding *i*) whether and how these interactions can encompass HLA class I protein diversity without the mechanisms enabling TCR polyspecificity and *ii*) how these interactions *in toto* might restrict or affect MHC class I evolution in regards to its role as a ligand for a panel of divergent receptors on various cell types.

The killer-cell immunoglobulin-like (KIR) family of receptors

The highly polymorphic KIR family of receptors is encoded on chromosome 19q13.4 within the leukocyte receptor complex (LRC) and expressed on NK and T cells. Members of the KIR family are Type I transmembrane glycoproteins that can have ectodomains comprising two (KIR2D…) or three (KIR3D…) immunoglobulin (Ig)-like domains (named D0, D1, and D2), can delivery

inhibitory or activating signals upon ligand engagement, and share >90% sequence identity in their extracellular domains (29). Inhibitory KIRs (designated by an “L”) have long cytoplasmic tails that contain two immunoreceptor tyrosine-based inhibitory motifs (ITIMs) and therefore have the capacity to inhibit cellular activity. Activating KIRs (designated by an “S”) have short cytoplasmic tails that contain a positively charged residue in the transmembrane (TM) region. This residue (Arg or Lys) interacts with a complementary charged residue in the TM of the immunoreceptor tyrosine-based activation motif (ITAM)-containing adaptor molecule DAP12 to deliver activating signals (30).

Inhibitory KIRs are known to bind various HLA-A, B, and C alleles; however, the ligands for most activating KIRs are unknown. While both KIRs and HLA molecules are highly polymorphic, HLA proteins contain certain shared motifs that mediate KIR recognition. The HLA-C C1 and C2 epitopes are defined by a sequence dimorphism (Lys/Asn) at position 80, situated on the α1 helix near the C-terminal end of the peptide binding cleft, which is complemented by a corresponding dimorphism (Met/Lys) at position 44 of KIR2D isoforms (31,32). The HLA-A/B Bw4 motif comprises residues 77-83 on the α1 helix of HLA-A and HLA-B molecules, and NK cell specificity is largely determined by the identity of the residue at position 80 (33,34). Two domain KIRs recognize the C1 and C2 epitopes whereas three domain KIRs recognize the Bw4 motif.

As with MHC class I proteins, KIR molecules and KIR/HLA combinations are highly correlated with disease susceptibility and outcome (35). While the majority of studies have focused on KIR in HIV-1, a role for KIR in immune responses to many other viruses has also been established, including hepatitis C virus (HCV) (36), hepatitis B virus (HBV) (37), Human Cytomegalovirus (HCMV) (38), herpes simplex virus type-1 (HSV-1) (39), and Epstein-Barr virus (EBV) (15). Studies on the role of KIRs in AIDS have identified the activating receptor KIR3DS1, its paired

inhibitory allele KIR3DL1, and HLA-B Bw480I (HLA-B alleles expressing the Bw4 epitope specifically with an isoleucine at position 80) as providing protective effects against HIV-1 pathogenesis. KIR3DL1 (97% identical to KIR3DS1) specifically binds HLA-B Bw480I complexes (40,41). Due to the close homology of KIR3DS1 to KIR3DL1, KIR3DS1 has been predicted to also recognize HLA-B Bw480I ligands. Evidence for the interaction of KIR3DS1 with HLA-B Bw480I comes from numerous genetic association studies that show that expression of KIR3DS1, either alone or in combination with HLA-Bw480I, is associated with a beneficial outcome during HIV infection. (42-48). These observations are supported by the finding that NK cells expressing KIR3DS1 are preferentially activated and lyse HIV-1 infected target cells in an HLA-B Bw4-80I dependent manner (44,49). However, demonstrating a direct interaction between activating KIR receptors and HLA/peptide complexes biochemically has remained elusive, at least partly due to the extreme difficulty in expressing soluble forms of activating KIR receptors suitable for biochemical studies, an observation that we can personally attest to.

During HIV-1 infection, mutations that map to T cell epitopes can also affect recognition of HLA by KIR3DL1. In one study (50), distinct HIV-1 epitopes differentially modulated the binding of KIR3DL1 to HLA-Bw4. Other studies have reported that HIV mutations emerging early during infection within an HLA Bw4 restricted T cell epitope abrogate binding to KIR3DL1 (51) and KIR2DL1 (52). Furthermore, changing the peptide presented by the HLA molecule can more efficiently abolish the inhibitory response than downregulation of HLA alone (53). These results suggest that detection of T cell escape variants by NK cells could contribute to the protective effect of the KIR3DL1/HLA-Bw4 compound genotype (50).

Major questions recently resolved or currently unresolved regarding KIR function, potentially addressable by structural approaches, include: understanding whether and how differences in peptide sequences presented by MHC class I proteins are discerned; understanding whether

and how recognition mechanisms of two- and three-domain KIRs differ; and identifying the ligands for activating KIRs (and why such highly-homologous receptors would have distinct specificities).

The structures of KIR2DL1 in complex with HLA-Cw4 and KIR2DL2 in complex with HLA-Cw3

While the two available KIR2D/HLA complex structures have been around long enough to have been extensively reviewed before (54-59), several salient details are worth summarizing to place more recent results in context. The two available KIR2/HLA complex crystal structures, of KIR2DL1/HLA-Cw4 (60) and KIR2DL2/HLA-Cw3 (61), display essentially the same overall arrangement of domains (Figure 2). The two complexes are arranged with the KIR moieties sitting on HLA in an orthogonal docking orientation to the HLA binding cleft, similar to that observed in TCR/HLA complexes, though more skewed to the C-terminal end of the groove, with the D1/D2 domains contacting the α1/α2 helices of the MHC peptide binding cleft and positions seven (P7) and eight (P8) of the bound peptide (nonamers in both cases). The KIR/HLA interface is dominated by charge complementarity; six loops on the electronegative binding surface of KIR (three from D0, two from D2 and one from the hinge loop connecting these two domains) interact with the electropositive binding surface on HLA, resulting in the formation of a network of hydrogen bonds and salt bridges. Specificities of KIR2DL1 for the C2 epitope and KIR2DL2 for the C1 epitope are achieved through different mechanisms: KIR2DL1 contains a shape-complementary pocket for Lys80 in HLA-Cw4, with the charged primary amine of the Lys80 side-chain contacting the polar side-chains of Ser184 and Glu187, and the aliphatic portion of the Lys80 side-chain contacting the apolar side-chain of Met44, in KIR2DL1. KIR2DL2, on the other hand, recognizes Asn80, which extends from the surface of HLA-Cw3 to

a much smaller degree than Lys80 in HLA-Cw4, through a hydrogen bond to Lys44 of KIR2DL2. In both complex structures, contacts between KIR moieties and the antigenic peptide specific for side-chains are sparse, with tenuous van der Waals contacts made by KIR residue Leu104 in both complexes and Gln71 in the KIR2DL2 complex, and only to the penultimate residue in the peptide. Mutational analyses have identified the interactions that contribute most to KIR/HLA binding; all of the mutations made to either the receptor or the peptide at positions shown to participate in contacts across the interface greatly reduced or completely abolished binding affinity (62). This low tolerance to mutation at the interface, however, is likely partly a reflection of the low observed equilibrium dissociation constants, in the 10 μM range, which provide little headroom for quantifying affinity reductions.

Are three domains better than two: KIR3DL1 recognition of HLA-B*5701

The first crystal structure of a three domain KIR, KIR3DL1, was recently determined in complex with HLA-B*5701 (Figures 2 and 3) (63). KIR3DL1 binds to HLA-B*5701 with its D1 and D2 domains positioned over the C-terminal end of the peptide binding cleft in an overall arrangement very similar to that in the KIR2D/HLA complexes. The D0, D1, and D2 domains together trace a zigzag path through KIR3DL1 bound in the complex, which allows D0 to extend down over the edge of the HLA α1 domain to interact directly with both the α1 and β_2m domains of HLA-B*5701. The long axis of the D0 domain is oriented almost perpendicular to the long axis of the peptide-binding cleft; this allows KIR3DL1 to make contacts with a relatively conserved region of the peptide-binding platform domain (outside of the polymorphic peptide- and TCR-binding surfaces), and extend to make contacts with invariant β_2m, providing, overall, ~30% of the binding surface area. Since this interaction involves residues conserved on both sides of the interface, a portion of HLA relatively invariant across multiple alleles and a conserved segment

of KIR, including many contacts to main-chain atoms, this interaction may be conserved across other three domain KIRs binding to HLA proteins. The surface on the peptide/HLA complex contacting the KIR3DL1 D1-D2 domains is relatively flat and allows for close positioning on HLA-B*5701; in particular, the D2 binding site is highly shape complementary to its cognate surface on HLA, excluding interfacial water molecules, where the D1 binding site is considerably less so. Analogous to the two domain KIR structures, the D1 domain makes contacts with the α1 helix and the peptide, and the D2 domain makes contacts with the α2 helix. Key contacts are made to a region of the HLA α2 helix also relatively conserved across alleles, and alanine substitution at any of these KIR D2 domain residues (Tyr200, Phe276, Glu282) or HLA-B*5701 residues (Ile142, Lys146, Ala149) abrogates binding of KIR2DL1 to HLA-Bw480I. Perhaps somewhat surprisingly at first glance, alanine substitutions in D1 (Lys136, Gly138, Ser140, Met165, Leu166, Ala167) at residues which make contact to the Bw4 epitope-defining residues, the epitope that imparts specificity of KIR3DL1 recognition of HLA alleles, did not greatly affect binding. This raises questions about the nature of the structural mechanism for specificity of this epitope. In HLA-B*5701, the side-chain of Ile80, a key residue associated with KIR3DL1 recognition, is positioned within a shallow depression ringed by the side-chains of Glu76 and Arg83 from the Bw4 motif on the α1 helix; Tyr84, also on the α1 helix; the C-terminal carboxylate and the side-chain of the penultimate residue (SerP8) of the bound peptide (LSSPVTKSF); and the side-chain of Lys146 on the α2 helix. The side-chain of Ile80 makes only one direct contact to KIR3DL1, a van der Waals contact to the side-chain of Leu166. Otherwise, Ile80 sits underneath a water-filled cavity at the KIR/HLA interface, where the ordered water molecules are also multiply-coordinated by the side-chains of Glu76 and Arg83, which penetrate into the water cluster. Alanine substitutions at any of the residues from the B*5701 α1 helix ringing this water filled cavity (Glu76, Arg79 and Arg83), all three of which also make direct contacts to KIR3DL1, severely reduced or abrogated binding. Alanine substitution at Ile80 also abrogated binding, and mutation to Ile80 to Thr, a common dimorphism found in

the Bw4 motif, reduced binding by approximately 40%. The side-chain specific, direct protein-protein contacts between KIR and HLA involving the Bw4 residues Arg79 and Arg83 are provided by *main-chain* atoms of KIR residues Gly138 and His278, partly explaining the apparent insensitivity of KIR3DL1 binding to mutations at Bw4 contacting residues.

The observation that KIR3DL1 is capable of considerable discrimination of the identity of the penultimate residue of the MHC-bound peptide (50,51,63) is difficult to reconcile with the complex structure. The lone KIR contact with the peptide at position eight (SerP8) is through a single side-chain to side-chain van der Waals contact with Leu166 of the D1 domain; longer P8 side-chains would also potentially be capable of direct interactions back to HLA. Experimental measurements of the effect of P8 substitutions on 3DL1 binding (in either B*5701 (63) or B*5703 (51) backgrounds) show that many replacements were tolerated with relatively small (less than ten-fold) changes on 3DL1 affinity, while others (Gln, Leu, Asp, Lys, Glu) much more greatly reduced affinities. Depending upon rotamer selection, there is room to accommodate various amino acids at P8, since this residue sits at the edge of the KIR/HLA interface with access to solvent. The puzzle is in the apparent lack of a pattern of allowed/disallowed amino acids; charge, hydrophobicity and side-chain bulkiness do not appear correlated with affinity changes. It is also not immediately apparent how such fine discrimination is possible without a larger number of neighboring contacts to provide the means to discriminate. There is a small pocket immediately adjacent to Leu166, bounded by Ser115, Pro199 and Glu282, that could accommodate longer P8 side-chains, seemingly almost ideal for lysine or arginine substitutions, both of which can be reasonably modeled into the complex – but lysine is among the most disfavored amino acids. Therefore, the nature of the selection rules, and their structural basis, remains unclear and merits further structural analyses.

Commentary on the need for experimentally-determined structures and the limits of modeling: the role of the D0 domain in KIR3D binding

Prior to the determination of the crystal structure of the complex of KIR3DL1 with an MHC class I protein, an extensive effort was undertaken to model the complex structure (Figure 4), precipitated by the observation that a set of sequence dimorphisms between KIR3DL1 alleles (*001, *005 and 015) affected binding to a particular MHC class I ligand, HLA-A*2402 complexed with a peptide from the HIV nef2 protein (64). It was clear, prior to these studies, that the D0 domain of KIR3D receptors contributed to MHC class I ligand binding (65). Though details of the interaction were lacking, prior alanine-scanning mutagenesis studies had, for instance, suggested that residues in the 47 to 54 region of the D0 domain affected HLA Bw4 binding (40). [KIR3DS1 does not bind HLA-A24nef likely because the nef peptide has cysteine at P8 which would sterically occlude 3DS1 binding, as discussed in greater detail below.] The effect of the sequence dimorphisms and domain swaps on ligand binding was determined and antibody binding was used to assess affects on inter- and intra-domain conformational changes. The sum total of these exhaustive mutagenesis and phylogenetic studies generated a list of constraints for positioning models of the D0 domain against the D1 and D2 domains (logically presumed to interact in the HLA complex with KIR3D receptors as they do with KIR2D receptors) and the A24nef ligand. For instance, the KIR D0 47 to 54 region was presumed to directly contact HLA. The D1 and D2 domains of KIR3DL1 (the *015 allele) were modeled on the crystal structure of KIR2DL1 domains D1 and D2 from the HLA-Cw4 complex structure. The D0 domain was modeled on the crystal structure of the D1 domain of KIR2DL2 from the HLA-Cw3 complex structure. Their methods employed multiple sequence alignments to identify the best homology model, moderately sophisticated computational modeling tools to adjust loops and accommodate sequence substitutions, and energy minimization to tweak the docking of

receptor domains onto the MHC class I ligand. The model was then thoroughly validated by additional mutagenesis/binding studies, producing a compelling structural analysis, one that we felt had gotten it right when we read it. In this model (Figure 4), the D0 domain packs tightly in the corner between the D1 and D2 domains, with KIR3DL1 perching on the C-terminal end of the peptide-binding groove like a tricorne hat, a quite logical and satisfying solution (64).

Unfortunately, despite the careful application of logical methodology and exhaustive experimental data, the modeled complex does not recapitulate many of the salient details of the overall domain organization of KIR3DL1, the structure of the D0 domain, or the details of the intermolecular interactions between the D0 domain and the HLA ligand (Figure 4). For instance, in the complex structure, the KIR D0 47 to 54 region does not contact HLA, but points completely away from ligand out into space, forming one of two short α-helices in the D0 domain that were not predicted during modeling. While we hail the efforts of these authors, who did yeoman's duty in this analysis, the pitfalls that likely affected the outcome provide a valuable cautionary tale. First, antibody binding is, at best, a very coarse measure of protein conformation. Second, it is very difficult in forward mutagenesis studies to distinguish between direct effects on contact residues from indirect effects on folding, or effects communicated at a distance to contact residues. Also, because of the uneven distribution of binding energy at protein-protein binding sites, direct contact residues may not show dramatic reductions in affinity when mutated to alanines. Third, homology modeling at low sequence identities (the D0 domain was 38% identical to the starting structure used for modeling) is always perilous unless using the most sophisticated computational algorithms available today (and sometimes even then), which are often restricted to dedicated computational biology groups because of the processor resources, and arcane computer code, required. Protein-protein docking still remains challenging for computational biology, even at the current state of the art, with incomplete tools for validating predicted results, demonstrating the continuing value in determining complex

structures experimentally.

Modeling KIR3DS1 interactions to define the unknowns

The physiological ligand for the activating counterpart of KIR3DL1, KIR3DS1, is not yet known but is hypothesized to be HLA-B Bw4 (49) (44), though studies have failed to demonstrate a direct interaction (66,67). As with most KIR activating/inhibitory pairs, the sequence identity between isoforms is quite high; 3DS1*010 and 3DL1*001 differ at only seven positions: I47V, V92M, S58G, G138W, L166R, P163S, P199L. We have modeled a hypothetical complex between KIR3DS1*010 and HLA-B*5701 by replacing residues corresponding to the 3DS1/3DL1 substitutions in the KIR3DL1*001/HLA-B*5701 crystal structure (Figure 3). Three of the seven 3DS1/3DL1 substitutions (S58G, V92M, I47V), while predicted to be on the surface of KIR3DS1*010 D0 domain, are distant from the KIR3DL1*001/HLA-B*5701 interface and should not affect MHC class I binding directly or indirectly. Three substitutions (G138W, P199L, L166R) occur at positions where residues make direct contacts with HLA-B*5701 in the KIR3DL1*001 complex structure and one substitution (P163S) occurs in the D1 loop (161-166). Proline, instead of serine, at this position in 3DL1*001 may serve to reorder the D1 loop and affect the positioning of residue 166, a potential indirect effect on binding. Experimentally swapping the 3DS1*010 residues into 3DL1*001 at the positions directly contacting HLA-B*5701 showed that two substitutions (G138W, P199L) did not result in a substantial decrease in HLA-B*5701 tetramer binding (63). In another study, substitutions in KIR3DL1*015 to those present KIR3DS1*010 either completely abolished (P163S, L166R, P199L) or severely decreased binding (G138W) to HLA-A24 (64). In the modeled complex, there is space at the interface to accommodate these substitutions without affecting MHC class I interactions. However, replacement of leucine 166 (in KIR3DL1) with arginine (in KIR3DS1) abrogates tetramer binding

(63). The side-chain of L166 is positioned directly above the P8 residue; in the KIR3DL1/HLA-B*5701 complex structure, the L166 side-chain makes a van der Waals contact with the hydroxyl of the P8 serine residue. In our model, R166 makes substantial steric clashes with any non-glycine residue modeling into the P8 position, with the caveat that the global KIR/MHC class I domain arrangement is held rigid in these analyses (smallish domain-by-domain rearrangements at the interface might be able to accommodate some of the less severe clashes). A glycine at P8 also sterically clashes, but to a minimal degree that could be accommodated by minimal adjustment of the KIR/MHC class I interaction at the local level. We also note that an alignment of all known KIR3DS1 and KIR3DL1 alleles shows that no residue is wholly unique to either KIR3DS1 or KIR3DL1 (for instance, even R166 can be found in KIR3DL1*054). This suggests that not all KIR3DL1 will bind HLA-A/B and that a combination of substitutions likely accounts for the absence of KIR3DS1 binding to HLA-A/B.

HLA-B*5703 structures complexed with three of five known HIV p24-derived T cell, B*57-restricted epitope peptides are available (68): ISPRTLNAW (ISP), KAFSPEVIPMF (KAF-11), and KAFSPEVI (KAF-8). These peptides vary widely in length and primary sequence and adopt markedly different conformations and associated ordered water structures in the HLA-B*5703 peptide-binding groove. Docking these complex structures onto HLA-B*5701 in the KIR3DL1 complex structure or the KIR3DS1 complex model (a valid exercise, since there are no sequence differences between these two class I sub-alleles within 12 Å of the peptide C-terminal residue directly underlying KIR) shows that the only irreconcilable steric clash with 3DL1 is with the penultimate methionine residue in the KAF-11 peptide, despite the wide range of conformations displayed; all three peptides would be predicted to clash with 3DS1.

The leukocyte immunoglobulin-like receptor (LILR) family

The LILRs (also leukocyte inhibitory receptors (LIRs), immunoglobulin-like transcripts (ILTs) or CD85a-m) are Type I transmembrane glycoproteins with either two or four tandem Ig extracellular domains and, like KIR, are encoded within the LRC, consist of both activating and inhibitory forms, and recognize MHC class I proteins (69-72). Inhibitory LILRs signal through ITIMs in their cytoplasmic tails while activating LILRs associate with the ITAM containing adaptor signaling protein FcεRIγ through a positively-charged arginine in their TMs (73). These receptors are expressed on NK cells, T cells, granulocytes, plasmacytoid DCs and cells of the myeloid lineage (74). Of the 13 LILR genes (two of which are pseudogenes), LILRA1, LILRA2, LILRA3, LILRB1, and LILRB2 are known to, or predicted to, bind classical MHC class I proteins, whereas LILRA4, LILRA5, LILRA6, LILRB3, LILRB4, and LILRB5 do not, or are not predicted to, recognize MHC I (74). LILRB1 and LILRB2 also bind the less polymorphic, nonclassical MHC class I proteins HLA-E, HLA-F, and HLA-G (75-77). HLA-E normally presents a restricted peptide repertoire derived from the leader sequences of other MHC class I proteins (78), and so serves as an indirect check for normal class I expression, but, under conditions of cellular stress, can present peptides derived from heat-shock proteins, pathogen-associated proteins or the defective processing of antigens (79). HLA-F, while associated with many disease states, is much less understood, but has been shown to be a cell-surface marker of activated lymphocytes (80) and to bind free HLA class I heavy chains when it is itself in a peptide-free state (81) (no HLA-F-binding peptides have been definitively identified to date). HLA-G is highly expressed on the trophoblast and plays a key role in maintaining tolerance at the feto-maternal interface (82), but has since been demonstrated to also have a tolerogenic function in a variety of pathological conditions (83).

The potential importance of LILRs in immune responses to viral infections beyond their role as

class I sensors was suggested by the finding that LILRB1 can act as a receptor for the HCMV protein UL18, a viral HLA class I homolog (69,84,85); LILRB1 binds to UL18 with 1000-fold greater affinity than to MHC class I proteins (86). While it might be expected that engaging an inhibitory LILR through a HLA class I homologous viral decoy ligand on the surface of an infected cell would inhibit cell lysis, the role in NK cell mediated lysis of HCMV infected cells is unclear: UL18 has been found to either decrease or increase target cell lysis by NK cells depending on the context (87). However, an evolutionary analysis of primate lineages showed that some NKG2-type NK receptors (discussed below) are evolving under positive selection. Such signatures of positive selection suggest that the protein involved is in genetic conflict with, for instance, a pathogen-associated evasion protein. Following on this observation, a search for a possible viral NKG2 interactor led to the discovery that UL18 also binds to the activating NK receptor NKG2C (88), which complicates the picture of how UL18 expression may affect NK cell activation.

LILRs may also play roles specifically in HIV-1 infections. Although LILRs have been shown to bind many HLA class I alleles, there is evidence that the affinity is influenced by allelic polymorphisms. For example, allelic subtypes of HLA-B*35 can be grouped into Px and Py subsets, differing by as few as a single amino acid, but which differentially influence HIV-1 disease progression: the Px subtype is a strong predictor of accelerated HIV-1 disease progression, while the Py allele has no effect (89). Since the Px and Py subsets are highly homologous, often with completely conserved peptide repertoire specificities, it seems unlikely that differences in disease progression would be attributable to CD8+ T cell responses (90). Supporting the alternative hypothesis that the effect is NK mediated, LILRB2 has been shown to bind more strongly to HLA-B*35 Px than HLA-B*35 Py subtypes as measured by tetramer staining and surface plasmon resonance (SPR) biosensor analyses (91). In this latter study, binding through LILRB2 expressed on DCs to HLA-B*35 Px-bearing cells resulted in decreased

maturation and cytokine secretion. In a similar vein, HLA-B*57 and HLA-B*27, two alleles that have consistently been associated with better control of HIV-1 disease progression (92), have been shown to have the weakest binding to LILRB2 when compared to over 90 other HLA class I alleles (93). As seen with KIR recognition of HLA, peptide variation may also modulate LILR recognition of MHC class I proteins. In one example, a typical escape mutation in an HLA-B*2705-restricted cytotoxic T cell epitope from HIV-1 gag (KRWIILGLNK, escape: L6M) was shown to substantially enhance LILRB2 binding of HLA-B*2705 tetramers presenting the mutant peptide (94). While these studies are highly suggestive, they have been performed either with DCs or biochemically, so direct associations between AIDS disease progression and LILRB2 signaling on NK cells remains to be demonstrated.

Structures of the LIRLB1/HLA-A2 and LILBR2/HLA-G complexes: a conservative viewpoint

The broad specificity of LILRB1 and LILRB2 for numerous HLA alleles is explained by the conserved nature of their recognition sites on MHC class I proteins (Figure 2). While KIR receptor footprints encompass the highly polymorphic α1/α2 platform domain of MHC class I proteins, LILRB1 and LILRB2 bind to a surface composed of sections of the relatively non-polymorphic α3 domain and the invariant β_2m light chain (74). The complex crystal structures for the D1-D2 domains of LILRB1 and LILRB2 bound to HLA-A2 and HLA-G respectively, show very similar binding modes, with two main interaction surfaces: D1 makes contacts with the HLA α3 domain, while the hinge region linking the two domains makes contacts with β_2m (95,96). LILRB1 and LILRB2 interactions with the MHC class I α3 and β_2m domains bury roughly comparable surface areas at the interfaces, with more area buried at the β_2m interfaces in both examples, but with binding energy more evenly distributed between the α3 and β_2m interactions

(LILRB2: 471 Å^2 buried on α3 (ΔG_{calc} = -3.2 kcal/mol), 635 Å^2 on β_2m (ΔG_{calc} = -4.1 kcal/mol); *versus* LILRB1: 272 Å^2 buried on α3 (ΔG_{calc} = -1.9 kcal/mol), 584 Å^2 on β_2m (ΔG_{calc} = -1.8 kcal/mol); values calculated with PISA (97)). The interesting observation that LILRB2, but not LILRB1, can bind to β_2m-free HLA heavy chains (14,19) is not explained by the relative surface areas buried, but is potentially explained by the difference in calculated solvation free energies of binding (ΔG_{calc}), a function of the nature of the interactions within the interfaces and not just the total area buried. LILRB1 and LILB2 are oriented head-to-tail with HLA in the complexes, with their N-terminal domains pointed towards the C-terminal, membrane proximal domains of the HLA protein. This orientation is consistent with a *trans* mode of binding, where the receptor on the effector cell engages an HLA molecule on the target cell. While the LILR D1 domains do extend below the plane of the MHC α3 domains in the complexes, the eight-residue spacer sequence between the MHC α3 and TM domains is long enough to allow LILRs to fully clear the membrane without canting over the MHC protein in this binding mode. Complexation with HLA triggers changes within the interdomain angle of LILRB1 (shifts of 14 to 19°) which are not observed in LILRB2 during binding (85,95,96,98,99). In general, however, more extensive conformational changes are seen in LILBR2 during complex formation than in LILBR1: for example, free LILRB2 contains one 3_{10} helix (residues 52 to 55) involving residues at the binding interface; upon complex formation, the interface is restructured with two 3_{10} helices (residues 46 to 50 and 53 to 57). Overall, the structure of LILRB2 in the complex state most closely resembles LILRB1 in the free state. However, these conformational changes presumed to occur concurrent with binding may also reflect the very different crystallization conditions used for the various proteins and/or the process of crystallization selecting particular conformers out of a larger ensemble of conformers accessible to flexible proteins.

While the LILR story appears to have reached consensus, with recognition apparently focused on conserved elements of MHC class I proteins, there is evidence that HLA polymorphisms and

peptide identity do play a role in modulating LILR binding (94,100). Quantitative SPR binding analyses have confirmed that the membrane-distal D1 and D2 domains are primarily responsible for ligand binding (86,95). However, the lesson of the KIR2D *versus* KIR3D interactions should not go unheeded; until structures of intact four-Ig domain LILRs, alone and in complex with HLA ligands, are available, the role that the membrane-proximal D3 and D4 domains play in recognition remains formally unknown, with the potential that these domains play minor modulatory or selective roles in binding, possibly in surprising ways.

LILRB1/UL18: recognizing a mimic is easy

HCMV UL18 is a Type I transmembrane glycoprotein homolog of human MHC class I proteins that analogously associates with β_2m and binds peptides within a groove formed by UL18's version of an α1/α2 platform domain, with the major difference that UL18 is much more heavily glycosylated than MHC class I proteins (101-103). Despite sharing only ~21% sequence identity with MHC class I molecules, the crystal structure of a minimally glycosylated UL18 construct bound to an actin-derived peptide (ALPHAILRL) in complex with the D1-D2 ectodomain of LILRB1 (84) shows a high degree of structural homology to the LILR/classical and LILR/nonclassical MHC class I complexes at the levels of overall domain organization, contact surfaces, tertiary and secondary structure. However, the orientation between LILR D1-D2 domains with respect to each other differs in detail, with LILRB1 adopting an interdomain angle in the UL18 complex intermediate between that in HLA-A2 complex and LILRB2 in the HLA-G complex. As in the LILRB1/HLA-A2 complex, the linking hinge between the D1 and D2 domains contacts β_2m (497/519 $Å^2$ buried on β_2m in the two complexes in the crystallographic asymmetric unit; values calculated with PISA (97)) while the D1 domain contacts the MHC α3 domain (579/595 $Å^2$ buried on α3). The nature of the intermolecular contacts between β_2m and

the LILRB1 D1-D2 linking hinge were very similar in both the LILRB1/UL18 and /HLA-A2 complexes, but the heavy chain/D1 interactions differed, effectively creating a more favorable binding interface for the latter, including additional salt bridges (two to HLA-A2 *versus* nine/eight to UL18), hydrogen bonds (three to HLA-A2 *versus* nine/seven to UL18), and van der Waals contacts, contrasting the van der Waals-dominated HLA-A2/D1 interface. Overall, recognition of UL18 by LILRB1 is driven by interactions involving the UL18 heavy chain, while detection of HLA is driven by recognition of the invariant β_2m light chain. Of course, this results because HCMV can only access the UL18 sequence, not the host-encoded β_2m sequence, to evolve a tighter-binding LILR decoy ligand. Known variable regions within the UL18 α1 domain previously predicted to interact with LILRB1 were not observed to do so in the complex crystal structure (104,105). It may be possible that residues in the LILRB1 D3 or D4 domains could provide such contacts, not observed in the crystal structure because these domains were not included in the crystallization construct. However, analytical ultracentrifugation studies (86) have suggested that LILRB1 adopts an extended structure that would require an unlikely bend to reach back and make additional UL18 contacts (the C-terminus of the LILR D2 domains in the complexes all point completely away from the ligand proteins). Furthermore, the affinities for UL18 of LILRB1 D1-D2 constructs *versus* intact ectodomain D1-D4 constructs did not differ significantly (86), all arguing that the D1-D2 complex structure revealed all the details of the interaction.

The extracellular domain of UL18 has 13 predicted N-linked glycosylation sites (106) and, when expressed in mammalian cells, is heavily glycosylated (~55% carbohydrate by weight). A fully glycosylated UL18 was modeled bound to LILRB1 (84). In this model, the UL18 heavy chain was completely shielded by carbohydrate except for two regions; the binding site for LILRs and the docking site for β_2m. This observation led to the hypothesis that UL18 evolved a glycan shield to prevent binding of other mediators of immunity, such as MHC class I binding $CD8^+$ T cells or antibodies, while preserving the binding surface for LILRs. This strategy parallels that

employed by other viruses, notably the comparably-glycosylated HIV envelope protein.

NKG2-CD94 heterodimers

Members of the NKG2 family of C-type lectin-like receptors, expressed on NK cells and subsets of $CD8^+$ T cells, include NKG2A, -B, -C, -E, and -H, all of which obligately dimerize with CD94 (which also has a C-type lectin-like fold) (107-110). NKG2-CD94 heterodimers recognize HLA-E molecules (111-114) and are encoded by genes found within the NK complex on chromosome 12p13 (115). NKG2A/B are inhibitory receptors that contain ITIM motifs in their cytoplasmic tails (116); NKG2C/E/H are activating receptors that associate with DAP12 through a positively charged residue in their transmembrane region (117). The nonclassical MHC class I protein HLA-E is distinct from the classical class I proteins in that it binds a restricted repertoire of peptides excised from the leader peptides of classical and nonclassical (HLA-G) MHC class I proteins (118). Because HLA-E expression is dependent on the expression of other HLA class I molecules (classical MHC class I and HLA-E proteins do not fully fold and express on the cell surface without binding an antigenic peptide), recognition of HLA-E by NKG2A/C/E-CD94 receptors allows NK cells to monitor the expression of other MHC class I molecules on the cell. The interactions between the various NKG2-CD94 receptors and HLA-E/peptide complexes span a wide range of affinities, with patterns that suggest the system has been tuned to produce particular outputs (109). Responses of NK cells to viral infections in different contexts support this idea. For example, loss of the inhibitory NKG2A signal due to downregulation of HLA-E by vaccinia results in NK-mediated lysis of virus-infected cells (119). On the other hand, infection with hantavirus (120) and HCMV (121) results in upregulation of HLA-E on infected cells, which may be expected to thwart NK cell surveillance against virus infection due to signaling through the inhibitory NKG2A receptor. However, during hantavirus and HCMV infection, there is a

specific expansion and persistence of activating NKG2C+ NK cells (122) and, with hantavirus, a decrease in the numbers of inhibitory NKG2A+ NK cells (120).

The only NKG2 ectodomain structures currently available including the complete recognition module (NKG2 moiety + CD94) are two, essentially identical, contemporaneous structures of NKG2A-CD94 bound to HLA-E (88,123) (Figure 2) and an earlier structure of NKG2A-CD94 alone (107). These structures are important for revealing views of NKG2 C-type lectin-like folds, 'lectin-like' in the sense that, while the fold is retained, specific structural elements necessary for lectin activity have been lost, allowing these receptors to target protein ligands as opposed to carbohydrates. In the complexes, CD94 binds over the α1 helix of HLA-E and contributes all contacts to the C-terminal portion of the peptide, while NKG2A sits over the α2 helix, an orientation very reminiscent of TCR/MHC interactions. Recognition of the C-terminal half of the peptide is important since almost all of the limited variation among HLA-E restricted peptides is localized to the C-terminal residues, showing that CD94, the invariant element in NKG2 heterodimers, directly reads out peptide sequence differences, which can result in moderate affinity differences (109). Modulation of affinity differences between NKG2-CD94 family members is achieved primarily through adjustments to the orientation of the two halves of the heterodimer; residues at the heterodimer interface also display the highest signatures of evolutionary positive selection, consistent with the idea that positive selection due to genetic conflict with a pathogen-associated evasion molecule would drive escape from the interaction.

In addition to these ectodomain structures, there is a rather unique structure available for the heterotrimeric assembly of TM domains from DAP12 and NKG2C, determined in its membrane-embedded form by NMR spectroscopy (120,167). In this structure, the intramembrane contacts between TMs are dominated by a network of electrostatic interactions between five polar residues. Since such polar/electrostatic TM interactions are often found between activating

immunoreceptors and their cognate adaptor molecules, this structural motif potentially serves as an example of a key molecular organizing element.

However, structural and biochemical studies have raised, rather than addressed, two puzzling aspects of the NKG2-CD94 family, pertaining to the NKG2E and NKG2F isoforms. NKG2E-CD94 has been expressed recombinantly, purified and used in quantitative SPR binding analyses, displaying a pattern of affinities seemingly tuned, through unique structural mechanisms, to match those of the inhibitory receptor NKG2A-CD94 (109). However, despite the presence of a positively-charged residue (lysine) in its TM, NKG2E-CD94 has not been demonstrated to interact with DAP12 (a requirement for activating NKG2 cell-surface expression) or to be found on the surface of any cell type, under any condition (59). NKG2F, on the other hand, which contains a cytoplasmic ITIM-like sequence, is missing a large, otherwise-conserved section in the C-type lectin-like domain, which, from a structural perspective, seems indispensible for folding (124). NKG2F has also not been found on the surface of any cell type, under any condition, not surprisingly. Therefore, the question arises as to whether these two isoforms are expressed under any conditions and, if so, how the structural hurdles to their expression are overcome. If they are not functionally expressed, the question arises as to why they are retained in the genome in a form that is potentially functional, at least for NKG2E.

Receptors specific for non-MHC class I ligands

NKG2D homodimers

NKG2D is a Type II transmembrane C-type lectin-like protein encoded by the KLRK1 (killer cell lectin-like receptor subfamily member 1) gene on chromosome 12p12-13 and is broadly expressed on NK cells and cytotoxic T cells (116,125). In humans, NKG2D transmits activating signals through the adaptor molecule DAP10 upon recognition of stress induced MHC class I chain related molecules (MICA/B) and UL16 binding proteins (ULBP1-6) (126). NKG2D ligands are highly polymorphic and resemble MHC class I molecules in terms of domain organization and fold. Unlike MHC class I proteins, NKG2D ligands do not associate with β_2m or antigenic peptides (127-130). Despite its name, NKG2D is quite distinct from the rest of the NKG2 family, showing limited sequence identity to the otherwise closely-related family members, forming homodimers rather than obligate heterodimers with CD94, and using DAP10 as an adaptor rather than DAP12. NKG2D structural analyses are perhaps most valuable for revealing a novel mechanism enabling functional recognition polyspecificity, 'rigid adaptation' (23,131,132), and are otherwise quite near-and-dear to our hearts, but have been previously reviewed and analyzed almost *ad nauseam*, so we will not waste anymore space here on the matter.

Natural Cytotoxicity Receptors (NCR): the direct approach

NCRs were discovered as the result of a hunt expressly dedicated to finding HLA-independent activating receptors (133). The three requirements that putative NCRs needed to meet were: *i*) expression should be restricted to NK cells; *ii*) monoclonal antibody (mAb) mediated

crosslinking targeting the candidate NCR should trigger NK cell lysis in a redirected killing assay; and *iii*) mAb mediated masking of the receptor should inhibit NK cell mediated cytotoxicity. Currently, this group consists of the activating receptors NKp30, NKp44, and NKp46. NKp46 maps to chromosome 19q13.42 in the LRC whereas NKp44 and NKp30 are found on chromosome 6p21.3. NCRs are Type 1 transmembrane glycoproteins, members of the Ig superfamily, and, though they do not share close sequence homology with one another or with other NK receptors, they are structurally quite similar to KIRs and LILRs. NKp46 and NKp30 are expressed on resting and activating NK cells and NKp44 is selectively expressed on IL-2 activated NK cells (134,135). NCRs contain a positively-charged residue in the TM that enables them to bind to adaptor signaling molecules: DAP12, CD3ζ, and FcεRIγ (136-137).

NCRs have received much attention for their involvement in tumor cell lysis; however, the known and characterized ligands for these receptors are often viral in origin and include the IV, Newcastle disease virus (NDV), and Sendai virus (SV) hemagglutinins (HA) (14,138,139) and HCMV pp65 (140). NCRs have also been reported to mediate protective responses during infections with HCV (141,142), HSV (143), Ebola and Marburg viruses (144), Dengue and West Nile viruses (145) and vaccinia virus (146), though the ligand/s targeted in these contexts remain unknown and even controversial. Crystal structures for the extracellular domains of NKp46 (147) (148), NKp44 (149), and NKp30 (150) have been determined. However, no ligand bound complex structures of NCRs have been reported as yet. Therefore, the polyspecificity mechanisms through which these receptors are able to bind multiple, distinct ligands, including those upregulated during tumorigenesis or infection, or encoded by pathogens, is not known.

The ectodomain of NKp30 consists of a single I-type Ig domain. The I-type domain is structurally intermediate between the V and C-type domain architectures originally identified in antibody structures. Despite limited sequence similarity, NKp30 is structurally homologous to the cell

adhesion molecule Coxsackievirus and adenovirus receptor (CAR; RMSD = 2.2 Å) and the T cell coreceptors CTLA-4 (RMSD = 2.6 Å) and CD28 (RMSD = 2.6 Å). NKp30 has been postulated to bind a wide array of ligands including HCMV pp65 (140), Duffy-binding-like domain 1α of *Plasmodium falciparum* erythrocyte membrane protein-1 (PfEMP-1) (151), leukocyte antigen-B-associated transcript 3 (BAT3) (152), heparan sulfate and heparin (153), and the B7 family member B7H6 (154), the latter biochemically verified through detailed epitope mapping and mutational analyses (150). The B7H6 binding site on NKp30 is proposed to occupy the structurally analogous position on CD28 family members that serves as the common, promiscuous binding site for B7 ligands (155,156). NKp30 is also the target of an immunoevasion mechanism mediated by the HCMV tegument protein, pp65. pp65 binds to NKp30, causing it to dissociate from its signaling adaptor, CD3ζ, thus preventing the transduction of activating signals (140).

NKp44 consists of a single V-type Ig domain in the extracellular region. A unique feature in the otherwise typical two β-sheet sandwich V domain structure is the presence of a second disulfide bond, which stabilizes the 33-51 β-hairpin. The abutment of the β-hairpin with the 87-112 β-hairpin, a region corresponding to the complementary determining region of antibody and TCR variable domains, forms a prominent groove on the protein surface, proposed as a ligand binding site. One side of the groove is lined with the side-chains of four basic residues, which creates a positive face, while the other side is neutral.

NKp46 consists of two C2-type Ig extracellular domains, D1 and D2 (148). The two closest homologs to NKp46 are KIRs, which share, at most, ~37% sequence identity, and LILRs, which share ~45% sequence identity, when comparisons are made domain-by-domain. Structural comparisons of NKp46 with KIR2DLs, LILRB1 and LIRB2 show considerable structural homology: superposition RMSDs of 1.8 Å to 2.3 Å superimposed as intact D1-D2 units and

RMSDs of 0.9 Å to 1.4 Å when D1 or D2 domains are superimposed individually. The interdomain hinge angle (~85°) and D1/D2 domain orientation are also conserved between NKP46 and KIRs/LILRs. However, comparisons between the binding sites of KIR or LILR and NKp36 show that the residues involved in ligand recognition are not conserved, consistent with the observation that NCRs do not bind MHC class I proteins.

NKp44 and NKp46 both bind IV, NDV, and SV HAs through interactions largely mediated by terminal sialic acid groups (the cellular receptor for HAs) on N-linked NCR carbohydrates. Recognition of HA by NKp46 relies heavily on the O-linked glycan on the stalk region between the Ig ectodomains and TM, particularly at residue Thr225. In NKp46, Thr225 is also responsible for recognition of, as of yet, undefined tumor ligands, suggested to act by maintaining the correct conformation of the binding site (157). This generalized mechanism of viral HA recognition may allow NKp44 and NKp46 to bind HAs from a wide range of viruses due to the conserved function of HAs to bind sialic acid-containing glycoproteins (158). This recognition mechanism is unlikely the whole story, since binding in these cases does not formally rely on a structured ectodomain, merely the property that these NCRs are glycoproteins with terminal sialic acid groups. IV has also developed an immunoevasion strategy to counter this defense. While HA is an activating ligand of NKp44 and NKp46 when expressed on the surface of infected cells, IV infected cells also release soluble HA that can be taken up by NK cells, presumably through NCR interactions. This results in lysosomal degradation of the CD3ζ chain, preventing the transduction of activating signals from NKp44 and NKp46 (159). This immune evasion mechanism likely contributes to the pathogeneses of IV.

Coda

Fine-tuning of NK receptor signaling across the array of often cross-reactive (recognizing common ligands), often opposing receptors seems to enable a precise discrimination of even slightly altered classical and nonclassical MHC class I expression levels on cells stressed as a result of any of a number of pathologies. This process is complemented by altering the repertoire of peripheral NK cells, defined by the subsets and patterns of receptors expressed on any given NK cell, to generate maximally-effective responses to infection or other pathology. Unfortunately, neither the mechanism of signal integration, nor the process of NK cell development and selection, are fully understood as yet. For instance, NK cells pass through a process known as 'licensing' during development, an 'MHC-dependent process that results in enhanced capacity of NK cells to be activated when stimulated through an activation receptor' (quoting 160). Though first described some time ago, the exact role and outcome of licensing on NK cell activity remains an active topic of discussion (160,161), illustrating how much there is yet to learn about NK cell development and receptor signaling mechanisms. Structural studies, stepping in where they are most useful, have begun to address the details of the signal transduction process by revealing the minutiae of the TM interaction between activating receptors and adaptor proteins, but much ground remains to be covered. Biophysical results, particularly recent ones, have provided crucial insights into the structural relatedness of NK receptors to the larger landscape of known protein structures and snapshot views of the recognition events mediating NK activation and viral immunoevasion. A common theme clearly employed by NK receptors to recognize diverse, polymorphic ligands is to focus attention on the most structurally conserved elements, although variations on this theme clearly play fundamental roles in the process of generating functional polyspecificity. Parsing evolutionary relationships on the basis of structural results is always dicey, but a remarkable observation

bears further comment: classical and nonclassical MHC class I proteins are multifarious recognition platforms for a whole host of receptors, with nearly every accessible surface involved in some binding event (Figure 2, which does not even include maps of the binding sites of TCRs or CD8). While they likely evolved as ligands for TCRs, based on the ancestry of this immunological mechanism, more recent immunorecognition systems, particularly those employed by NK cells (and the viruses attempting to avoid them), have made effective use of MHC class I proteins as ligands. Undoubtedly, this places new constraints on the future evolution of MHC class I proteins, providing ample grist for many more reviews on this subject in the future.

Acknowledgements

I would like to thank Roland K. Strong for his assistance with manuscript preparation

Figures

Table 1. List of human NK cell receptors: structures, signaling, and viral involvement.

activating receptor	PDB accession code	ligand	signaling	virus	references
NKG2D	1HQ8, 1HYR (NKG2D/MICA), 1KCG (NKG2D/ULBP3), 1JSK (NKG2D/Rae-1β)	MICA/B, ULBP1-6	DAP10	HCMV, KSV, HBV, HIV-1, IV	126, 128, 129, 162-165
NKG2C-CD94	2L35 (DAP12/NKG2C)	HLA-E	DAP12	HCMV, HBV, HCV, hantavirus	112, 120, 122, 142, 166, 167,
NKp46 (NCR1)	1P6F, 1OLL	HA, heparan sulfate/ heparin	FcεRIγ, CD3ζ	Influenza, vaccinia, HCV, HSV, NDV, SeV	14, 139-143, 146-148, 153
NKp44 (NCR2)	1HKF	HA, heparan sulfate/ heparin, E-protein, PCNA, bacterial	DAP12	Influenza, vaccinia, HSV, NDV, SeV, Dengue, West Nile,	138, 139, 143, 146, 149, 153, 168

		cell wall elements		HIV-1	
NKp30 (NCR3)	3NOI	B7H6, BAT3, HCMV pp65, heparan sulfate/heparin, PfEMP-1	FcεRIγ, CD3ζ	Influenza, vaccinia, HCV, HSV, HCMV, Ebola, Marburg	140, 142-144, 146, 150-154, 169, 170
NKp80	---	Activation-induced C-type lectin	AICL	KSV	163, 171
KIR2DS1	---	HLA-Cw2,4,5,6 (C2 epitope)	DAP12	EBV	15
KIR2DS2	1M4K	---	DAP12	HCV, HSV-1	172, 173
KIR2DS4	3H8N	HLA-Cw4	DAP12	HCV	174-176
KIR3DS1	---	HLA-Bw4?	DAP12	HIV	42, 44, 49
KIR2DL4	3H8N	HLA-G	FceRIg, 1 ITIM	---	177
inhibitory receptor	**PDB accession code**	**ligand**	**signaling**	**virus**	**references**
NKG2A/B-CD94	1B6E (CD94), 3BDW (CD94/NKG2A), 3CDG and 3CII (CD94/NKG2A/HLA-E)	HLA-E	1 ITIM	HCMV, HIV-1, HCV	88, 107, 112, 123, 178, 179, 180

KIR2DL1	1NKR, 1IM9 (KIR2DL1/HLA-Cw4)	HLA-Cw2,4,5,6 (C2 epitope)	2 ITIM	EBV	15, 60, 181
KIR2DL2	2DL1, 2DL2, 1EFX (KIR2DL2/HLA-Cw3)	HLA-Cw1,3,7,8 (C1 epitope)	2 ITIM	HCV, HSV-1	61, 172, 182, 183
KIR2DL3	1B6U	HLA-Cw1,3,7,8 (C1 epitope)	2 ITIM	HCV	184, 36
KIR3DL1	3VH8	HLA-A, B (Bw4 epitope)	2 ITIM	HIV	49, 63, 185
LILRB1 (ILT2)	1UFU, 1VDJ, 1G0X, 1P7Q (LIR1/HLA-A2), 3D2U (LIR1/UL18)	HLA class I, HLA-G, HLA-F, HLA-E, UL18	4 ITIM	HCMV, HIV-1	75, 84, 85, 95
LILRB2 (ILT4)	2GW5, 2DYP (LIR2/HLA-G)	HLA class I, HLA-G, HLA-F, HLA-E, UL18	3 ITIM	HCMV, HIV-1	75, 96, 99

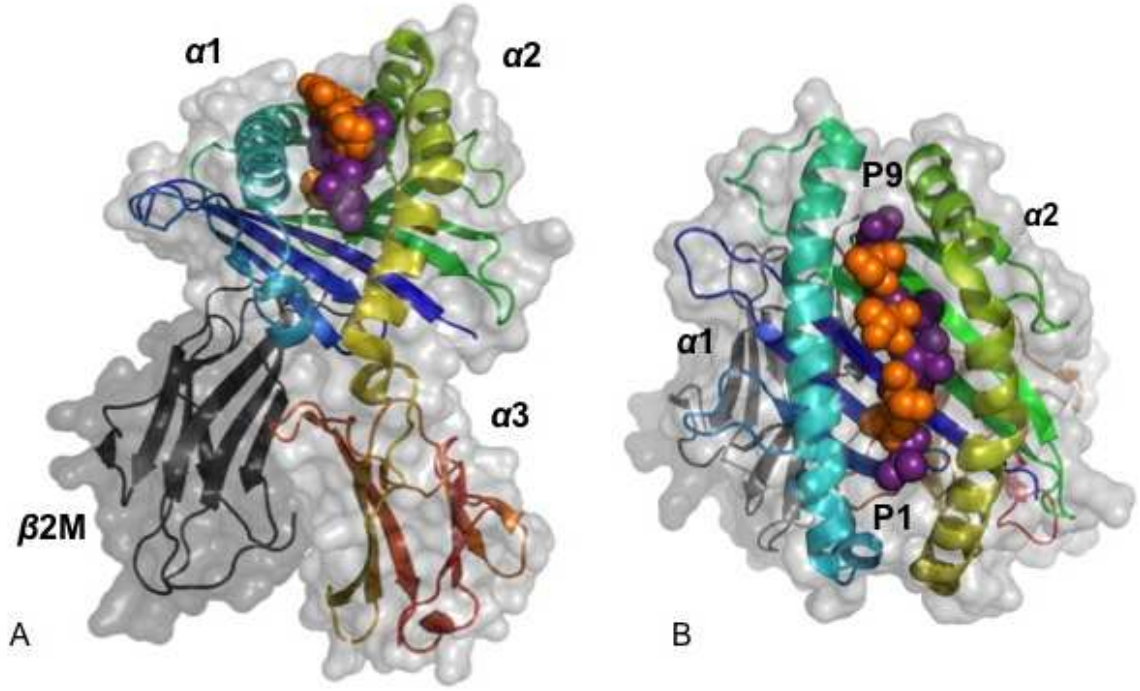

Figure 1: The structure and domain organization of MHC class I proteins. The structure of HLA-B*5701 is shown, in two orthogonal views (A and B), as a ribbon representation (α-helices shown as coils and β-strands shown as arrows) with a semi-transparent rendering of the molecular surface superimposed. The α-chain is colored blue-to-red from N- to C-terminus and the light chain (β_2m) is colored dark gray. Domains are labeled. The peptide is shown in a CPK representation, with alternating residues colored purple or orange. The view in (B) looks down onto the top of the protein, the view seen from the perspective of an incoming TCR.

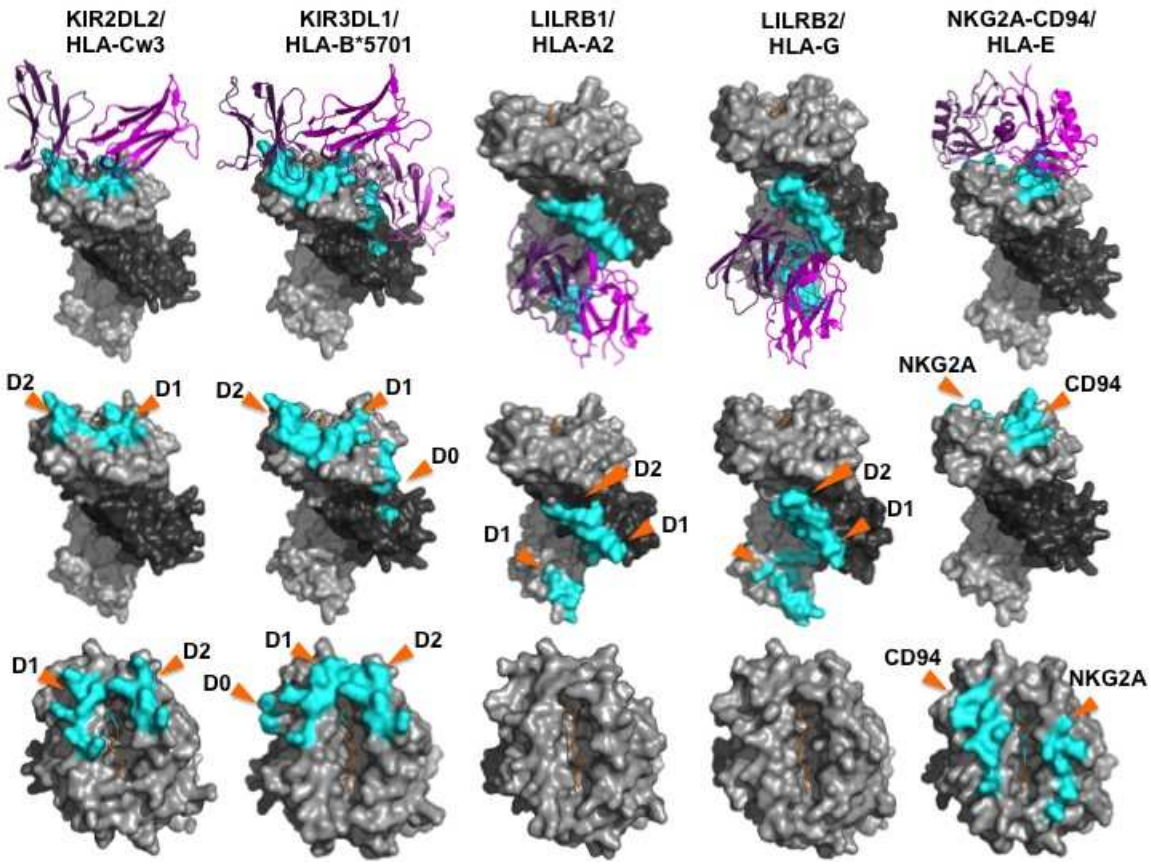

Figure 2: Interaction footprints for representative NK receptors on MHC class I proteins. Molecular surface representations of the class I ligand protein are shown, colored light gray for the heavy chain (α-chain) and dark gray for the light chain (β_2m); the bound antigenic peptide is shown as an orange squiggle. In the top row, the NK receptor is shown in ribbon representation, colored purple, with α-helices indicated as coils and β-strands indicated as arrows. The middle and bottom rows show the class I protein without the receptor, in two different orthogonal views. The interaction footprint of the receptor on the class I ligand is colored aqua, with the receptor domains contributing the contacts labeled in the middle and bottom rows. Contacts to the peptide are also colored aqua.

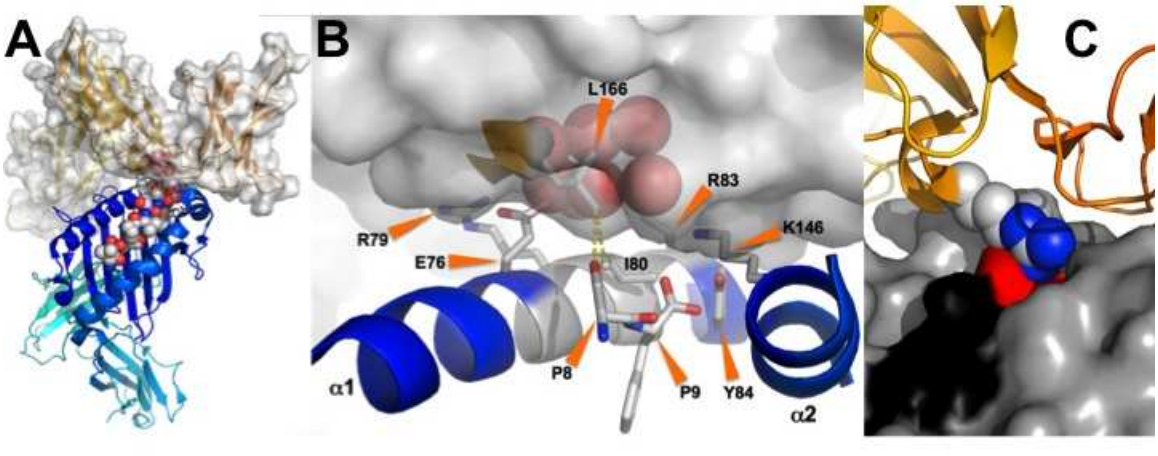

Figure 3: The KIR3DL1/HLA-B*5701 interface. (A) An overview of the complex is shown, with both class I ligand and KIR receptor shown in ribbon representations; a semi-transparent rendering of the molecular surface of KIR3DL1 is superimposed. The peptide is shown in CPK representation, colored by atom type (carbon: gray; oxygen: red; nitrogen: blue). (B) A zoomed-in view of the interface, centered on I80 in HLA-B. Only segments of the protein backbone ribbons and the side-chains of key residues (shown in licorice-stick representation, colored by atom-type as in (A) and labeled) are shown for clarity. The segment of the HLA-B heavy chain corresponding to the Bw4 epitope is colored gray. Direct contacts between Val166 in KIR3DL1 and Ile80 in HLA-B and P8 in the antigenic peptide are indicated with yellow dotted lines. The cluster of ordered water molecules in a pocket between the D1/HLA interface is shown as red spheres. (C) A model of the effect of substituting Val166 with arginine, the key 3DL1/3DS1 sequence difference. The side-chain of Arg166, placed in the least disallowed rotamer, is shown in CPK representation, colored by atom-type as in (A). The molecular surface of HLA-B*5701 is colored light gray for the platform domain, black for the peptide and red for P8 in the peptide. The steric clash is readily apparent, shown by the penetration of the Arg166 side-chain through the molecular surface.

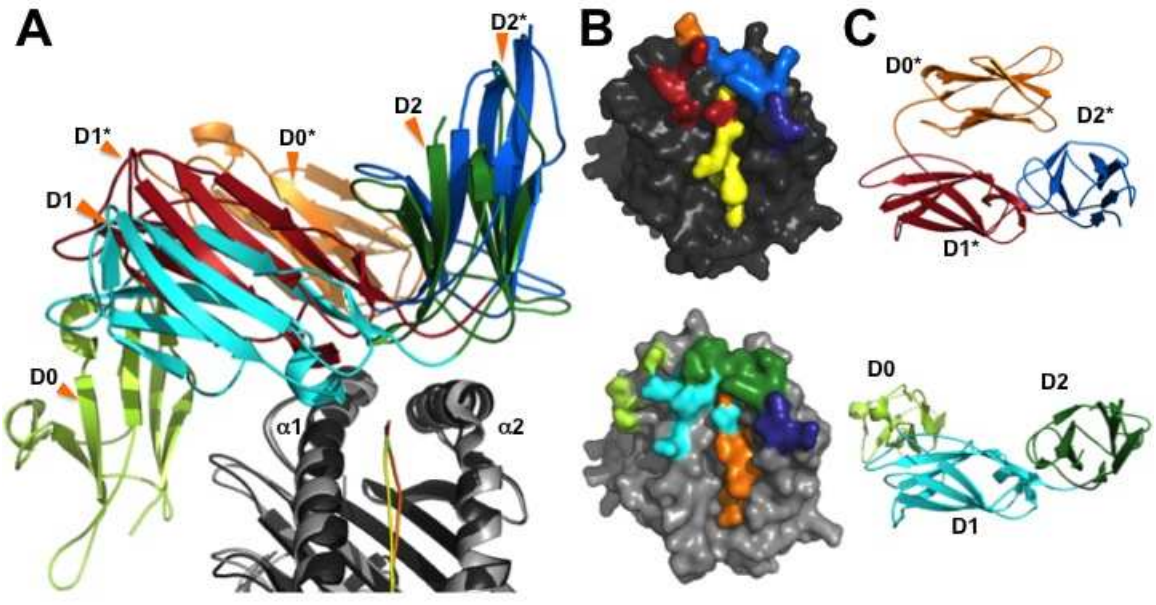

Figure 4: Comparison of the KIR3DL1*001/HLA-B*5701 complex structure with the KIR3DL1*015/HLA-A24 model. (A) A superposition of KIR3DL1*001/HLA-B*5701 and KIR3DL1*015/HLA-A24 is shown as a ribbon representation. KIR3DL1*001 is colored by domain (D0 in light green, D1 in aqua, and D2 in dark green); HLA-B*5701 is colored light grey with bound peptide in orange; KIR3DL1*015 is colored by domain and labels are qualified by asterisks (D0* orange, D1* red, D2* blue); HLA-A24 is colored dark grey with bound peptide in yellow. (B) The footprint of KIR3DL1*001 is shown on the molecular surface of HLA-B*5701 (top figure) and the footprint of KIR3DL1*015 is shown on the molecular surface of HLA-A24 (bottom figure) as viewed looking down on the peptide binding platform. Individual domain contacts are colored as in (A) with the exception of the D1-D2 loop of KIR which is colored dark blue. (C) Ribbon representations of KIR3DL1*001 and KIR3DL*015 are shown as viewed from above (rotated 90° from the position shown in (A)). Indiv idual domain contacts are colored as in (A).

References

1. Biron, CA, Byron, KS, Sullivan, JL. Severe herpesvirus infections in an adolescent without natural killer cells. N. Engl. J. Med. 1989; **320**: 1731–1735.

2. Bancroft, GJ. The role of natural killer cells in innate resistance to infection. Current Opinion in Immunology 1993; **5**: 503–510.

3. Biron, CA, Nguyen, KB, Pien, GC, Cousens, LP, Salazar-Mather, TP. Natural killer cells in antiviral defense: function and regulation by innate cytokines. Annu. Rev. Immunol. 1999; **17**: 189–220.

4. Cooper, MA, Fehniger, TA, Caligiuri, MA. The biology of human natural killer-cell subsets. Trends in Immunology 2001; **22**: 633–640.

5. Poli, A, Michel, T, Thérésine, M, Andrès, E, Hentges, F, Zimmer, J. CD56 brightnatural killer (NK) cells: an important NK cell subset. Immunology 2009; **126**: 458–465.

6. Bryceson, YT, Ljunggren, HG, Long, EO. Minimal requirement for induction of natural cytotoxicity and intersection of activation signals by inhibitory receptors. Blood 2009; **114**: 2657–2666.

7. Horng, T, Bezbradica, JS, Medzhitov, R. NKG2D signaling is coupled to the interleukin 15 receptor signaling pathway. Nat Immunol 2007; **8**: 1345–1352.

8. Storkus, WJ, Alexander, J, Payne, JA, Dawson, JR, Cresswell, P. Reversal of natural killing susceptibility in target cells expressing transfected class I HLA genes. Proceedings of the National Academy of Sciences of the United States of America 1989; **86**: 2361–2364.

9. Kärre, K, Ljunggren, HG, Piontek, G, Kiessling, R. Selective rejection of H-2-deficient

lymphoma variants suggests alternative immune defence strategy. Nature 1986; **319**: 675–678.

10. Shimizu, Y, DeMars, R. Demonstration by class I gene transfer that reduced susceptibility of human cells to natural killer cell-mediated lysis is inversely correlated with HLA class I antigen expression. Eur. J. Immunol. 1989; **19**: 447–451.

11. Ljunggren, HGH, Kärre, KK. In search of the “missing self”: MHC molecules and NK cell recognition. Immunol. Today 1990; **11**: 237–244.

12. Diefenbach, A, Raulet, DH. Strategies for target cell recognition by natural killer cells. Immunological Reviews 2001; **181**: 170–184.

13. Arase, H. Direct Recognition of Cytomegalovirus by Activating and Inhibitory NK Cell Receptors. Science 2002; **296**: 1323–1326.

14. Mandelboim, O, Lieberman, N, Lev, M, Paul, L, Arnon, TI, Bushkin, Y, Davis, DM, et al. Recognition of haemagglutinins on virus-infected cells by NKp46 activates lysis by human NK cells. Nature 2001; **409**: 1055–1060.

15. Stewart, CA, Laugier-Anfossi, F, Vély, F, Saulquin, X, Riedmuller, J, Tisserant, A, Gauthier, L, et al. Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-like receptors. Proceedings of the National Academy of Sciences of the United States of America 2005; **102**: 13224–13229.

16. Foley, B, De Santis, D, Lathbury, L, Christiansen, F, Witt, C. KIR2DS1-mediated activation overrides NKG2A-mediated inhibition in HLA-C C2-negative individuals. International Immunology 2008; **20**: 555–563.

17. Bjorkman, PJ, Saper, MA, Samraoui, B, Bennett, WS, Strominger, JL, Wiley, DC. Structure of the human class I histocompatibility antigen, HLA-A2. Nature 1987; **329**: 506–512.

18. Bjorkman, PJ. Finding the groove. Nat Immunol 2006; **7**: 787–789. Kimball, JW. Introduction to Immunology. Macmillan Library Reference, 1983.

20. Cohn, M. The T-cell receptor mediating restrictive recognition of antigen. Cell 1983; **33**: 657–669.

21. Shimonkevitz, R, Kappler, J, Marrack, P, Grey, H. Antigen recognition by H-2-restricted T cells. I. Cell-free antigen processing. J. Exp. Med. 1983; **158**: 303–316.

22. Bjorkman, PJ, Saper, MA, Samraoui, B, Bennett, WS, Strominger, JL, Wiley, DC. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature 1987; **329**: 512–518.

23. Wucherpfennig, KW, Allen, PM, Celada, F, Cohen, IR, De Boer, R, Garcia, KC, Goldstein, B, et al. Polyspecificity of T cell and B cell receptor recognition. Seminars in Immunology 2007; **19**: 216–224.

24. Marrack, P, Scott-Browne, JP, Dai, S, Gapin, L, Kappler, JW. Evolutionarily Conserved Amino Acids That Control TCR-MHC Interaction. Annu. Rev. Immunol. 2008; **26**: 171–203.

25. Garcia, KC, Adams, JJ, Feng, D, Ely, LK. The molecular basis of TCR germline bias for MHC is surprisingly simple. Nature Publishing Group 2009; **10**: 143–147.

26. Scott-Browne, JPJ, White, JJ, Kappler, JWJ, Gapin, LL, Marrack, PP. Germline-encoded amino acids in the alphabeta T-cell receptor control thymic selection. CORD Conference Proceedings 2009; **458**: 1043–1046.

27. Adams, JJ, Narayanan, S, Liu, B, Birnbaum, ME, Kruse, AC, Bowerman, NA, Chen, W, et al. T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. Immunity 2011; **35**: 681–693.

28. Scott-Browne, JP, Crawford, F, Young, MH, Kappler, JW, Marrack, P, Gapin, L. Evolutionarily conserved features contribute to αβ T cell receptor specificity. Immunity 2011; **35**: 526–535.

29. Marsh, S.G.E., Parham, P., Dupont, B., Geraghty, D.E., Trowsdale, J., Middleton, D., Vilches, C., Carrington, M., Witt, C., Guethlein, L.A., Shilling, H., Garcia, C.A., Hsu, K.C., & Wain, H. (2003). Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. 55(4), 220–226.

30. Lanier, LL, Corliss, BC, Wu, J, Leong, C, Phillips, JH. Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. Nature 1998; **391**: 703–707.

31. Mandelboim, O, Reyburn, HT, Valés-Gómez, M, Pazmany, L, Colonna, M, Borsellino, G, Strominger, JL. Protection from lysis by natural killer cells of group 1 and 2 specificity is mediated by residue 80 in human histocompatibility leukocyte antigen C alleles and also occurs with empty major histocompatibility complex molecules. J. Exp. Med. 2010; **184**: 913–922.

32. Winter, CC, Long, EO. A single amino acid in the p58 killer cell inhibitory receptor controls the ability of natural killer cells to discriminate between the two groups of HLA-C allotypes. J. Immunol. 1997; **158**: 4026–4028.

33. Gumperz, JE, Litwin, V, Phillips, JH, Lanier, LL, Parham, P. The Bw4 public epitope of HLA-B molecules confers reactivity with natural killer cell clones that express NKB1, a putative HLA receptor. J. Exp. Med. 1995; **181**: 1133–1144.

34. Gumperz, JE, Barber, LD, Valiante, NM, Percival, L, Phillips, JH, Lanier, LL, Parham, P. Conserved and variable residues within the Bw4 motif of HLA-B make separable contributions to recognition by the NKB1 killer cell-inhibitory receptor. J. Immunol. 1997; **158**: 5237–5241.

35. Parham, P. MHC class I molecules and kirs in human history, health and survival. Nat Rev Immunol 2005; **5**: 201–214.

36. Romero, V, Azocar, J, Zúñiga, J, Clavijo, OP, Terreros, D, Gu, X, Husain, Z, et al. Interaction of NK inhibitory receptor genes with HLA-C and MHC class II alleles in Hepatitis C virus infection outcome. Molecular immunology 2008; **45**: 2429–2436.

37. Singh, R, Kaul, R, Kaul, A, Khan, K. A comparative review of HLA associations with hepatitis B and C viral infections across global populations. World J. Gastroenterol. 2007; **13**: 1770–1787.

38. López-Botet, M, Angulo, A, Guma, M. Natural killer cell receptors for major histocompatibility complex class I and related molecules in cytomegalovirus infection. Tissue Antigens 2004; **63**: 195–203.

39. Estefanía, E, Gómez-Lozano, N, Portero, F, de Pablo, R, Solís, R, Sepúlveda, S, Vaquero, M, et al. Influence of KIR gene diversity on the course of HSV-1 infection: resistance to the disease is associated with the absence of KIR2DL2 and KIR2DS2. Tissue Antigens 2007; **70**: 34–41.

40. Khakoo, SI, Geller, R, Shin, S, Jenkins, JA, Parham, P. The D0 Domain of KIR3D Acts as a Major Histocompatibility Complex Class I Binding Enhancer. Journal of Experimental Medicine 2002; **196**: 911–921.

41. Carr, WH, Pando, MJ, Parham, P. KIR3DL1 polymorphisms that affect NK cell inhibition by HLA-Bw4 ligand. J. Immunol. 2005; **175**: 5222–5229.

42. Martin, MP, Gao, X, Lee, J-H, Nelson, GW, Detels, R, Goedert, JJ, Buchbinder, S, et al. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. Nat Genet

2002; **31**: 429–434.

43. Qi, Y, Martin, MP, Gao, X, Jacobson, L, Goedert, JJ, Buchbinder, S, Kirk, GD, et al. KIR/HLA Pleiotropism: Protection against Both HIV and Opportunistic Infections. PLoS Pathog 2006; **2**: e79.

44. Alter, G, Martin, MP, Teigen, N, Carr, WH, Suscovich, TJ, Schneidewind, A, Streeck, H, et al. Differential natural killer cell mediated inhibition of HIV-1 replication based on distinct KIR/HLA subtypes. Journal of Experimental Medicine 2007; **204**: 3027–3036.

45. Gaudieri, S, DeSantis, D, McKinnon, E, Moore, C, Nolan, D, Witt, CS, Mallal, SA, et al. Killer immunoglobulin-like receptors and HLA act both independently and synergistically to modify HIV disease progression. Genes Immun 2005;

46. Barbour, JD, Sriram, U, Caillier, SJ, Levy, JA, Hecht, FM, Oksenberg, JR. Synergy or Independence? Deciphering the Interaction of HLA Class I and NK Cell KIR Alleles in Early HIV-1 Disease Progression. PLoS Pathog 2007; **3**: e43.

47. Long, BR, Ndhlovu, LC, Oksenberg, JR, Lanier, LL, Hecht, FM, Nixon, DF, Barbour, JD. Conferral of Enhanced Natural Killer Cell Function by KIR3DS1 in Early Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 2008; **82**: 4785–4792.

48. Bouleta, S, Sharafia, S, Simica, N, Bruneaub, J, Routya, J-P, Tsoukasa, CM, Bernarda, NF. Increased proportion of KIR3DS1 homozygotes in HIV-exposed uninfected individuals. 2008; 1–5.

49. Alter, G, Rihn, S, Walter, K, Nolting, A, Martin, M, Rosenberg, ES, Miller, JS, et al. HLA Class I Subtype-Dependent Expansion of KIR3DS1+ and KIR3DL1+ NK Cells during Acute Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 2009; **83**: 6798–6805.

50. Fadda, L, O'Connor, GM, Kumar, S, Piechocka-Trocha, A, Gardiner, CM, Carrington, M, McVicar, DW, et al. Common HIV-1 Peptide Variants Mediate Differential Binding of KIR3DL1 to HLA-Bw4 Molecules. Journal of Virology 2011; **85**: 5970–5974.

51. Brackenridge, S, Evans, EJ, Toebes, M, Goonetilleke, N, Liu, MKP, di Gleria, K, Schumacher, TN, et al. An Early HIV Mutation within an HLA-B*57-Restricted T Cell Epitope Abrogates Binding to the Killer Inhibitory Receptor 3DL1. Journal of Virology 2011; **85**: 5415–5422.

52. Alter, G, Heckerman, D, Schneidewind, A, Fadda, L, Kadie, CM, Carlson, JM, Oniangue-Ndza, C, et al. HIV-1 adaptation to NK-cell-mediated immune pressure. Nature 2011; **476**: 96–100.

53. Faddaa, L, Borhisa, G, Ahmeda, P, Cheenta, K, Pageonb, SV, Cazalyc, A, Stathopoulosa, S, et al. Peptide antagonism as a mechanism for NK cell activation. 2010; 1–6.

54. Boyington, JC, Sun, PD. A structural perspective on MHC class I recognition by killer cell immunoglobulin-like receptors. Molecular immunology 2002; **38**: 1007–1021.

55. Natarajan, K, Dimasi, N, Wang, J, Mariuzza, RA, Margulies, DH. Structure and function of natural killer cell cell receptors: multiple molecular solutions to self, nonself discrimination. Annu. Rev. Immunol. 2002; **20**: 853–885.

56. Radaev, SS, Sun, PDP. Structure and function of natural killer cell surface receptors. Annu Rev Biophys Biomol Struct 2003; **32**: 93–114.

57. Sun, PDP. Structure and function of natural-killer-cell receptors. Immunol. Res. 2003; **27**: 539–548.

58. Joyce, MG, Sun, PD. The Structural Basis of Ligand Recognition by Natural Killer Cell

Receptors. Journal of Biomedicine and Biotechnology 2011; **2011**: 1–15.

59. Lanier, LL. NK cell recognition. Annu. Rev. Immunol. 2005; **23**: 225–274.

60. Fan, QR, Long, EO, Wiley, DC. Crystal structure of the human natural killer cell inhibitory receptor KIR2DL1-HLA-Cw4 complex. Nat Immunol 2001; **2**: 452–460.

61. Boyington, JC, Motyka, SA, Schuck, P, Brooks, AG, Sun, PD. Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. Nature 2000; **405**: 537–543.

62. Boyington, JC, Brooks, AG, Sun, PD. Structure of killer cell immunoglobulin-like receptors and their recognition of the class I MHC molecules. Immunological Reviews 2001; **181**: 66–78.

63. Vivian, JP, Duncan, RC, Berry, R, O'Connor, GM, Reid, HH, Beddoe, T, Gras, S, et al. Killer cell immunoglobulin-like receptor 3DL1-mediated recognition of human leukocyte antigen B. Nature 2011; **479**: 401–405.

64. Sharma, D, Bastard, K, Guethlein, LA, Norman, PJ, Yawata, N, Yawata, M, Pando, M, et al. Dimorphic Motifs in D0 and D1+D2 Domains of Killer Cell Ig-Like Receptor 3DL1 Combine to Form Receptors with High, Moderate, and No Avidity for the Complex of a Peptide Derived from HIV and HLA-A*2402. The Journal of Immunology 2009; **183**: 4569–4582.

65. Rojo, SS, Wagtmann, NN, Long, EOE. Binding of a soluble p70 killer cell inhibitory receptor to HLA-B*5101: requirement for all three p70 immunoglobulin domains. Eur. J. Immunol. 1997; **27**: 568–571.

66. Gillespie, GMA, Bashirova, A, Dong, T, McVicar, DW, Rowland-Jones, SL, Carrington, M. Lack of KIR3DS1 Binding to MHC Class I Bw4 Tetramers in Complex with CD8 +T Cell Epitopes. AIDS Research and Human Retroviruses 2007; **23**: 451–455.

67. Carr, WH, Rosen, DB, Arase, H, Nixon, DF, Michaelsson, J, Lanier, LL. Cutting Edge: *KIR3DS1,* a Gene Implicated in Resistance to Progression to AIDS, Encodes a DAP12-Associated Receptor Expressed on NK Cells That Triggers NK Cell Activation. 2006; 1–5.

68. Stewart-Jones, GBE, Gillespie, G, Overton, IM, Kaul, R, Roche, P, McMichael, AJ, Rowland-Jones, S, et al. Structures of three HIV-1 HLA-B*5703-peptide complexes and identification of related HLAs potentially associated with long-term nonprogression. J. Immunol. 2005; **175**: 2459–2468.

69. Cosman, D, Fanger, N, Borges, L, Kubin, M, Chin, W, Peterson, L, Hsu, ML. A novel immunoglobulin superfamily receptor for cellular and viral MHC class I molecules. Immunity 1997; **7**: 273–282.

70. Colonna, M, Navarro, F, Bellón, T, Llano, M, García, P, Samaridis, J, Angman, L, et al. A common inhibitory receptor for major histocompatibility complex class I molecules on human lymphoid and myelomonocytic cells. J. Exp. Med. 1997; **186**: 1809–1818.

71. Volz, A, Wende, H, Laun, K, Ziegler, A. Genesis of the ILT/LIR/MIR clusters within the human leukocyte receptor complex. Immunological Reviews 2001; **181**: 39–51.

72. Barrow, AD, Trowsdale, J. The extended human leukocyte receptor complex: diverse ways of modulating immune responses. Immunological Reviews 2008; **224**: 98–123.

73. Nakajima, H, Samaridis, J, Angman, L, Colonna, M. Human myeloid cells express an activating ILT receptor (ILT1) that associates with Fc receptor gamma-chain. J. Immunol. 1999; **162**: 5–8.

74. Brown, D, Trowsdale, J, Allen, R. The LILR family: modulators of innate and adaptive immune pathways in health and disease. Tissue Antigens 2004; **64**: 215–225.

75. Lepin, EJ, Bastin, JM, Allan, DS, Roncador, G, Braud, VM, Mason, DY, van der Merwe, PA, et al. Functional characterization of HLA-F and binding of HLA-F tetramers to ILT2 and ILT4 receptors. Eur. J. Immunol. 2000; **30**: 3552–3561.

76. Navarro, F, Llano, M, Bellón, T, Colonna, M, Geraghty, DE, López-Botet, M. The ILT2(LIR1) and CD94/NKG2A NK cell receptors respectively recognize HLA-G1 and HLA-E molecules co-expressed on target cells. Eur. J. Immunol. 1999; **29**: 277–283.

77. Gonen-Gross, T, Achdout, H, Gazit, R, Hanna, J, Mizrahi, S, Markel, G, Goldman-Wohl, D, et al. Complexes of HLA-G protein on the cell surface are important for leukocyte Ig-like receptor-1 function. J. Immunol. 2003; **171**: 1343–1351.

78. Lee, N, Goodlett, DR, Ishitani, A, Marquardt, H, Geraghty, DE. HLA-E surface expression depends on binding of TAP-dependent peptides derived from certain HLA class I signal sequences. J. Immunol. 1998; **160**: 4951–4960.

79. van Hall, T, Oliveira, CC, Joosten, SA, Ottenhoff, TH. The other Janus face of Qa-1 and HLA-E: diverse peptide repertoires in times of stress. Microbes Infect. 2010; 12;910-18 .

80. Lee, N, Ishitani, A, Geraghty, DE. HLA-F is a surface marker on activated lymphocytes. Eur. J. Immunol. 2010; **40**: 2308–2318.

81. Goodridge, JP, Burian, A, Lee, N, Geraghty, DE. HLA-F complex without peptide binds to MHC class I protein in the open conformer form. The Journal of Immunology 2010; **184**: 6199–6208.

82. Rouas-Freiss, NN, Gonçalves, RMR, Menier, CC, Dausset, JJ, Carosella, EDE. Direct evidence to support the role of HLA-G in protecting the fetus from maternal uterine natural killer cytolysis. Proceedings of the National Academy of Sciences of the United States of America

1997; **94**: 11520–11525.

83. Carosella, EDE, Gregori, SS, Rouas-Freiss, NN, LeMaoult, JJ, Menier, CC, Favier, BB. The role of HLA-G in immunity and hematopoiesis. CORD Conference Proceedings 2011; **68**: 353–368.

84. Yang, Z, Bjorkman, PJ. Structure of UL18, a peptide-binding viral MHC mimic, bound to a host inhibitory receptor. Proceedings of the National Academy of Sciences 2008; **105**: 10095–10100.

85. Chapman, TL, Heikema, AP, West, AP, Bjorkman, PJ. Crystal structure and ligand binding properties of the D1D2 region of the inhibitory receptor LIR-1 (ILT2). Immunity 2000; **13**: 727–736.

86. Chapman, TL, Heikeman, AP, Bjorkman, PJ. The inhibitory receptor LIR-1 uses a common binding interaction to recognize class I MHC molecules and the viral homolog UL18. Immunity 1999; **11**: 603–613.

87. Wagner, CS, Ljunggren, H-G, Achour, A. Immune modulation by the human cytomegalovirus-encoded molecule UL18, a mystery yet to be solved. J. Immunol. 2008; **180**: 19–24.

88. Kaiser, BK, Pizarro, JC, Kerns, J, Strong, RK. Structural basis for NKG2A/CD94 recognition of HLA-E. Proceedings of the National Academy of Sciences 2008; **105**: 6696–6701.

89. Gao, X, Nelson, GW, Karacki, P, Martin, MP, Phair, J, Kaslow, R, Goedert, JJ, et al. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. N. Engl. J. Med. 2001; **344**: 1668–1675.

90. Lichterfeld, M, Yu, XG. The emerging role of leukocyte immunoglobulin-like receptors

(LILRs) in HIV-1 infection. Journal of Leukocyte Biology 2012; **91**: 27–33.

91. Huang, J, Goedert, JJ, Sundberg, EJ, Cung, TDH, Burke, PS, Martin, MP, Preiss, L, et al. HLA-B*35-Px-mediated acceleration of HIV-1 infection by increased inhibitory immunoregulatory impulses. Journal of Experimental Medicine 2009; **206**: 2959–2966.

92. McMichael, AJ, Yvonne Jones, E. First-Class Control of HIV-1. Science 2010; **330**: 1488–1490.

93. Jones, DC, Kosmoliaptsis, V, Apps, R, Lapaque, N, Smith, I, Kono, A, Chang, C, et al. HLA Class I Allelic Sequence and Conformation Regulate Leukocyte Ig-Like Receptor Binding. The Journal of Immunology 2011; **186**: 2990–2997.

94. Lichterfeld, M, Kavanagh, DG, Williams, KL, Moza, B, Mui, SK, Miura, T, Sivamurthy, R, et al. A viral CTL escape mutation leading to immunoglobulin-like transcript 4 mediated functional inhibition of myelomonocytic cells. Journal of Experimental Medicine 2007; **204**: 2813–2824.

95. Willcox, BE, Thomas, LM, Bjorkman, PJ. Crystal structure of HLA-A2 bound to LIR-1, a host and viral major histocompatibility complex receptor. Nat Immunol 2003; **4**: 913–919.

96. Shiroishi, M, Kuroki, K, Rasubala, L, Tsumoto, K, Kumagai, I, Kurimoto, E, Kato, K, et al. Structural basis for recognition of the nonclassical MHC molecule HLA-G by the leukocyte Ig-like receptor B2 (LILRB2/LIR2/ILT4/CD85d). Proceedings of the National Academy of Sciences of the United States of America 2006; **103**: 16412–16417.

97. Krissinel, E, Henrick, K. Inference of macromolecular assemblies from crystalline state. Journal of Molecular Biology 2007; **372**: 774–797.

98. Kuroki, K, Tsuchiya, N, Shiroishi, M, Rasubala, L, Yamashita, Y, Matsuta, K, Fukazawa, T, et al. Extensive polymorphisms of LILRB1 (ILT2, LIR1) and their association with HLA-DRB1

shared epitope negative rheumatoid arthritis. Human Molecular Genetics 2005; **14**: 2469–2480.

99. Willcox, BE, Thomas, LM, Chapman, TL, Heikema, AP, West, AP, Bjorkman, PJ. Crystal structure of LIR-2 (ILT4) at 1.8 A: differences from LIR-1 (ILT2) in regions implicated in the binding of the Human Cytomegalovirus class I MHC homolog UL18. BMC Struct. Biol. 2002; **2**: 6.

100. Gruda, R, Achdout, H, Stern-Ginossar, N, Gazit, R, Betser-Cohen, G, Manaster, I, Katz, G, et al. Intracellular cysteine residues in the tail of MHC class I proteins are crucial for extracellular recognition by leukocyte Ig-like receptor 1. J. Immunol. 2007; **179**: 3655–3661.

101. Park, B, Kim, Y, Shin, J, Lee, S, Cho, K, Früh, K, Lee, S, et al. Human cytomegalovirus inhibits tapasin-dependent peptide loading and optimization of the MHC class I peptide cargo for immune evasion. Immunity 2004; **20**: 71–85.

102. Browne, H, Smith, G, Beck, S, Minson, T. A complex between the MHC class I homologue encoded by human cytomegalovirus and beta 2 microglobulin. Nature 1990; **347**: 770–772.

103. Fahnestock, ML, Johnson, JL, Feldman, RM, Neveu, JM, Lane, WS, Bjorkman, PJ. The MHC class I homolog encoded by human cytomegalovirus binds endogenous peptides. Immunity 1995; **3**: 583–590.

104. Wagner, CS, Rölle, A, Cosman, D, Ljunggren, H-G, Berndt, KD, Achour, A. Structural Elements Underlying the High Binding Affinity of Human Cytomegalovirus UL18 to Leukocyte Immunoglobulin-like Receptor-1. Journal of Molecular Biology 2007; **373**: 695–705.

105. Cerboni, C, Achour, A, Wärnmark, A, Mousavi-Jazi, M, Sandalova, T, Hsu, M-L, Cosman, D, et al. Spontaneous mutations in the human CMV HLA class I homologue UL18 affect its binding to the inhibitory receptor LIR-1/ILT2/CD85j. Eur. J. Immunol. 2006; **36**: 732–741.

106. Beck, S, Barrell, BG. Human cytomegalovirus encodes a glycoprotein homologous to MHC class-I antigens. Nature 1988; **331**: 269–272.

107. Sullivan, LC, Clements, CS, Beddoe, T, Johnson, D, Hoare, HL, Lin, J, Huyton, T, et al. The Heterodimeric Assembly of the CD94-NKG2 Receptor Family and Implications for Human Leukocyte Antigen-E Recognition. Immunity 2007; **27**: 900–911.

108. Brostjan, C, Bellón, T, Sobanov, Y, López-Botet, M, Hofer, E. Differential expression of inhibitory and activating CD94/NKG2 receptors on NK cell clones. J. Immunol. Methods 2002; **264**: 109–119.

109. Kaiser, BK, Barahmand-Pour, F, Paulsene, W, Medley, S, Geraghty, DE, Strong, RK. Interactions between NKG2x immunoreceptors and HLA-E ligands display overlapping affinities and thermodynamics. J. Immunol. 2005; **174**: 2878–2884.

110. Lazetic, S, Chang, C, Houchins, JP, Lanier, LL, Phillips, JH. Human natural killer cell receptors involved in MHC class I recognition are disulfide-linked heterodimers of CD94 and NKG2 subunits. J. Immunol. 1996; **157**: 4741–4745.

111. Borrego, F, Ulbrecht, M, Weiss, EH, Coligan, JE, Brooks, AG. Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. J. Exp. Med. 1998; **187**: 813–818.

112. Braud, VM, Allan, DS, O'Callaghan, CA, Söderström, K, D'Andrea, A, Ogg, GS, Lazetic, S, et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. Nature 1998; **391**: 795–799.

113. Brooks, AG, Borrego, F, Posch, PE, Patamawenu, A, Scorzelli, CJ, Ulbrecht, M, Weiss,

EH, et al. Specific recognition of HLA-E, but not classical, HLA class I molecules by soluble CD94/NKG2A and NK cells. J. Immunol. 1999; **162**: 305–313.

114. Lee, N, Llano, M, Carretero, M, Ishitani, A, Navarro, F, López-Botet, M, Geraghty, DE. HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. Proceedings of the National Academy of Sciences of the United States of America 1998; **95**: 5199–5204.

115. Suto, Y, Yabe, T, Maenaka, K, Tokunaga, K, Tadokoro, K, Juji, T. The human natural killer gene complex (NKC) is located on chromosome 12p13.1-p13.2. Immunogenetics 1997; **46**: 159–162.

116. Houchins, JP, Yabe, T, McSherry, C, Bach, FH. DNA sequence analysis of NKG2, a family of related cDNA clones encoding type II integral membrane proteins on human natural killer cells. J. Exp. Med. 1991; **173**: 1017–1020.

117. Lanier, LL, Corliss, B, Wu, J, Phillips, JH. Association of DAP12 with activating CD94/NKG2C NK cell receptors. Immunity 1998; **8**: 693–701.

118. Braud, VM, Allan, DS, Wilson, D, McMichael, AJ. TAP- and tapasin-dependent HLA-E surface expression correlates with the binding of an MHC class I leader peptide. Curr. Biol. 1998; **8**: 1–10.

119. Brooks, CR, Elliott, T, Parham, P, Khakoo, SI. The inhibitory receptor NKG2A determines lysis of vaccinia virus-infected autologous targets by NK cells. J. Immunol. 2006; **176**: 1141–1147.

120. Bjorkstrom, NK, Lindgren, T, Stoltz, M, Fauriat, C, Braun, M, Evander, M, Michaelsson, J, et al. Rapid expansion and long-term persistence of elevated NK cell numbers in humans infected with hantavirus. Journal of Experimental Medicine 2011; **208**: 13–21.

121. Tomasec, P, Braud, VM, Rickards, C, Powell, MB, McSharry, BP, Gadola, S, Cerundolo, V, et al. Surface expression of HLA-E, an inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40. Science 2000; **287**: 1031.

122. Guma, M. Imprint of human cytomegalovirus infection on the NK cell receptor repertoire. Blood 2004; **104**: 3664–3671.

123. Petrie, EJ, Clements, CS, Lin, J, Sullivan, LC, Johnson, D, Huyton, T, Heroux, A, et al. CD94-NKG2A recognition of human leukocyte antigen (HLA)-E bound to an HLA class I leader sequence. Journal of Experimental Medicine 2008; **205**: 725–735.

124. Plougastel, B, Trowsdale, J. Cloning of NKG2-F, a new member of the NKG2 family of human natural killer cell receptor genes. Eur. J. Immunol. 1997; **27**: 2835–2839.

125. Bauer, S, Groh, V, Wu, J, Steinle, A, Phillips, JH, Lanier, LL, Spies, T. Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. Science 1999; **285**: 727–729.

126. Champsaur, M, Lanier, LL. Effect of NKG2D ligand expression on host immune responses. Immunological Reviews 2010; **235**: 267–285.

127. Li, PP, Willie, STS, Bauer, SS, Morris, DLD, Spies, TT, Strong, RKR. Crystal structure of the MHC class I homolog MIC-A, a gammadelta T cell ligand. Immunity 1999; **10**: 577–584.

128. Li, P, Morris, DL, Willcox, BE, Steinle, A, Spies, T, Strong, RK. Complex structure of the activating immunoreceptor NKG2D and its MHC class I-like ligand MICA. Nat Immunol 2001; **2**: 443–451.

129. Radaev, S, Rostro, B, Brooks, AG, Colonna, M, Sun, PD. Conformational plasticity revealed by the cocrystal structure of NKG2D and its class I MHC-like ligand ULBP3. Immunity 2001; **15**: 1039–1049.

130. Holmes, MA, Li, P, Petersdorf, EW, Strong, RK. Structural studies of allelic diversity of the MHC class I homolog MIC-B, a stress-inducible ligand for the activating immunoreceptor NKG2D. J. Immunol. 2002; **169**: 1395–1400.

131. McFarland, BJ, Kortemme, T, Yu, SF, Baker, D, Strong, RK. Symmetry recognizing asymmetry: analysis of the interactions between the C-type lectin-like immunoreceptor NKG2D and MHC class I-like ligands. Structure 2003; **11**: 411–422.

132. McFarland, B, Strong, S. Thermodynamic analysis of degenerate recognition by the NKG2D immunoreceptor: not induced fit but rigid adaptation. Immunity 2003; 19;772-4.

133. Moretta, A, Biassoni, R, Bottino, C, Mingari, MC, Moretta, L. Natural cytotoxicity receptors that trigger human NK-cell-mediated cytolysis. Immunol. Today 2000; **21**: 228–234.

134. Cantoni, C, Bottino, C, Vitale, M, Pessino, A, Augugliaro, R, Malaspina, A, Parolini, S, et al. NKp44, a triggering receptor involved in tumor cell lysis by activated human natural killer cells, is a novel member of the immunoglobulin superfamily. J. Exp. Med. 1999; **189**: 787–796.

135. Pessino, A, Sivori, S, Bottino, C, Malaspina, A, Morelli, L, Moretta, L, Biassoni, R, et al. Molecular cloning of NKp46: a novel member of the immunoglobulin superfamily involved in triggering of natural cytotoxicity. J. Exp. Med. 1998; **188**: 953–960.

136. Pende, D, Parolini, S, Pessino, A, Sivori, S, Augugliaro, R, Morelli, L, Marcenaro, E, et al. Identification and molecular characterization of NKp30, a novel triggering receptor involved in natural cytotoxicity mediated by human natural killer cells. J. Exp. Med. 1999; **190**: 1505–1516.

137. Vitale, M, Bottino, C, Sivori, S, Sanseverino, L, Castriconi, R, Marcenaro, E, Augugliaro, R, et al. NKp44, a novel triggering surface molecule specifically expressed by activated natural killer cells, is involved in non-major histocompatibility complex-restricted tumor cell lysis. J. Exp.

Med. 1998; **187**: 2065–2072.

138. Arnon, TI, Lev, M, Katz, G, Chernobrov, Y, Porgador, A, Mandelboim, O. Recognition of viral hemagglutinins by NKp44 but not by NKp30. Eur. J. Immunol. 2001; **31**: 2680–2689.

139. Jarahian, M, Watzl, C, Fournier, P, Arnold, A, Djandji, D, Zahedi, S, Cerwenka, A, et al. Activation of Natural Killer Cells by Newcastle Disease Virus Hemagglutinin-Neuraminidase. Journal of Virology 2009; **83**: 8108–8121.

140. Arnon, TI, Achdout, H, Levi, O, Markel, G, Saleh, N, Katz, G, Gazit, R, et al. Inhibition of the NKp30 activating receptor by pp65 of human cytomegalovirus. Nat Immunol 2005; **6**: 515–523.

141. De Maria, A, Fogli, M, Mazza, S, Basso, M, Picciotto, A, Costa, P, Congia, S, et al. Increased natural cytotoxicity receptor expression and relevant IL-10 production in NK cells from chronically infected viremic HCV patients. Eur. J. Immunol. 2007; **37**: 445–455.

142. Ahlenstiel, G, Titerence, RH, Koh, C, Edlich, B, Feld, JJ, Rotman, Y, Ghany, MG, et al. Natural Killer Cells Are Polarized Toward Cytotoxicity in Chronic Hepatitis C in an Interferon-Alfa–Dependent Manner. YGAST 2010; **138**: 325–335.e2.

143. Chisholm, SE, Howard, K, Gómez, MV, Reyburn, HT. Expression of ICP0 Is Sufficient to Trigger Natural Killer Cell Recognition of Herpes Simplex Virus–Infected Cells by Natural Cytotoxicity Receptors. J INFECT DIS 2007; **195**: 1160–1168.

144. Fuller, CL, Ruthel, G, Warfield, KL, Swenson, DL, Bosio, CM, Aman, MJ, Bavari, S. NKp30-dependent cytolysis of filovirus-infected human dendritic cells. Cell Microbiol 2007; **9**: 962–976.

145. Hershkovitz, O, Rosental, B, Rosenberg, LA, Navarro-Sanchez, ME, Jivov, S, Zilka, A,

Gershoni-Yahalom, O, et al. NKp44 Receptor Mediates Interaction of the Envelope Glycoproteins from the West Nile and Dengue Viruses with NK Cells. The Journal of Immunology 2009; **183**: 2610–2621.

146. Chisholm, SE, Reyburn, HT. Recognition of vaccinia virus-infected cells by human natural killer cells depends on natural cytotoxicity receptors. Journal of Virology 2006; **80**: 2225–2233.

147. Foster, CE. Crystal Structure of the Human Natural Killer (NK) Cell Activating Receptor NKp46 Reveals Structural Relationship to Other Leukocyte Receptor Complex Immunoreceptors. Journal of Biological Chemistry 2003; **278**: 46081–46086.

148. Ponassi, M, Cantoni, C, Biassoni, R, Conte, R, Spallarossa, A, Pesce, A, Moretta, A, et al. Structure of the human NK cell triggering receptor NKp46 ectodomain. Biochemical and Biophysical Research Communications 2003; **309**: 317–323.

149. Cantoni, C, Ponassi, M, Biassoni, R, Conte, R, Spallarossa, A, Moretta, A, Moretta, L, et al. The Three-Dimensional Structure of the Human NK Cell Receptor NKp44, a Triggering Partner in Natural Cytotoxicity. Structure 2003; **11**: 725–734.

150. Joyce, MG, Tran, P, Zhuravleva, MA, Jaw, J, Colonna, M, Sun, PD. Crystal structure of human natural cytotoxicity receptor NKp30 and identification of its ligand binding site. Proceedings of the National Academy of Sciences 2011; **108**: 6223–6228.

151. Mavoungou, E, Held, J, Mewono, L, Kremsner, PG. A Duffy Binding–Like Domain Is Involved in the NKp30□Mediated Recognition of Plasmodiumfalciparum–Parasitized Erythrocytes by Natural Killer Cells. J INFECT DIS 2007; **195**: 1521–1531.

152. Pogge von Strandmann, E, Simhadri, VR, Tresckow, von, B, Sasse, S, Reiners, KS, Hansen, HP, Rothe, A, et al. Human Leukocyte Antigen-B-Associated Transcript 3 Is Released

from Tumor Cells and Engages the NKp30 Receptor on Natural Killer Cells. Immunity 2007; **27**: 965–974.

153. Hecht, M-L, Rosental, B, Horlacher, T, Hershkovitz, O, De Paz, JL, Noti, C, Schauer, S, et al. Natural Cytotoxicity Receptors NKp30, NKp44 and NKp46 Bind to Different Heparan Sulfate/Heparin Sequences. J. Proteome Res. 2009; **8**: 712–720.

154. Brandt, CS, Baratin, M, Yi, EC, Kennedy, J, Gao, Z, Fox, B, Haldeman, B, et al. The B7 family member B7-H6 is a tumor cell ligand for the activating natural killer cell receptor NKp30 in humans. Journal of Experimental Medicine 2009; **206**: 1495–1503.

155. Schwartz, JC, Zhang, X, Fedorov, AA, Nathenson, SG, Almo, SC. Structural basis for co-stimulation by the human CTLA-4/B7-2 complex. Nature 2001; **410**: 604–608.

156. Stamper, CC, Zhang, Y, Tobin, JF, Erbe, DV, Ikemizu, S, Davis, SJ, Stahl, ML, et al. Crystal structure of the B7-1/CTLA-4complex that inhibits humanimmune responses. 2001; 1–4.

157. Arnon, TI. The mechanisms controlling the recognition of tumor- and virus-infected cells by NKp46. Blood 2004; **103**: 664–672.

158. Skehel, JJ, Wiley, DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu. Rev. Biochem. 2000; **69**: 531–569.

159. Mao, H, Tu, W, Liu, Y, Qin, G, Zheng, J, Chan, PL, Lam, KT, et al. Inhibition of Human Natural Killer Cell Activity by Influenza Virions and Hemagglutinin. Journal of Virology 2010; **84**: 4148–4157.

160. Elliott, JM, Yokoyama, WM. Unifying concepts of MHC-dependent natural killer cell education. Trends in Immunology 2011; **32**: 364–372.

161. Orr, MT, Lanier, LL. Natural killer cell licensing during viral infection. Adv. Exp. Med. Biol. 2011; **780**: 37–44.

162. LisniÄ, VJ, Krmpotić, A, Jonjić, S. Modulation of natural killer cell activity by viruses. Current Opinion in Microbiology 2010; **13**: 530–539.

163. Thomas, M, Boname, JM, Field, S, Nejentsev, S, Salio, M, Cerundolo, V, Wills, M, et al. Down-regulation of NKG2D and NKp80 ligands by Kaposi's sarcoma-associated herpesvirus K5 protects against NK cell cytotoxicity. Proceedings of the National Academy of Sciences 2008; **105**: 1656–1661.

164. Wolan, DW, Teyton, L, Rudolph, MG, Villmow, B, Bauer, S, Busch, DH, Wilson, IA. Crystal structure of the murine NK cell-activating receptor NKG2D at 1.95 A. Nat Immunol 2001; **2**: 248–254.

165. Li, P, McDermott, G, Strong, RK. Crystal structures of RAE-1beta and its complex with the activating immunoreceptor NKG2D. Immunity 2002; **16**: 77–86.

166. Oliviero, B, Varchetta, S, Paudice, E, Michelone, G, Zaramella, M, Mavilio, D, De Filippi, F, et al. Natural Killer Cell Functional Dichotomy in Chronic Hepatitis B and Chronic Hepatitis C Virus Infections. YGAST 2009; **137**: 1151–1160.e7.

167. Call, ME, Wucherpfennig, KW, Chou, JJ. The structural basis for intramembrane assembly of an activating immunoreceptor complex. Nature Publishing Group 2010; **11**: 1023–1029.

168. Fausther-Bovendo, H, Sol-Foulon, N, Candotti, D, Agut, H, Schwartz, O, Debré, P, Vieillard, V. HIV escape from natural killer cytotoxicity: nef inhibits NKp44L expression on CD4+ T cells. AIDS 2009; **23**: 1077–1087.

169. Colonna, M, Borsellino, G, Falco, M, Ferrara, GB, Strominger, JL. HLA-C is the inhibitory

ligand that determines dominant resistance to lysis by NK1- and NK2-specific natural killer cells. Proceedings of the National Academy of Sciences of the United States of America 1993; **90**: 12000–12004.

170. Jarahian, M, Fiedler, M, Cohnen, A, Djandji, D, Hämmerling, GJ, Gati, C, Cerwenka, A, et al. Modulation of NKp30- and NKp46-Mediated Natural Killer Cell Responses by Poxviral Hemagglutinin. PLoS Pathog 2011; **7**: e1002195.

171. Welte, S, Kuttruff, S, Waldhauer, I, Steinle, A. Mutual activation of natural killer cells and monocytes mediated by NKp80-AICL interaction. Nat Immunol 2006; **7**: 1334–1342.

172. Askar, M, Avery, R, Corey, R, Lopez, R, Thomas, D, Pidwell, D, Eghtesad, B, et al. Lack of killer immunoglobulin-like receptor 2DS2 (KIR2DS2) and KIR2DL2 is associated with poor responses to therapy of recurrent hepatitis C virus in liver transplant recipients. Liver Transpl 2009; **15**: 1557–1563.

173. Saulquin, X, Gastinel, LN, Vivier, E. Crystal Structure of the Human Natural Killer Cell Activating Receptor KIR2DS2 (CD158j). Journal of Experimental Medicine 2003; **197**: 933–938.

174. Katz, G, Markel, G, Mizrahi, S, Arnon, TI, Mandelboim, O. Recognition of HLA-Cw4 but not HLA-Cw6 by the NK cell receptor killer cell Ig-like receptor two-domain short tail number 4. J. Immunol. 2001; **166**: 7260–7267.

175. Zúñiga, J, Romero, V, José Azocar, D, Terreros, Inés Vargas-Rojas, M, Torres-García, D, et al. Protective KIR-HLA interactions for HCV infection in intravenous drug users. Molecular immunology 2009; **46**: 2723.

176. Graef, T, Moesta, AK, Norman, PJ, Abi-Rached, L, Vago, L, Older Aguilar, AM, Gleimer, M, et al. KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for

HLA-A*11 while diminishing avidity for HLA-C. Journal of Experimental Medicine 2009; **206**: 2557–2572.

177. Rajagopalan, S, Long, EO. A human histocompatibility leukocyte antigen (HLA)-G-specific receptor expressed on all natural killer cells. J. Exp. Med. 1999; **189**: 1093–1100.

178. Boyington, JC, Riaz, AN, Patamawenu, A, Coligan, JE, Brooks, AG, Sun, PD. Structure of CD94 reveals a novel C-type lectin fold: implications for the NK cell-associated CD94/NKG2 receptors. Immunity 1999; **10**: 75–82.

179. Nattermann, JJ, Nischalke, HDH, Hofmeister, VV, Kupfer, BB, Ahlenstiel, GG, Feldmann, GG, Rockstroh, JJ, et al. HIV-1 infection leads to increased HLA-E expression resulting in impaired function of natural killer cells. Antivir Ther 2005; **10**: 95–107.

180. Jinushi, M, Takehara, T, Tatsumi, T, Kanto, T, Miyagi, T, Suzuki, T, Kanazawa, Y, et al. Negative regulation of NK cell activities by inhibitory receptor CD94/NKG2A leads to altered NK cell-induced modulation of dendritic cell functions in chronic hepatitis C virus infection. J. Immunol. 2004; **173**: 6072–6081.

181. Fan, QR, Mosyak, L, Winter, CC, Wagtmann, N, Long, EO, Wiley, DC. Structure of the inhibitory receptor for human natural killer cells resembles haematopoietic receptors (KIR2DL1). Nature 1997; **389**: 96–100.

182. Snyder, GA, Brooks, AG, Sun, PD. Crystal structure of the HLA-Cw3 allotype-specific killer cell inhibitory receptor KIR2DL2. Proceedings of the National Academy of Sciences of the United States of America 1999; **96**: 3864–3869.

183. Moraru, M, Cisneros, E, Gómez-Lozano, N, de Pablo, R, Portero, F, Canizares, M, Vaquero, M, et al. Host Genetic Factors in Susceptibility to Herpes Simplex Type 1 Virus

Infection: Contribution of Polymorphic Genes at the Interface of Innate and Adaptive Immunity. The Journal of Immunology 2012; **188**: 4412–4420.

184. Maenaka, K, Juji, T, Stuart, DI, Jones, EY. Crystal structure of the human p58 killer cell inhibitory receptor (KIR2DL3) specific for HLA-Cw3-related MHC class I. Structure 1999; **7**: 391–398.

185. Martin, MP, Qi, Y, Gao, X, Yamada, E, Martin, JN, Pereyra, F, Colombo, S, et al. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. Nat Genet 2007; **39**: 733–740.

Chapter 3: Autoreactivity and exceptional CDR plasticity (but not unusual polyspecificity) hinder elicitation of the anti-HIV antibody 4E10

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: ^{671}NWFDITNWLW^{680}; core epitope: NWF^{D}/NIT), immediately adjacent to the viral membrane in the gp41 membrane proximal external region (MPER) [4, 10]. bNAbs targeting the MPER, such 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: ^{662}ELDKWA^{667}) 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_HDJ_H 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 β2 glycoprotein I (β2GPI) 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 □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 $4E10H^{+/+}$ knock-in mice was evaluated by flow cytometry and compared with both wild-type B6 and control $KL25H^{+/+}$ knock-in mice (**Fig. 1A**). $KL25H^{+/+}$ 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 $B220^{+}$ cells from $4E10H^{+/+}$ knock-in mice revealed a dramatic reduction of both $B220^{+}IgM^{+}IgD^{-}$ pre-B cells and $B220^{+}IgM^{+}IgD^{+}$ 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 $B220^{+}$ 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$^{+/+}$, KL25H$^{+/+}$, 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')$_2$, or in culture media alone (no stimulus). The following day, B cell proliferation was assessed by flow cytometric measurement of dye dilution in B220^{+} cells (**Fig. 1B**). B220^{+} 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$^{+/+}$, KL25H$^{+/+}$, 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$^{+/+}$ and B6 B cells proliferated in response to IgM cross-linking, 4E10H$^{+/+}$ 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 the 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.

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 interactions. 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 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 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 Li_2SO_4 (d_{min} = 2.4Å), 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 scrunched 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 Fab complex structure (2FX7.pdb [9]), showed that the largest structural differences are also between HCDR1 and 3, with movements of ~12Å between the Cα positions of W100H in HCDR3 and ~6Å between the

Cα positions of F29H in HCDR1 (**Fig. 5A and B**; pair-wise superpositions resulted in a root mean square deviations (RMSD) on all common Cαs of 0.76Å for the heavy chain, with HCDR1 and 3 deleted, and 0.32Å 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 (αPDI) 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 αPDI 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 αPDI 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 $-\text{Log}_{10}$ *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^{D}/NIT) 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 $-\text{Log}_{10}$ *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 $-\text{Log}_{10}$ *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, 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_3R; NP_002214 [69, 70]) [71]. MEME motif analysis [72] revealed a highly conserved core motif (DEQGSK^{V}/I^{N}/SDFF) within these three peptides quite unrelated to the 4E10 MPER epitope (**Fig. 7B** inset). While no homologous crystal structure is available, Phyre^2 threading [73] predicts that this sequence forms part of an extended helix in IP_3R. 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 β 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 (σ = 0.82); b12 library average = 0.33 (σ = 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 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_3R [76] show that IP_3R1 is markedly enriched in Purkinje neurons in the cerebellum, IP_3R2 is overwhelmingly found in glial cells and IP_3R3 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_3R 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_3R1 rabbit polyclonal Ab (LifeSpan Biosciences LS-B2627) and appropriate secondary Abs. The commercial anti-IP_3R1 polyclonal was raised against a synthetic peptide corresponding to C-terminal residues of human IP_3R1, 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_3R1 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 is localized to the nucleus [77]. The staining patterns of both 4E10 and the anti-IP_3R1 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 □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 α-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-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 lipidic 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 interactions. 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_4^{2-} and PO_4^{2-} 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 7Å away from the sulfate site, and the main-chain nitrogen atoms 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 others 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 reveal 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_3R2, with additional hits from two related IP_3R types (IP_3R1 and IP_3R3) within the top-scoring five peptides (**Fig. 8**), providing evidence that one, two or all three IP_3R 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 $-Log_{10}$ *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_3R were identified as top-scoring PhIP-Seq hits and validated by SPR. In neuronal cell lines, IP_3R1 and IP_3R3 are expressed predominantly in the ER and perinuclear regions, while IP_3R2 is expressed primarily in the nucleus [85]. IP_3R3 shows an evenly diffuse cytoplasmic staining pattern, but the IP_3R1 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_3Rs) prevented comparative immunofluorescence analyses. We therefore sought to determine if the PhIP-Seq results identifying IP_3R1, IP_3R2 and IP_3R3 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_3R1 control polyclonal antibody. The 4E10 staining pattern observed was most consistent in terms of cell specificity with IP_3R1 IHC patterns reported in previous studies [76], strongly arguing that IP_3R1 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_3R, 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_3R 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 18g 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 X 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 X 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')$_2$) in B cell media, in a 37℃ CO_2 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 herein have been reviewed and approved by the Fred Hutchinson Cancer Research Center (FHCRC) Institutional Animal Care and Use Committee (IACUC), file #1695. FHCRC 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-α-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-α-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 μm polycarbonate membrane (Avanti) 21 times.

Protein expression and purification

HCDR3-grafted Ubq fusion constructs (**Fig. 3**) with C-terminal His_6-purification tags were engineered, cytoplasmically expressed in *E. coli* BL21(DE3) RIL cells (Invitrogen), and purified employing the methodology described in [92]. 4E10 [G(L50)E] 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℃ 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 μg/mL, 5 μL/min, until approximately 1000 SPR response units (RUs) were captured, then allowed to equilibrate for four hours at 10 μL/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) ± 2.5 mM $CaCl_2$, 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

μM (**Fig. 3D**), were injected in duplicate analyses at 50 μL/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 μg/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 μM and 5 or 10 μM (depending on maximum responses) at 50 μL/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 μL/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 μL/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℃ by the hanging-drop vapor-diffusion method with a well solution of 0.8 M Li_2SO_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 CuKα radiation on an in-house R-AXIS IV++ image plate detector with HR optics (Rigaku) at -170℃ and at the Advanced Li ght 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 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 μL 0.1 M $NaPO_4$, pH 7.4. Beads were rocked at ambient temperatures for 24 hrs with 60 μg of either 4E10 or b12 IgG in 1 M $(NH_4)_2SO_4$ 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 $-Log_{10}$ *P*-values.

IHC

Four micron sections of formalin-fixed and paraffin-embedded tissues were cut and baked for 1 hr at 60℃, 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_3R1 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_3R1 IgG (LifeSpan Biosciences) at a dilution of 1:200 (2.5 μg/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 μg/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.

Acknowledgements

I would like to thank Roland K. Strong[1] for his assistance with data analysis and manuscript preparation; Kevin Larimore[4] and Philip D. Greenberg[2,3,4] for their work on mouse knock-in experiments; H. Benjamin Larman[5] and Stephen J. Elledge[5] for their work on PhIP-Seq experiments; Della Friend[1] for her assistance with SPR experiments and analysis; Chris Carrico for suggesting Ubq as an HCDR3 scaffold; Colin Correnti[1], Bonnie Cuthbert[1], and Madeline LaPatra[1] for assistance with protein production; Peter B. Rupert[1] for his assistance with crystallography; 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.

[1]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

[2]Department of Immunology, University of Washington, Seattle, WA 98195

[3]Department of Medicine, University of Washington, Seattle, WA 98195

[4]Program in Immunology, Cancer Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

[5]Department of Genetics, Harvard University Medical School, and Division of Genetics, Howard Hughes Medical Institute, Brigham and Women's Hospital, Boston, MA 02115

Figures

Table 1: Crystallographic data collection and refinement statistics

	Synchrotron	In-house
Data collection*		
Space group	$C222_1$	$C222_1$
Cell dimensions		
a, *b*, *c* (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)
R_{merge}	0.072 (0.382)	0.066 (0.438)
$I / \sigma I$	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)
Refinement		
No. reflections (unique)	46,684 (2,305)	49,263 (4,863)
R_{work} / R_{free}	0.216/0.245	--
No. non-hydrogen atoms		
Protein	6,452	--
Ligand/ion	81	--
Water	296	--
B-factors (Å^2)		
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 ESUc; Å)	0.161	--
Accession code	4LLV	

* Values in parentheses are for the highest-resolution shell.

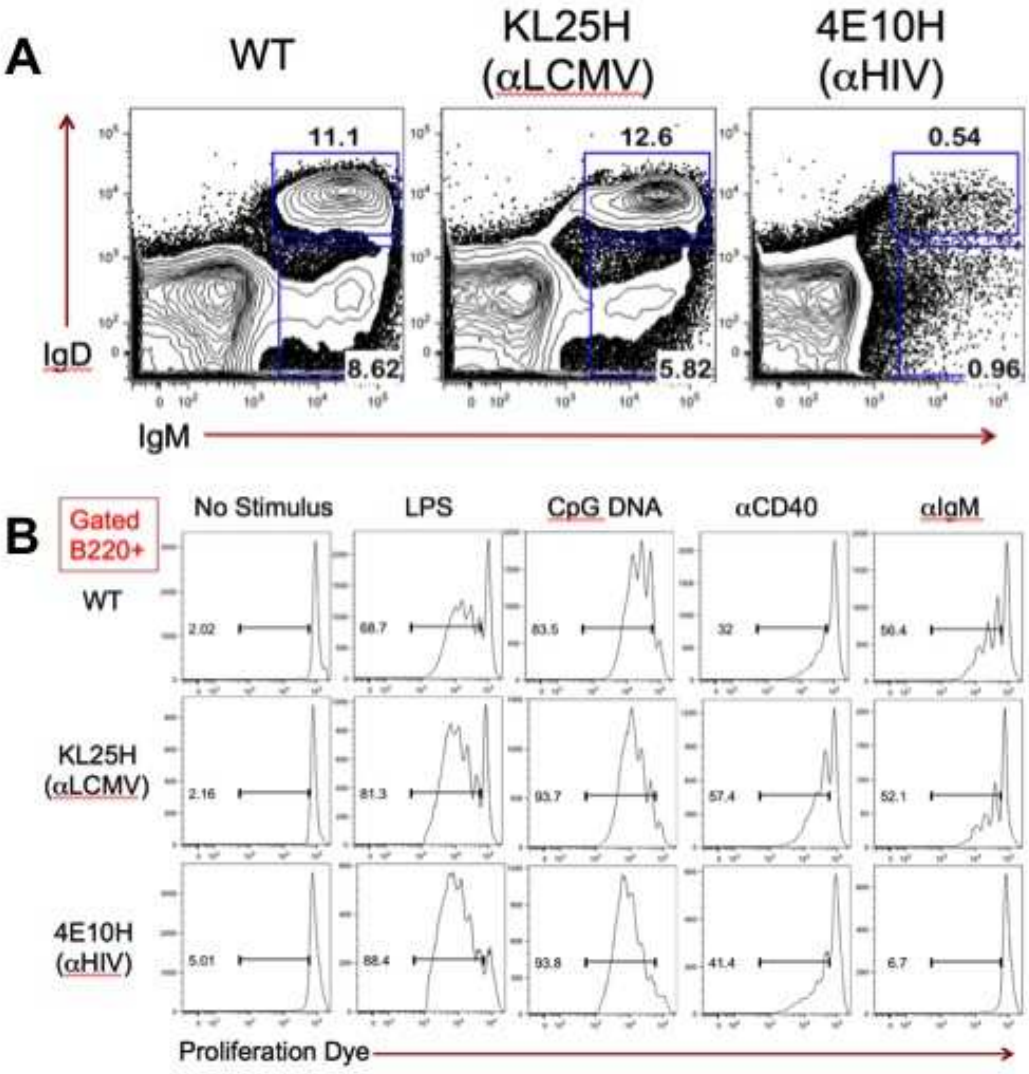

Figure 1. Analysis of 4E10H knock-in mice. (**A**) The percentages of IgM^+IgD^- pre-B and IgM^+IgD^+ immature B cells within the $B220^+$ cell population in the bone marrow of B6 WT, B6-KL25H and B6-4E10H mice are shown. Marrow cells were isolated and stained for cell surface markers as described in the methods section. FACS plots were previously gated $B220^+$ and on live cell size based on forward and side scatter. Numbers represent the percentages of $B220^+$ bone marrow cells falling within each box gate. (**B**) The proliferation of $B220^+$ 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 $B220^+$ B cells. Histograms were previously gated $B220^+$ and on live cell size based on forward and side scatter. Numbers represent the percentage of $B220^+$ cells falling within the bar gate on each plot, indicating dilution of the proliferation dye as a result of cell division.

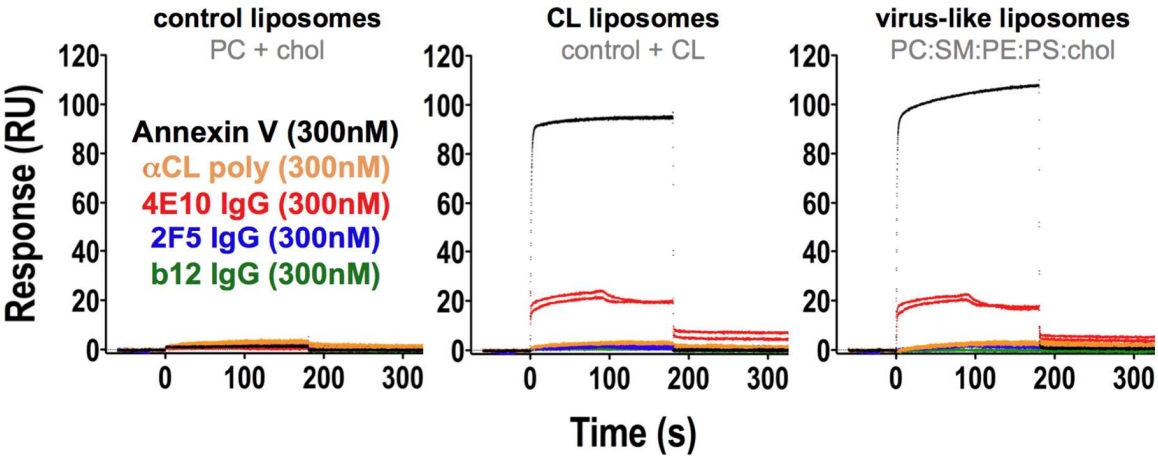

Figure 2. SPR analyses of 4E10/liposome interactions. Corrected SPR responses are shown (response units: RU; duplicate runs) of Annexin V (black), αCL polyclonal Ab (orange), 4E10 IgG (red), 2F5 IgG (blue), and b12 IgG (green) analytes binding to liposomes incorporating biotinylated lipids captured on streptavidin-coated biosensor chips. All analyte concentrations were 300 nM. The Annexin V analyte buffer included 2.5 mM $CaCl_2$. Liposome compositions are indicated above each frame.

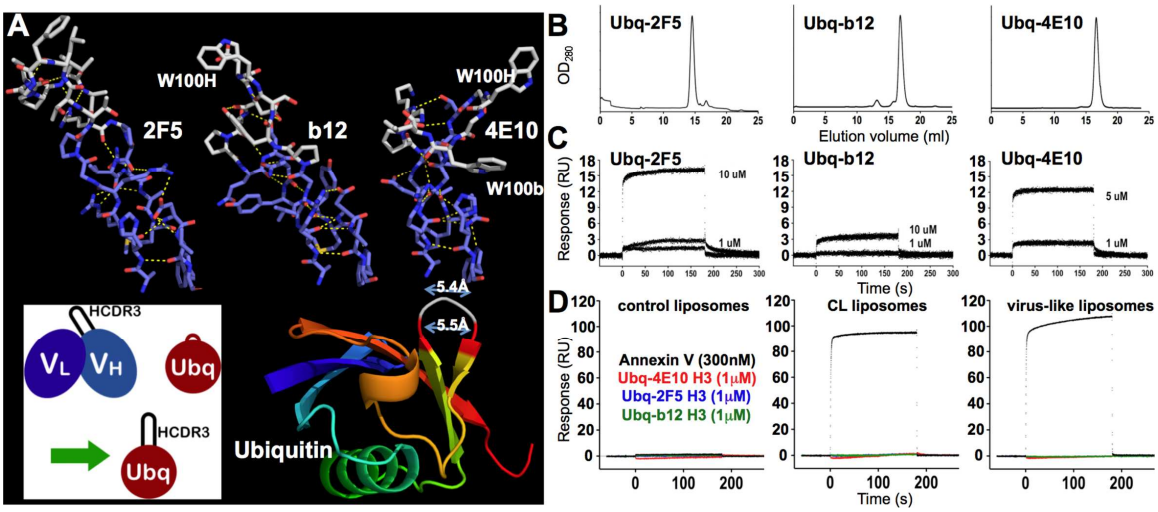

Figure 3. Design and characterization of HCDR3-grafted Ubq fusion proteins. (**A**) A ribbon representation of Ubq (1UBQ.pdb [106]) is shown at bottom right, colored prismatically from N- (blue) to C-terminus (red); the targeted loop for HCDR3 engraftment is highlighted in gray with red bounding residues. 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: AHRRGP**TT**T**LFGVPIARG**PVNAMDVW; b12: ARVG**PYSWDDSP**QYNYYMDVW; 4E10: AREGTT**GWGWLGKP**IGAFAHW; exposed residues **bolded**), were characterized as "hydrophobic" with the Sigma-Aldrich PEPscreen Library Design Tool (Φ_{2F5} = 0.52; Φ_{b12} = 0.51; Φ_{4E10} = 0.56). Key residues and Cα-Cα distances (Å) 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 monodispersivity. (**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 >>10 μM 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.

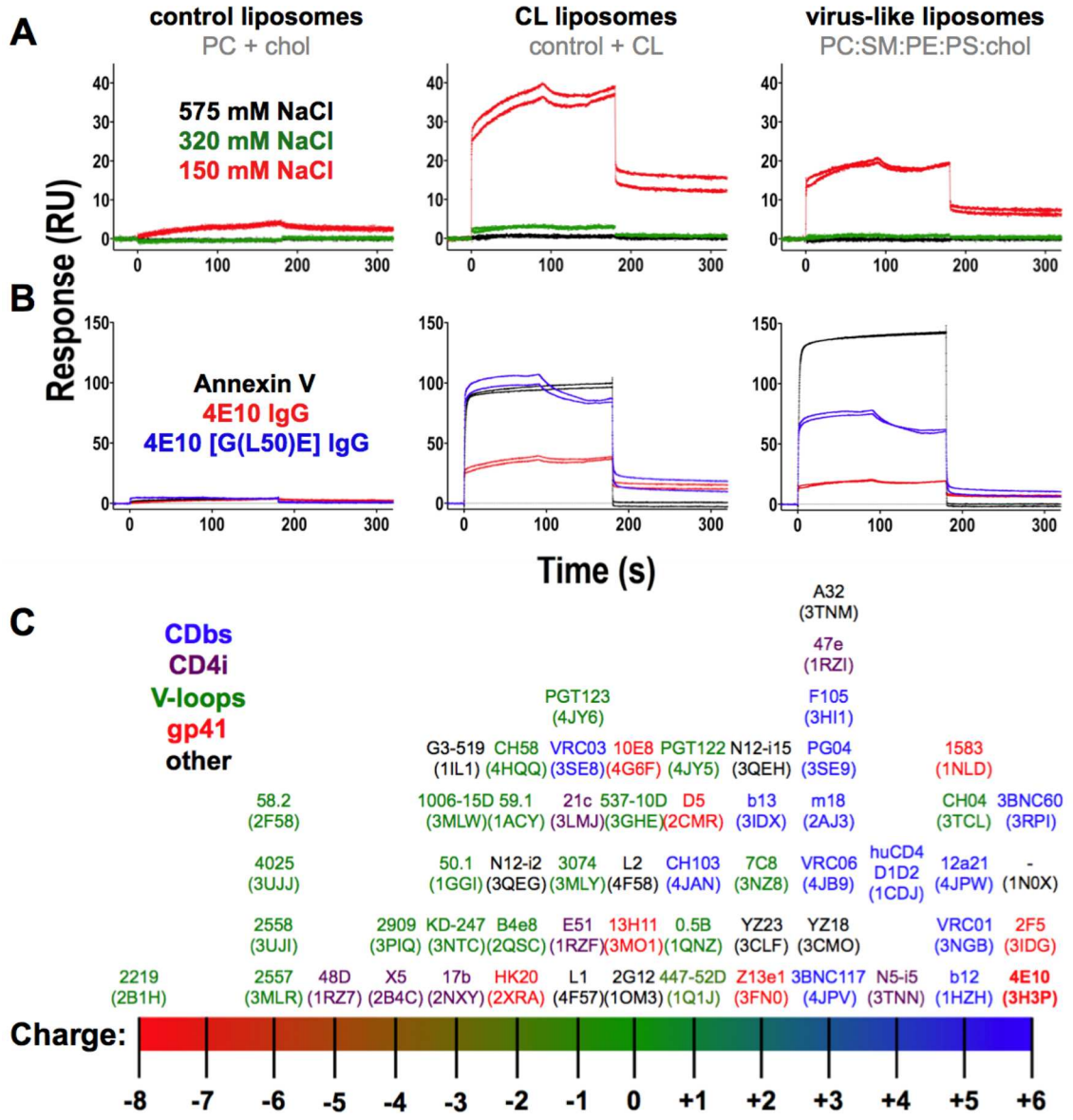

Figure 4. Analysis of the effects of salt concentration, mutation and overall cassette charge on 4E10/liposome interactions. Corrected SPR responses are shown for Annexin V or 4E10 IgG analytes (300 nM; duplicate runs) to liposomes incorporating biotinylated lipids captured on streptavidin-coated biosensor 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.

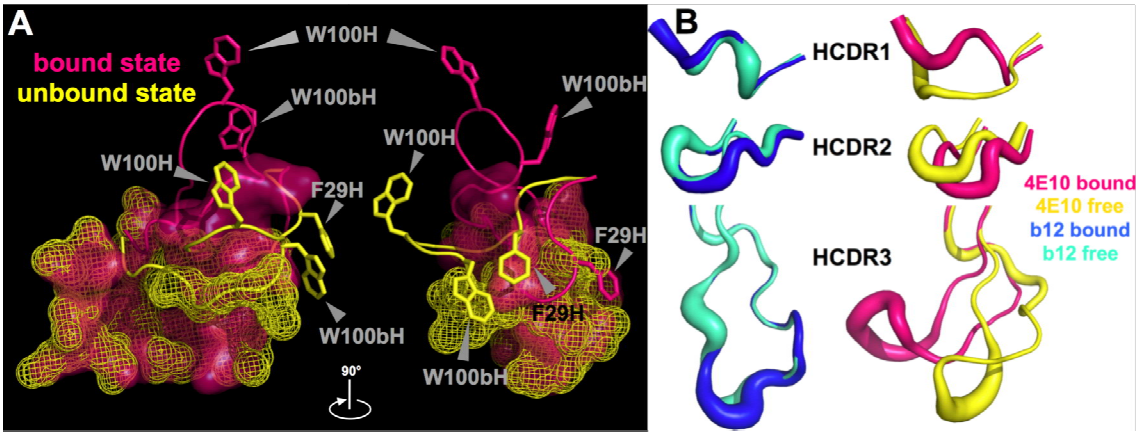

Figure 5. Structure of 4E10 free of bound antigen. (**A**) Superposition of residues from the 4E10 epitope binding site and HCDR1 and 3 from bound (semi-transparent molecular surface in pink; 2FX7.pdb [9]) and unbound (mesh molecular surface in yellow; 4LLV.pdb) 4E10 Fv are shown; HCDR1 and 3 are shown 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]. (**B**) 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).

Figure 6. IF microscopy of HEp-2 cells. Fixed and permeabilized HEp-2 cells were stained with (**A**) 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.

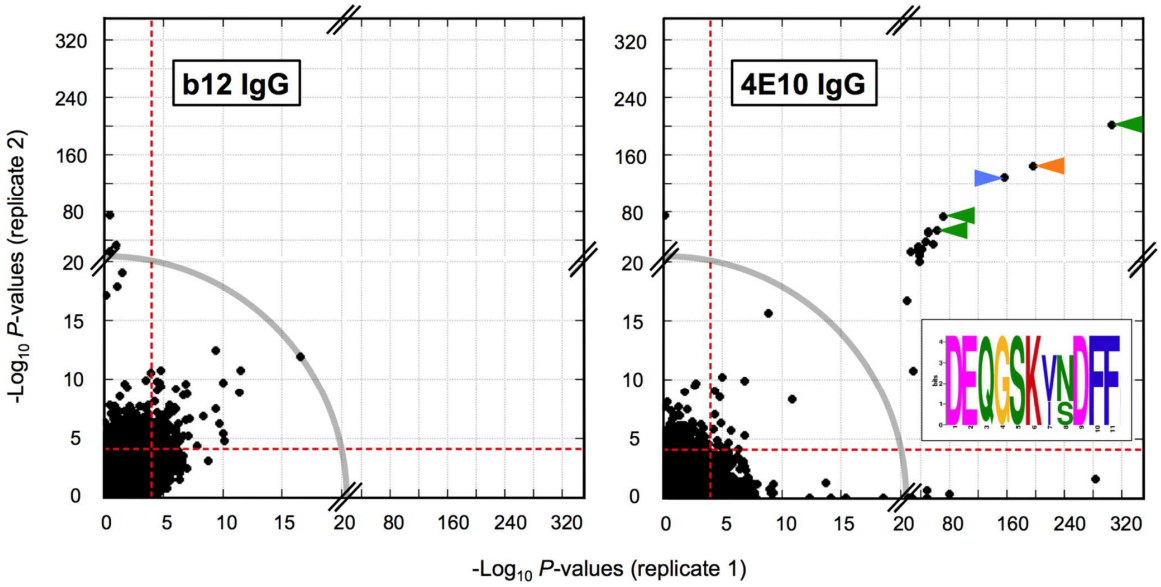

Figure 7. PhIP-Seq analyses of 4E10 and b12. PhIP-Seq results are plotted as $-\text{Log}_{10}$ *P* values, one replicate on the abscissa, the other on the ordinate; note the discontinuity in axis scales. The top five scoring 4E10 peptides are highlighted with arrows; the three peptides derived from IP_3R 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.

Rank	Avg. -log*P*	Peptide sequence	Parent protein	Φ	Charge
b12 IgG			*average*:	0.1 ± 0.7	+9 ± 4
1	14.3	YKVDVTWTRARGASRGWRSRHQLKGRNGWRNSRVYK	TP53-target gene 5 protein	0.12	+10.1
2	11.1	KKAKVAKPKKAAKSAAKAVKPKAAKPKVVKPKKAAP	histone H1.2	-0.08	+15
3	11.0	KSPQTWPRRTRPRSPEPAAPSGVRGSTWTRRRDTPR	metaxin-1 isoform 1	0.02	+8
4	10.2	NEEGKENRDRDRDYSRRRGGPPRRGRGASRGREFRG	ubiquitin associated protein 2-like	-0.34	+6
5	9.89	AVLKYENNVMNIRQFNCSPHPYWLPNFMDVFTWSLP	serine/threonine-protein phosphatase 2B catalytic subunit α isoform 1	0.81	- 0 -
6	8.41	SSSSSQSRSRSSSRERSRSRGSKSRSSSRSHRGSSS	zinc finger Ran-binding domain-containing protein 2 isoform 2	-0.32	+9.1
7	8.26	AMAGRIYISGMAPRPSLAKKQRFRHRNRKGYRSQRG	vitronectin precursor	0.09	+11.1
8	7.96	RTATQRKIFQCNKHMKVFHKYSNRNKVRHTKKKTFK	zinc finger protein 729	0.10	+13.2
9	7.84	PKGEPTRRGRGGTFRRGGRDPGGRPSRPSTLRRPAY	protein PRRC2C	-0.08	+9
10	7.80	HRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKR	epidermal growth factor-like protein 7 precursor	0.29	+9
11	7.68	QHPFYNRPFFAPYLYTPRYYPGGSQHLISRPSVKTS	kinase D-interacting substrate of 220 kDa	0.50	+4.2
12	7.68	AKYNPRKHRAKRHPRRPPRSPGMEPPFSHRCFPRYI	telomerase protein component 1	0.16	+10.2
13	7.63	AKKVKTPQPKKAAKSPAKAKAPKPKAAKPKSGKPKV	histone H1.3	-0.11	+14
14	7.61	RRKAPTKKSPLSGVSQEGKQVQRNAANARERARMRV	transcription factor 21	-0.06	+9
15	7.49	SPPKSLFSQRPQRMRSRSRSFSRHRSCSRSPYSRSR	peroxisome proliferator-activated receptor gamma coactivator 1α	0	+11
4E10 IgG			*average*:	0.4 ± 0.1	-1 ± 2
1*	255	SKCRVFNTTERDEQGSKVNDFFQQTEDLYNEMKWQK	**inositol 1,4,5-trisphosphate receptor type 2**	0.17	-1
2*	170	PPHELTEEEKQQILHSEEFLSFFDHSTRIVERALSE	cytoplasmic dynein 1 intermediate chain 2	0.28	-5.7
3*	143	PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEIVR	complement receptor type 1 isoform S precursor	0.48	+1.9
4*	72.5	TERDEQGSKINDFFLRSEDLFNEMNWQKKLRAQPVL	**inositol 1,4,5-trisphosphate receptor type 1 isoform 2**	0.28	-1
5*	58.0	LTEETKHRLFTTTEQDEQGSKVSDFFDQSSFLHNEM	**inositol 1,4,5-trisphosphate receptor type 3**	0.21	-4.8
6	51.4	TFGHLGRLRELSLRNNALSALSGDIFAASPALYRLD	TLR4 interactor with leucine rich repeats precursor	0.48	+1.1
7	50.5	ASSRHQRIQDFNYTDHTLGRIILNAMNETNFFGVTG	γ-aminobutyric acid type B receptor subunit 2 precursor	0.45	+0.2
8	46.0	GAQPPFDAQSPLDSQPQPSGQPWNFHASTSWYWRQS	nuclear fragile X mental retardation-interacting protein 1	0.42	-0.9
9	42.3	IITVNACSQSDQLNPEPGENSISEEEYSKNWFTVSK	transient receptor potential cation channel subfamily M member 6 isoform a	0.39	-4
10	35.3	ESSQHTLNYWKEQSLNVSQDLDTIRSNINNFFEFSK	nesprin-2 isoform 1	0.41	-1.9
11	34.1	RLLLCSAATDGSLAFWDLTTMLDHDSTVLEPPVDPG	WD repeat-containing protein 6	0.55	-5
12	33.9	NGHINFKKFWEISRQIHEFMTWTQVECPFEKDKKIQ	ras-GEF domain-containing family member 1A	0.43	+1.1
13	31.8	VYEYKYGPCANHVEPSENCLKGYVNNSLSFFDLSEL	anoctamin-3	0.55	-3
14	29.3	PGYENDSVEDLKEVTSISSRKRGKRRYFWEYSEQLT	UPF0474 protein C5orf41 isoform 1	0.12	- 0 -
15	25.4	AAEPVRAKLDRQRRVFQPSPLASQFELPGDFFNLTA	UBX domain-containing protein 6 isoform 1	0.36	+1
16	20.2	QFNTLYLRVRGDGRPWMVNIKEDTDFFQRTNQMYSY	complex I intermediate-associated protein 30, mitochondrial precursor	0.42	+1
17	18.4	TTGEYKSTFNGNRPSSSDRHLIPVAFVSEKWFEISC	adherens junction-associated protein 1	0.41	- 0 -

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 (Φ) 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_3R isoform peptides (labeled in bold), identified by MEME [72], is underlined. Asterisks indicate peptide sequences selected for expression as fusion constructs for subsequent validation.

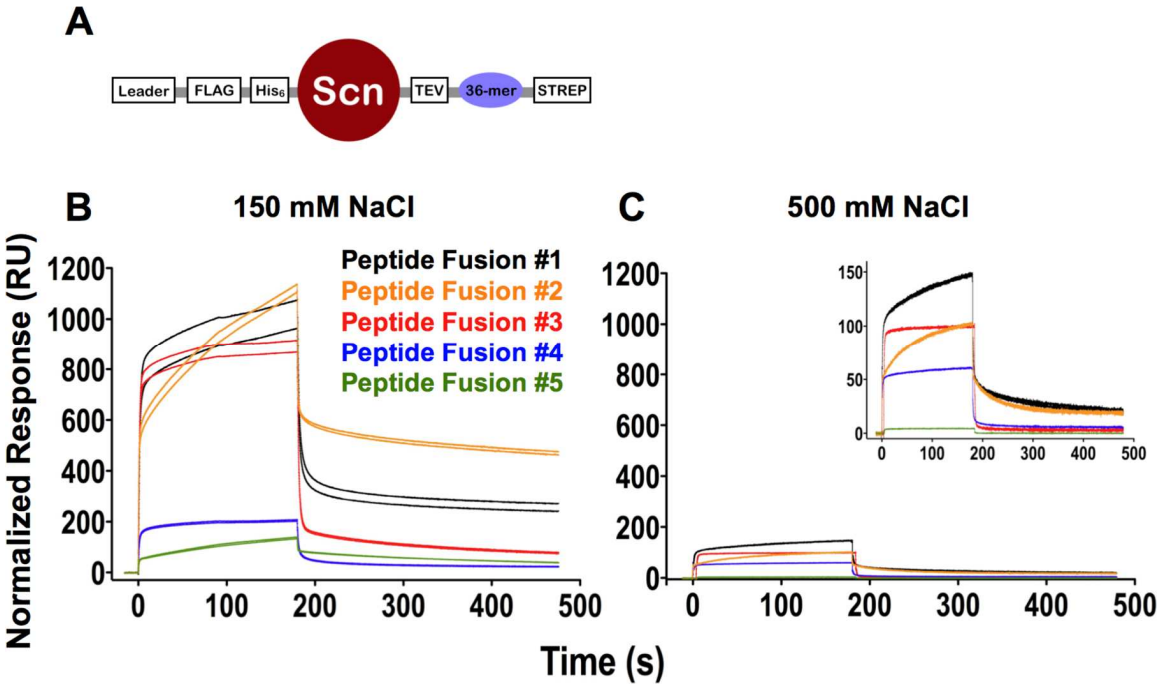

Figure 9. SPR analyses of Scn-peptide fusions. (**A**) The generalized siderocalin fusion construct used for PhIP-Seq peptide expression is shown. (**B**) Corrected and normalized SPR responses for 4E10 IgG binding to the top five peptides immobilized on a CM5 biosensor chip are plotted. (**C**) Corrected 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.

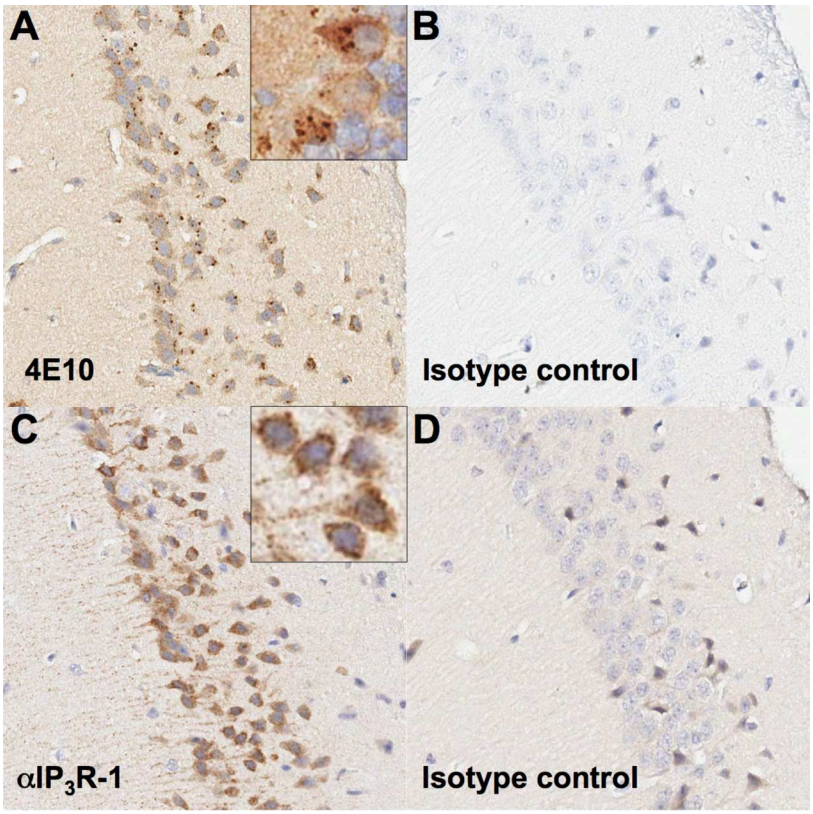

Figure 10. IHC of serial mouse cerebellar sections. (**A**) Immunohistochemical staining with 4E10 IgG in mouse cerebellum sections reveals perinuclear, nodular immunoreactivity in Purkinje neurons; inset shows a magnified view. (**B**) Human IgG Isotype control staining of mouse cerebellum is shown. (**C**) Staining for IP_3R1 with an anti-IP_3R1 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.

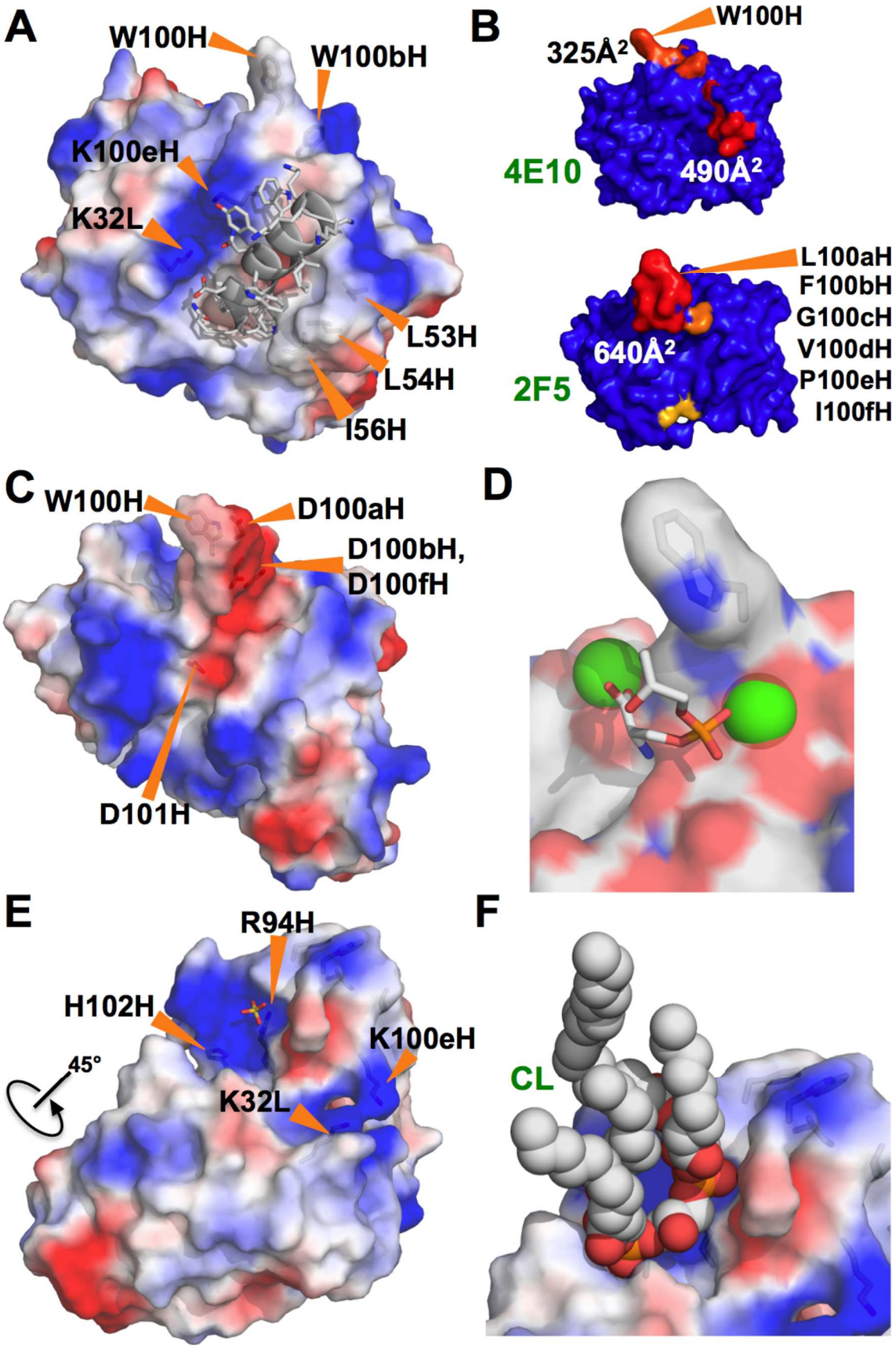
A
W100H
W100bH
K100eH
K32L
L53H
L54H
I56H
B
W100H
325Å²
4E10
490Å²
L100aH
F100bH
G100cH
V100dH
P100eH
I100fH
640Å²
2F5
C
W100H
D100aH
D100bH,
D100fH
D101H
D
E
R94H
H102H
K100eH
45°
K32L
F
CL

Figure 11: Structures of 4E10, b12, 2F5, 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_L domain is in the upper left and the V_H 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_L domains at left and the V_H 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_4 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].

References

1. Wyatt, R. & J. Sodroski (1998) 'The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens' *Science* **280**(5371), p. 1884-8
2. Montefiori, D., Q. Sattentau, J. Flores, J. Esparza & J. Mascola (2007) 'Antibody-based HIV-1 vaccines: recent developments and future directions' *PLoS Med* **4**(12), p. e348
3. Schief, W.R., Y.E. Ban & L. Stamatatos (2009) 'Challenges for structure-based HIV vaccine design' *Curr Opin HIV AIDS* **4**(5), p. 431-40
4. Zwick, M.B., A.F. Labrijn, M. Wang, C. Spenlehauer, E.O. Saphire, J.M. Binley, J.P. Moore, G. Stiegler, H. Katinger, D.R. Burton & P.W. Parren (2001) 'Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41' *J Virol* **75**(22), p. 10892-905
5. D'Souza, M.P., S.J. Geyer, C.V. Hanson, R.M. Hendry & G. Milman (1994) 'Evaluation of monoclonal antibodies to HIV-1 envelope by neutralization and binding assays: an international collaboration' *AIDS* **8**(2), p. 169-81
6. Stiegler, G., R. Kunert, M. Purtscher, S. Wolbank, R. Voglauer, F. Steindl & H. Katinger (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**(18), p. 1757-65
7. Binley, J.M., T. Wrin, B. Korber, M.B. Zwick, M. Wang, C. Chappey, G. Stiegler, R. Kunert, S. Zolla-Pazner, H. Katinger, C.J. Petropoulos & D.R. Burton (2004) 'Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies' *J Virol* **78**(23), p. 13232-52

8. Brunel, F.M., M.B. Zwick, R.M. Cardoso, J.D. Nelson, I.A. Wilson, D.R. Burton & P.E. Dawson (2006) 'Structure-function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody' *J Virol* **80**(4), p. 1680-7

9. Cardoso, R.M., F.M. Brunel, S. Ferguson, M. Zwick, D.R. Burton, P.E. Dawson & I.A. Wilson (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**(5), p. 1533-44

10. Cardoso, R.M., M.B. Zwick, R.L. Stanfield, R. Kunert, J.M. Binley, H. Katinger, D.R. Burton & I.A. Wilson (2005) 'Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41' *Immunity* **22**(2), p. 163-73

11. Walker, L.M., M. Huber, K.J. Doores, E. Falkowska, R. Pejchal, J.P. Julien, S.K. Wang, A. Ramos, P.Y. Chan-Hui, M. Moyle, J.L. Mitcham, P.W. Hammond, O.A. Olsen, P. Phung, S. Fling, C.H. Wong, S. Phogat, T. Wrin, M.D. Simek, W.C. Koff, I.A. Wilson, D.R. Burton & P. Poignard (2011) 'Broad neutralization coverage of HIV by multiple highly potent antibodies' *Nature* **477**(7365), p. 466-70

12. Walker, L.M., S.K. Phogat, P.Y. Chan-Hui, D. Wagner, P. Phung, J.L. Goss, T. Wrin, M.D. Simek, S. Fling, J.L. Mitcham, J.K. Lehrman, F.H. Priddy, O.A. Olsen, S.M. Frey, P.W. Hammond, G. Miiro, J. Serwanga, A. Pozniak, D. McPhee, O. Manigart, L. Mwananyanda, E. Karita, A. Inwoley, W. Jaoko, J. Dehovitz, L.G. Bekker, P. Pitisuttithum, R. Paris, S. Allen, S. Kaminsky, T. Zamb, M. Moyle, W.C. Koff, P. Poignard & D.R. Burton (2009) 'Broad and Potent Neutralizing Antibodies from an African Donor Reveal a New HIV-1 Vaccine Target' *Science*

13. Li, Y., K. Svehla, M.K. Louder, D. Wycuff, S. Phogat, M. Tang, S.A. Migueles, X. Wu, A. Phogat, G.M. Shaw, M. Connors, J. Hoxie, J.R. Mascola & R. Wyatt (2009) 'Analysis of

neutralization specificities in polyclonal sera derived from human immunodeficiency virus type 1-infected individuals' *J Virol* **83**(2), p. 1045-59

14. Sather, D.N., J. Armann, L.K. Ching, A. Mavrantoni, G. Sellhorn, Z. Caldwell, X. Yu, B. Wood, S. Self, S. Kalams & L. Stamatatos (2009) 'Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection' *J Virol* **83**(2), p. 757-69
15. Muster, T., F. Steindl, M. Purtscher, A. Trkola, A. Klima, G. Himmler, F. Ruker & H. Katinger (1993) 'A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1' *J Virol* **67**(11), p. 6642-7
16. Haynes, B.F., M.A. Moody, L. Verkoczy, G. Kelsoe & S.M. Alam (2005) 'Antibody polyspecificity and neutralization of HIV-1: a hypothesis' *Hum Antibodies* **14**(3-4), p. 59-67
17. Verkoczy, L., G. Kelsoe, M.A. Moody & B.F. Haynes (2011) 'Role of immune mechanisms in induction of HIV-1 broadly neutralizing antibodies' *Curr Opin Immunol* **23**(3), p. 383-90
18. Alam, S.M., M. McAdams, D. Boren, M. Rak, R.M. Scearce, F. Gao, Z.T. Camacho, D. Gewirth, G. Kelsoe, P. Chen & B.F. Haynes (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**(7), p. 4424-35
19. Haynes, B.F., J. Fleming, E.W. St Clair, H. Katinger, G. Stiegler, R. Kunert, J. Robinson, R.M. Scearce, K. Plonk, H.F. Staats, T.L. Ortel, H.X. Liao & S.M. Alam (2005) 'Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies' *Science* **308**(5730), p. 1906-8
20. Yang, G., T.M. Holl, Y. Liu, Y. Li, X. Lu, N.I. Nicely, T.B. Kepler, S.M. Alam, H.X. Liao, D.W. Cain, L. Spicer, J.L. Vandeberg, B.F. Haynes & G. Kelsoe (2013) 'Identification of

autoantigens recognized by the 2F5 and 4E10 broadly neutralizing HIV-1 antibodies' *J Exp Med*

21. Alam, S.M., M. Morelli, S.M. Dennison, H.X. Liao, R. Zhang, S.M. Xia, S. Rits-Volloch, L. Sun, S.C. Harrison, B.F. Haynes & B. Chen (2009) 'Role of HIV membrane in neutralization by two broadly neutralizing antibodies' *Proc Natl Acad Sci U S A* **106**(48), p. 20234-9
22. Scherer, E.M., D.P. Leaman, M.B. Zwick, A.J. McMichael & D.R. Burton (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**(4), p. 1529-34
23. Cohn, M. (2008) 'An in depth analysis of the concept of "polyspecificity" assumed to characterize TCR/BCR recognition' *Immunol Res* **40**(2), p. 128-47
24. Wucherpfennig, K.W., P.M. Allen, F. Celada, I.R. Cohen, R. De Boer, K.C. Garcia, B. Goldstein, R. Greenspan, D. Hafler, P. Hodgkin, E.S. Huseby, D.C. Krakauer, D. Nemazee, A.S. Perelson, C. Pinilla, R.K. Strong & E.E. Sercarz (2007) 'Polyspecificity of T cell and B cell receptor recognition' *Semin Immunol*
25. Herzog, S. & H. Jumaa (2012) 'Self-recognition and clonal selection: autoreactivity drives the generation of B cells' *Curr Opin Immunol* **24**(2), p. 166-72
26. Shlomchik, M.J. (2008) 'Sites and stages of autoreactive B cell activation and regulation' *Immunity* **28**(1), p. 18-28
27. Vettermann, C. & H.M. Jack (2010) 'The pre-B cell receptor: turning autoreactivity into self-defense' *Trends Immunol* **31**(5), p. 176-83
28. Verkoczy, L., M. Diaz, T.M. Holl, Y.B. Ouyang, H. Bouton-Verville, S.M. Alam, H.X. Liao, G. Kelsoe & B.F. Haynes (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**(1), p. 181-6

29. Kallenberg, C.G., J. van der Meulen, G.W. Pastoor, J.A. Snijder, T.E. Feltkamp & T.H. The (1983) 'Human fibroblasts, a convenient nuclear substrate for detection of anti-nuclear antibodies including anti-centromere antibodies' *Scand J Rheumatol* **12**(3), p. 193-200

30. Houtkooper, R.H. & F.M. Vaz (2008) 'Cardiolipin, the heart of mitochondrial metabolism' *Cell Mol Life Sci* **65**(16), p. 2493-506

31. Osman, C., D.R. Voelker & T. Langer (2011) 'Making heads or tails of phospholipids in mitochondria' *J Cell Biol* **192**(1), p. 7-16

32. Suh-Lailam, B.B., A. Cromar, K.W. Davis & A.E. Tebo (2012) 'APhL antibody ELISA as an alternative to anticardiolipin test for the diagnosis of antiphospholipid syndrome' *Int J Clin Exp Pathol* **5**(3), p. 210-5

33. Scherer, E.M., M.B. Zwick, L. Teyton & D.R. Burton (2007) 'Difficulties in eliciting broadly neutralizing anti-HIV antibodies are not explained by cardiolipin autoreactivity' *Aids* **21**(16), p. 2131-9

34. Xu, H., L. Song, M. Kim, M.A. Holmes, Z. Kraft, G. Sellhorn, E.L. Reinherz, L. Stamatatos & R.K. Strong (2010) 'Interactions between lipids and human anti-HIV antibody 4E10 can be reduced without ablating neutralizing activity' *J Virol* **84**(2), p. 1076-88

35. Tong, T., E.T. Crooks, K. Osawa & J.M. Binley (2012) 'HIV-1 virus-like particles bearing pure env trimers expose neutralizing epitopes but occlude nonneutralizing epitopes' *J Virol* **86**(7), p. 3574-87

36. Ferguson, C.G., R.D. James, C.S. Bigman, D.A. Shepard, Y. Abdiche, P.S. Katsamba, D.G. Myszka & G.D. Prestwich (2005) 'Phosphoinositide-containing polymerized liposomes: stable membrane-mimetic vesicles for protein-lipid binding analysis' *Bioconjug Chem* **16**(6), p. 1475-83

37. Martinez, V., M.C. Diemert, M. Braibant, V. Potard, J.L. Charuel, F. Barin, D. Costagliola, E. Caumes, J.P. Clauvel, B. Autran & L. Musset (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**(1), p. 123-32

38. Sene, D., J.C. Piette & P. Cacoub (2009) '[Antiphospholipid antibodies, antiphospholipid syndrome and viral infections]' *Rev Med Interne* **30**(2), p. 135-41

39. Galrao, L., C. Brites, M.L. Atta, A. Atta, I. Lima, F. Gonzalez, F. Magalhaes & M. Santiago (2007) 'Antiphospholipid antibodies in HIV-positive patients' *Clin Rheumatol* **26**(11), p. 1825-30

40. Vcelar, B., G. Stiegler, H.M. Wolf, W. Muntean, B. Leschnik, S. Mehandru, M. Markowitz, C. Armbruster, R. Kunert, M.M. Eibl & H. Katinger (2007) 'Reassessment of autoreactivity of the broadly neutralizing HIV antibodies 4E10 and 2F5 and retrospective analysis of clinical safety data' *AIDS* **21**(16), p. 2161-70

41. Singh, H., K.A. Henry, S.S. Wu, A. Chruscinski, P.J. Utz & J.K. Scott (2011) 'Reactivity profiles of broadly neutralizing anti-HIV-1 antibodies are distinct from those of pathogenic autoantibodies' *AIDS* **25**(10), p. 1247-57

42. Burton, D.R., C.F. Barbas, 3rd, M.A. Persson, S. Koenig, R.M. Chanock & R.A. Lerner (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**(22), p. 10134-7

43. Roben, P., J.P. Moore, M. Thali, J. Sodroski, C.F. Barbas, 3rd & D.R. Burton (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**(8), p. 4821-8

44. Sajadi, M.M., G.K. Lewis, M.S. Seaman, Y. Guan, R.R. Redfield & A.L. DeVico (2012) 'Signature biochemical properties of broadly cross-reactive HIV-1 neutralizing antibodies in human plasma' *J Virol* **86**(9), p. 5014-25

45. Larman, H.B., Z. Zhao, U. Laserson, M.Z. Li, A. Ciccia, M.A. Gakidis, G.M. Church, S. Kesari, E.M. Leproust, N.L. Solimini & S.J. Elledge (2011) 'Autoantigen discovery with a synthetic human peptidome' *Nat Biotechnol* **29**(6), p. 535-41

46. Hangartner, L., B.M. Senn, B. Ledermann, U. Kalinke, P. Seiler, E. Bucher, R.M. Zellweger, K. Fink, B. Odermatt, K. Burki, R.M. Zinkernagel & H. Hengartner (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**(22), p. 12883-8

47. Hodnik, V. & G. Anderluh (2010) 'Capture of intact liposomes on biacore sensor chips for protein-membrane interaction studies' *Methods Mol Biol* **627**, p. 201-11

48. Harris, E.N. & S.S. Pierangeli (2002) 'Revisiting the anticardiolipin test and its standardization' *Lupus* **11**(5), p. 269-75

49. Myszka, D.G. (1999) 'Improving biosensor analysis' *J Mol Recognit* **12**(5), p. 279-84

50. Brown, B.K., N. Karasavvas, Z. Beck, G.R. Matyas, D.L. Birx, V.R. Polonis & C.R. Alving (2007) 'Monoclonal antibodies to phosphatidylinositol phosphate neutralize human immunodeficiency virus type 1: role of phosphate-binding subsites' *J Virol* **81**(4), p. 2087-91

51. Dolinsky, T.J., P. Czodrowski, H. Li, J.E. Nielsen, J.H. Jensen, G. Klebe & N.A. Baker (2007) 'PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations' *Nucleic Acids Res* **35**(Web Server issue), p. W522-5

52. Dolinsky, T.J., J.E. Nielsen, J.A. McCammon & N.A. Baker (2004) 'PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations' *Nucleic Acids Res* **32**(Web Server issue), p. W665-7

53. Li, H., A.D. Robertson & J.H. Jensen (2005) 'Very fast empirical prediction and rationalization of protein pKa values' *Proteins* **61**(4), p. 704-21

54. Unni, S., Y. Huang, R.M. Hanson, M. Tobias, S. Krishnan, W.W. Li, J.E. Nielsen & N.A. Baker (2011) 'Web servers and services for electrostatics calculations with APBS and PDB2PQR' *J Comput Chem* **32**(7), p. 1488-91

55. Patten, P.A., N.S. Gray, P.L. Yang, C.B. Marks, G.J. Wedemayer, J.J. Boniface, R.C. Stevens & P.G. Schultz (1996) 'The immunological evolution of catalysis' *Science* **271**(5252), p. 1086-91

56. Wedemayer, G.J., P.A. Patten, L.H. Wang, P.G. Schultz & R.C. Stevens (1997) 'Structural insights into the evolution of an antibody combining site' *Science* **276**(5319), p. 1665-9

57. Romesberg, F.E., B. Spiller, P.G. Schultz & R.C. Stevens (1998) 'Immunological origins of binding and catalysis in a Diels-Alderase antibody' *Science* **279**(5358), p. 1929-33

58. Li, Y., H. Li, F. Yang, S.J. Smith-Gill & R.A. Mariuzza (2003) 'X-ray snapshots of the maturation of an antibody response to a protein antigen' *Nat Struct Biol* **10**(6), p. 482-8

59. Terzyan, S., P.A. Ramsland, E.W. Voss, Jr., J.N. Herron & A.B. Edmundson (2004) 'Three-dimensional structures of idiotypically related Fabs with intermediate and high affinity for fluorescein' *J Mol Biol* **339**(5), p. 1141-51

60. Manivel, V., N.C. Sahoo, D.M. Salunke & K.V. Rao (2000) 'Maturation of an antibody response is governed by modulations in flexibility of the antigen-combining site' *Immunity* **13**(5), p. 611-20

61. Sagawa, T., M. Oda, M. Ishimura, K. Furukawa & T. Azuma (2003) 'Thermodynamic and kinetic aspects of antibody evolution during the immune response to hapten' *Mol Immunol* **39**(13), p. 801-8

62. Li, Y., C.A. Lipschultz, S. Mohan & S.J. Smith-Gill (2001) 'Mutations of an epitope hot-spot residue alter rate limiting steps of antigen-antibody protein-protein associations' *Biochemistry* **40**(7), p. 2011-22

63. Mohan, S., N. Sinha & S.J. Smith-Gill (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**(5), p. 3221-36

64. Acierno, J.P., B.C. Braden, S. Klinke, F.A. Goldbaum & A. Cauerhff (2007) 'Affinity maturation increases the stability and plasticity of the Fv domain of anti-protein antibodies' *J Mol Biol* **374**(1), p. 130-46

65. Zimmermann, J., F.E. Romesberg, C.L. Brooks, 3rd & I.F. Thorpe (2010) 'Molecular description of flexibility in an antibody combining site' *J Phys Chem B* **114**(21), p. 7359-70

66. Kallewaard, N.L., B.A. McKinney, Y. Gu, A. Chen, B.V. Prasad & J.E. Crowe, Jr. (2008) 'Functional maturation of the human antibody response to rotavirus' *J Immunol* **180**(6), p. 3980-9

67. Schmidt, A.G., H. Xu, A.R. Khan, T. O'Donnell, S. Khurana, L.R. King, J. Manischewitz, H. Golding, P. Suphaphiphat, A. Carfi, E.C. Settembre, P.R. Dormitzer, T.B. Kepler, R. Zhang, M.A. Moody, B.F. Haynes, H.X. Liao, D.E. Shaw & S.C. Harrison (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**(1), p. 264-9

68. Hager-Braun, C., H. Katinger & K.B. Tomer (2006) 'The HIV-neutralizing monoclonal antibody 4E10 recognizes N-terminal sequences on the native antigen' *J Immunol* **176**(12), p. 7471-81

69. Pruitt, K.D., T. Tatusova, G.R. Brown & D.R. Maglott (2012) 'NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy' *Nucleic Acids Res* **40**(Database issue), p. D130-5

70. Pruitt, K.D., T. Tatusova & D.R. Maglott (2007) 'NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins' *Nucleic Acids Res* **35**(Database issue), p. D61-5

71. Mikoshiba, K. (2002) '[IP3 receptor, a Ca2+ oscilator--role of IP3 receptor in development and neural plasticity]' *Nihon Yakurigaku Zasshi* **120**(1), p. 6P-10P

72. Bailey, T.L. & C. Elkan (1994) 'Fitting a mixture model by expectation maximization to discover motifs in biopolymers' *Proc Int Conf Intell Syst Mol Biol* **2**, p. 28-36

73. Kelley, L.A. & M.J. Sternberg (2009) 'Protein structure prediction on the Web: a case study using the Phyre server' *Nat Protoc* **4**(3), p. 363-71

74. Bandaranayake, A.D., C. Correnti, B.Y. Ryu, M. Brault, R.K. Strong & D.J. Rawlings (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**, p. e143

75. Schmidt, T.G. & A. Skerra (2007) 'The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins' *Nat Protoc* **2**(6), p. 1528-35

76. Sharp, A.H., F.C. Nucifora, Jr., O. Blondel, C.A. Sheppard, C. Zhang, S.H. Snyder, J.T. Russell, D.K. Ryugo & C.A. Ross (1999) 'Differential cellular expression of isoforms of inositol 1,4,5-triphosphate receptors in neurons and glia in brain' *J Comp Neurol* **406**(2), p. 207-20

77. Das, B.K., L. Xia, L. Palandjian, O. Gozani, Y. Chyung & R. Reed (1999) 'Characterization of a protein complex containing spliceosomal proteins SAPs 49, 130, 145, and 155' *Mol Cell Biol* **19**(10), p. 6796-802

78. Burton, D.R. (2002) 'Antibodies, viruses and vaccines' *Nat Rev Immunol* **2**(9), p. 706-13

79. Killian, J.A. & G. von Heijne (2000) 'How proteins adapt to a membrane-water interface' *Trends Biochem Sci* **25**(9), p. 429-34

80. de Planque, M.R., J.A. Kruijtzer, R.M. Liskamp, D. Marsh, D.V. Greathouse, R.E. Koeppe, 2nd, B. de Kruijff & J.A. Killian (1999) 'Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane alpha-helical peptides' *Journal of Biological Chemistry* **274**(30), p. 20839-46

81. Yau, W.M., W.C. Wimley, K. Gawrisch & S.H. White (1998) 'The preference of tryptophan for membrane interfaces' *Biochemistry* **37**(42), p. 14713-8

82. Palsdottir, H., C.G. Lojero, B.L. Trumpower & C. Hunte (2003) 'Structure of the yeast cytochrome bc1 complex with a hydroxyquinone anion Qo site inhibitor bound' *J Biol Chem* **278**(33), p. 31303-11

83. Chakrabarti, P. (1993) 'Anion binding sites in protein structures' *J Mol Biol* **234**(2), p. 463-82

84. Copley, R.R. & G.J. Barton (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**(4), p. 321-9

85. Duncan, R.S., S.Y. Hwang & P. Koulen (2007) 'Differential inositol 1,4,5-trisphosphate receptor signaling in a neuronal cell line' *Int J Biochem Cell Biol* **39**(10), p. 1852-62

86. Corti, D., J. Voss, S.J. Gamblin, G. Codoni, A. Macagno, D. Jarrossay, S.G. Vachieri, D. Pinna, A. Minola, F. Vanzetta, C. Silacci, B.M. Fernandez-Rodriguez, G. Agatic, S. Bianchi, I. Giacchetto-Sasselli, L. Calder, F. Sallusto, P. Collins, L.F. Haire, N. Temperton, J.P. Langedijk, J.J. Skehel & A. Lanzavecchia (2011) 'A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins' *Science* **333**(6044), p. 850-6

87. Dreyfus, C., D.C. Ekiert & I.A. Wilson (2013) 'Structure of a classical broadly neutralizing stem antibody in complex with a pandemic h2 influenza virus hemagglutinin' *J Virol* **87**(12), p. 7149-54

88. Dreyfus, C., N.S. Laursen, T. Kwaks, D. Zuijdgeest, R. Khayat, D.C. Ekiert, J.H. Lee, Z. Metlagel, M.V. Bujny, M. Jongeneelen, R. van der Vlugt, M. Lamrani, H.J. Korse, E. Geelen, O. Sahin, M. Sieuwerts, J.P. Brakenhoff, R. Vogels, O.T. Li, L.L. Poon, M. Peiris, W. Koudstaal, A.B. Ward, I.A. Wilson, J. Goudsmit & R.H. Friesen (2012) 'Highly conserved protective epitopes on influenza B viruses' *Science* **337**(6100), p. 1343-8

89. Ekiert, D.C., G. Bhabha, M.A. Elsliger, R.H. Friesen, M. Jongeneelen, M. Throsby, J. Goudsmit & I.A. Wilson (2009) 'Antibody recognition of a highly conserved influenza virus epitope' *Science* **324**(5924), p. 246-51

90. Sui, J., W.C. Hwang, S. Perez, G. Wei, D. Aird, L.M. Chen, E. Santelli, B. Stec, G. Cadwell, M. Ali, H. Wan, A. Murakami, A. Yammanuru, T. Han, N.J. Cox, L.A. Bankston, R.O. Donis, R.C. Liddington & W.A. Marasco (2009) 'Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses' *Nat Struct Mol Biol* **16**(3), p. 265-73

91. Guenaga, J. & R.T. Wyatt (2012) 'Structure-guided alterations of the gp41-directed HIV-1 broadly neutralizing antibody 2F5 reveal new properties regarding its neutralizing function' *PLoS Pathog* **8**(7), p. e1002806

92. Correia, B.E., Y.E. Ban, M.A. Holmes, H. Xu, K. Ellingson, Z. Kraft, C. Carrico, E. Boni, D.N. Sather, C. Zenobia, K.Y. Burke, T. Bradley-Hewitt, J.F. Bruhn-Johannsen, O. Kalyuzhniy, D. Baker, R.K. Strong, L. Stamatatos & W.R. Schief (2010) 'Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope' *Structure* **18**(9), p. 1116-26

93. Gibson, D.G., L. Young, R.Y. Chuang, J.C. Venter, C.A. Hutchison, 3rd & H.O. Smith (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases' *Nat Methods* **6**(5), p. 343-5

94. Gibson, D.G., H.O. Smith, C.A. Hutchison, 3rd, J.C. Venter & C. Merryman (2010) 'Chemical synthesis of the mouse mitochondrial genome' *Nat Methods* **7**(11), p. 901-3

95. Otwinowski, Z. & W. Minor, *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Meth. Enzymol.*, C.W.C. Jr. and R.M. Sweet, Editors. 1997, Academic Press: NY. p. 307-26.

96. McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni & R.J. Read (2007) 'Phaser crystallographic software' *Journal of Applied Crystallography* **40**, p. 658-674

97. Potterton, E., P. Briggs, M. Turkenburg & E. Dodson (2003) 'A graphical user interface to the CCP4 program suite' *Acta Crystallogr D Biol Crystallogr* **59**(Pt 7), p. 1131-7

98. Emsley, P. & K. Cowtan (2004) 'Coot: model-building tools for molecular graphics' *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1), p. 2126-32

99. Murshudov, G.N., A.A. Vagin & E.J. Dodson (1997) 'Refinement of macromolecular structures by the maximum-likelihood method' *Acta Crystallogr D Biol Crystallogr* **53**(Pt 3), p. 240-55

100. Painter, J. & E.A. Merritt (2006) 'TLSMD web server for the generation of multi-group TLS models' *Journal of Applied Crystallpgraphy* **39**, p. 109-111

101. Painter, J. & E.A. Merritt (2006) 'Optimal description of a protein structure in terms of multiple groups undergoing TLS motion' *Acta Crystallogr D Biol Crystallogr* **62**(Pt 4), p. 439-50

102. Davis, I.W., A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, L.W. Murray, W.B. Arendall, 3rd, J. Snoeyink, J.S. Richardson & D.C. Richardson (2007) 'MolProbity:

all-atom contacts and structure validation for proteins and nucleic acids' *Nucleic Acids Res* **35**(Web Server issue), p. W375-83

103. Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov & P.E. Bourne (2000) 'The Protein Data Bank' *Nucleic Acids Research* **28**(1), p. 235-42

104. Azoitei, M.L., B.E. Correia, Y.E. Ban, C. Carrico, O. Kalyuzhniy, L. Chen, A. Schroeter, P.S. Huang, J.S. McLellan, P.D. Kwong, D. Baker, R.K. Strong & W.R. Schief (2011) 'Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold' *Science* **334**(6054), p. 373-6

105. Larman, H.B., U. Laserson, L. Querol, K. Verhaeghen, N.L. Solimini, G.J. Xu, P.L. Klarenbeek, G.M. Church, D.A. Hafler, R.M. Plenge, P.A. Nigrovic, P.L. De Jager, I. Weets, G.A. Martens, K.C. O'Connor & S.J. Elledge (2013) 'PhIP-Seq characterization of autoantibodies from patients with multiple sclerosis, type 1 diabetes and rheumatoid arthritis' *J Autoimmun*

106. Vijay-Kumar, S., C.E. Bugg & W.J. Cook (1987) 'Structure of ubiquitin refined at 1.8 A resolution' *J Mol Biol* **194**(3), p. 531-44

107. Fraczkiewicz, R. & W. Braun (1998) 'Exact and Efficient Analytical Calculation of the Accessible Surface Areas and Their Gradients for Macromolecules' *J. Comp. Chem.* **19**, p. 319-333

108. Bryson, S., J.P. Julien, R.C. Hynes & E.F. Pai (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**(22), p. 11862-75

109. DeLano, W.L., *The PyMOL Molecular Graphics System*. 2002, DeLano Scientific: San Carlos, CA, USA.

110. Saphire, E.O., P.W. Parren, R. Pantophlet, M.B. Zwick, G.M. Morris, P.M. Rudd, R.A. Dwek, R.L. Stanfield, D.R. Burton & I.A. Wilson (2001) 'Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design' *Science* **293**(5532), p. 1155-9

111. Schneider, T.D. & R.M. Stephens (1990) 'Sequence logos: a new way to display consensus sequences' *Nucleic Acids Res* **18**(20), p. 6097-100

112. Wootton, J.C. (1994) 'Non-globular domains in protein sequences: automated segmentation using complexity measures' *Comput Chem* **18**(3), p. 269-85

113. Pettit, F.K., E. Bare, A. Tsai & J.U. Bowie (2007) 'HotPatch: a statistical approach to finding biologically relevant features on protein surfaces' *J Mol Biol* **369**(3), p. 863-79

114. Swairjo, M.A., N.O. Concha, M.A. Kaetzel, J.R. Dedman & B.A. Seaton (1995) 'Ca(2+)-bridging mechanism and phospholipid head group recognition in the membrane-binding protein annexin V' *Nat Struct Biol* **2**(11), p. 968-74

Chapter 4: Ontogeny of recognition specificity and functionality for the broadly neutralizing anti-HIV antibody 4E10

An effective HIV vaccine will likely need to elicit broadly-neutralizing antibodies (bnAbs) that target the viral envelope protein (Env) as part of a protective immune response (reviewed in[1-6]). Env-derived and reverse-engineered immunogen-based vaccines have consistently failed to elicit bnAbs, however. Possible explanations include that: (1) immunogens may be unable to bind germline-encoded precursors (GEPs) of bnAbs with sufficient affinity to initiate B cell activation and affinity maturation, which has an ~micromolar threshold value, affected also by kinetics and avidity[7-9]; (2) V_H and V_L genes compatible with the development of anti-HIV bnAbs may not be common in vaccinee GEP repertoires (either in humans or animal models); (3) some bnAbs have been demonstrated to be autoreactive, which hinders their elicitation due to deletion through tolerance mechanisms; (4) the unusual characteristics inherent to anti-HIV bnAbs, such as long complementarity determining regions (CDRs), functionally-required polyspecificity, and a high degree of somatic mutation (typically observed in Abs elicited in response to chronic infections, including anti-HIV bnAbs), may not be easily achieved through conventional vaccination strategies; (5) imperfect immunogens may elicit off-target (non-neutralizing or non-Env) or humoral responses with limited breadth; and, finally, (6) neutralization mechanisms may involve complexities beyond simply binding a particular epitope on Env, which may be difficult to recapitulate, since selection during Ab maturation is based only on binding properties, not higher-order functionalities[10].

The focus of our studies is the bnAb 4E10 with a conserved, linear epitope (consensus clade B sequence: ^{671}NWFDITNWLW^{680}; core epitope: NWF$^{D}/_{N}$IT) immediately adjacent to the viral membrane in the Env gp41 subunit membrane proximal external region (MPER)[11,12]. While 4E10 displays admirable breadth[13], has been the target of successful reverse-engineered immunogen design[14], has recognizable GEPs present at finite frequencies in human naïve repertoires[15], and arguably does not require a high degree of polyspecificity to neutralize HIV[16,17], its viability as a vaccine target is hampered by limited potency, demonstrated autoreactivity and exceptional combining site flexibility[17-19]. The neutralization mechanism of 4E10 also has not been clearly defined and may involve higher order effects[17,20]. The ontogeny of 4E10 needs to be elucidated in order to understand how these properties were acquired and to what degree they impose constraints that might hinder re-elicitation.

In general, mutations acquired during Ab maturation occur predominantly in the CDRs, which make up the six loops (CDR1, 2, and 3 on the heavy chain (HCDRs) and CDR1, 2, and 3 on the light chain (LCDRs)) comprising the Ab combining site[21,22]. Therefore, the CDRs are responsible for the majority of direct contacts to an antigen, as opposed to the intervening framework regions (FWRs), which form the immunoglobulin β-sheet structure stabilizing the combining site, helping define CDR loop conformations. While CDR mutations are typically thought to more directly affect antigen binding and neutralization, anti-HIV bnAbs consistently depend on FWR substitutions to a surprising degree, though 4E10 is an exception to this exception[23]. Anti-HIV bnAbs are notorious for their high degree of somatic hypermutation, the product of a long process of affinity maturation against a rapidly mutating virus during a persistent, chronic infection[24]. While typical affinity-matured Abs have acquired 15 to 20 V_H mutations, anti-HIV bnAbs accumulate up to 100 V_H mutations[4]. These mutations are crucial because reversion to germline sequences drastically reduces epitope affinity, neutralization potency and breadth. In many cases, bnAb GEPs are unable to bind Env, though the actual eliciting isolate may not be

known[25-30]. In addition, anti-HIV bnAbs can contain extraordinarily long HCDR3s, up to 33 residues long versus an average of 13 for non-HIV bnAbs[31,32], which are most likely the product of junctional insertions during initial VDJ gene segment recombination. Using phylogenetically-inferred GEP sequences (**Fig. 1**)[15,33], 4E10 has acquired between 33 and 35 mutations during maturation, 20 to 22 in V_H, depending on initial gene segment selection, and 13 in V_L, and has an HCDR3 22 residues long, values at the less exceptional end of the anti-HIV bnAb spectrum and not unheard of for conventional Abs.

Affinity-matured Abs typically display univalent equilibrium binding constants (K_D) ranging from 10^{-6} to 10^{-10} M [34] that are orders-of-magnitude stronger than their GEPs. Multiple approaches, including computational analyses and structural and biophysical comparisons of affinity-matured Abs and their associated GEPs (*e.g.* [35-46]), have generated a consensus model for the molecular mechanism of affinity maturation, perhaps better understood as binding optimization, that traces its roots back to Pauling[47]. In the consensus model, the antigen specificities of the naïve, germline-encoded repertoire, while diverse and extensive, are further extended by encoding a high degree of polyspecificity into GEPs. This is accomplished partly through increased combining site plasticity, more formally stated as the ability of the CDRs of GEPs to dynamically sample a broader ensemble of structural conformers. In response to immunogen stimulation, Ab binding properties are iteratively optimized through cycles of somatic hypermutation and selection, resulting in improved binding affinities, kinetics and thermodynamics. While mutations have been observed to improve or add direct contacts to antigen, typically improving enthalpies of interaction and off-rates (k_d), a majority of measurably favorable mutations do not directly contact antigen. These mutations indirectly optimize binding by: (1) increasing shape complementarity between paratopes and epitopes through more global or indirect effects on structure; (2) increasing antibody stability, typically measured as solution thermostabilities (T_m), thus compensating for deleterious affects of other mutations that improve affinity; and (3)

structurally rigidifying the combining site conformer optimal for binding antigen from the accessible ensemble, typically improving entropies of interaction and on-rates (k_a). For antibodies that possibly leverage combining site plasticity as part of an instructive encounter or "induced fit" model, improvements in affinity can be more evenly distributed between k_d and k_a, though the consensus of studies of binding proteins and enzymes suggests that conformer selection is the preeminent recognition mechanism[48].

Surprisingly, comparisons of the bound and unbound structures of 4E10 revealed that this affinity-matured anti-HIV bnAb incorporates considerable HCDR conformational flexibility[17], in excess of what has been observed in most other antibodies, mature or GEP, including deviations from canonical CDR structure [49], suggesting that the ontogeny of 4E10 may be an exception to the consensus model and may pose unique challenges as a vaccine target. In order to fully understand the ontogeny of this unique bnAb and consequences for vaccine development, we characterized the unbound and complex crystal structures, and the biophysical, functional and binding properties, of 4E10 and an ensemble of its most likely GEPs. GEPs showed detectable, but extremely weak, binding to soluble Env gp140s and extremely limited neutralization potency, though some reverse engineered epitope-scaffolds (ESs) showed robust GEP affinities, well above the B cell activation threshold. 4E10 and GEP paratopes displayed a remarkable degree of structural conservation in the antigen-bound state, with little improvement in overall shape complementarity. Multi-log improvements in affinity for engineered immunogens were the result of improved off-rates or combined improvements in on- and off-rates, due largely to a small number of enhanced contacts to antigen. FWR mutations had little discernable affect on global or local structure. In violation of the consensus model of ontogeny, 4E10 thermostability was significantly worse than its GEPs; while 4E10 and GEPs displayed similarly constrained V_H/V_L interdomain movements upon binding, 4E10 maturation involved negligible combining site rigidification, with both 4E10 and GEPs sampling extensive HCDR

conformer ensembles. The narrowing of polyspecificity assumed to concur with maturation was not observed with 4E10, as both 4E10 and its GEPs showed similar patterns of limited polyspecificity to phage-displayed human peptidomes (PhIP-Seq)[50]. While 4E10 is demonstrably autoreactive, GEPs exhibited a distinct profile of autoantigen recognition by PhIP-Seq, suggesting that autoreactivity was acquired during ontogeny. In many respects, 4E10 provides a divergent example of Ab ontogeny, broadening the known range of affinity maturation pathways and challenging the generality of the existing paradigm. Retained combining site flexibility, and discrepancies in GEP binding of engineered versus Env-derived antigens, suggest that higher order mechanisms of neutralization are in play and that conventional vaccination protocols are unlikely to generate 4E10-equivalent Abs.

Results

Prediction of an ensemble of likeliest GEP sequences

In the absence of access to the original donor, computational identification of a single GEP sequence with high confidence for many bNAbs, including 4E10, is complicated by extensive editing, TdT N-nucleotide insertion during rearrangement and imprecise RNA splicing [25], leading to our decision to study an ensemble of the 12 likeliest candidates. The 4E10 heavy and light chain nucleotide sequences were analyzed using a suite of web-based programs to compositely identify segments with the fewest nucleotide mismatches (**Fig. 1A**). Each derived GEP candidate consists of a single V_L domain sequence assigned with high confidence (IGKV3-20*01 plus IGKJ1*01), one V_H gene segment also confidently assigned (IGHV1-69*06), plus one of six likely D segments (IGHD1-1*01, IGHD6-19*01, IGHD7-27*01, IGHD6-25*01, IGHD1-26*01, IGHD4-17*01) and either of two likely J segments (IGHJ4*02, IGHJ1*01). The CDR3 junctions and presumed N-nucleotide insertions in the mature Ab were retained in the candidate GEP Abs. Due to the prediction that the sequence differences introduced by alternate heavy chain J gene segments may not affect any discernable property of the GEPs, based on comparative modeling, the initial ensemble was limited to eight GEPs (IGHJ4*02 paired with all six D segments plus IGHJ1*01 paired with IGHD1-1*01 and IGHD6-19*01), with the intention of generating additional GEPs if the IGHJ4*02/IGHJ1*01 substitution exhibited any differences in structure or binding properties on the IGHD1-1*01 or IGHD6-19*01 backgrounds. These eight GEPs also recapitulate some of the variability seen in potential 4E10 GEPs identified by deep sequencing of uninfected individuals (**Fig. 1B**) [15], providing an additional justification for the ensemble approach. However, GEPs in our ensemble differ from one another by no more than four mutations, though the mutations are quite non-conservative.

GEP protein production and validation

GEPs were engineered as cleavable, single-chain, variable domain cassette (Fv; V_H + V_L) constructs to ease expression, analysis, crystallizability and prevent monobody-diabody interchange, following protocols developed for 4E10[16]. The prior study confirmed that these Fv constructs retained the structural and binding properties of Fab fragments of intact 4E10, validating this approach. All eight GEP Fvs expressed at high levels as bacterial inclusion bodies and, in all but one case (GEP 5), were refolded *in vitro* with what in our experience were exceptionally high efficiencies (20 to 40% final yields). GEP 5 was not included in subsequent experiments because the extremely poor *in vitro* refolding results suggested that this was not a viable Ab sequence. GEP constructs were stable and monodisperse in solution, running exclusively as monomers by size exclusion chromatography (SEC). Reduced/non-reduced PAGE analysis of GEPs confirmed purity and proper disulfide bond formation. GEP T_m values (**Fig. 2**), determined by circular dichroism spectroscopy as in[16], narrowly ranged from 64.2 ℃ to 67.0 ℃, showing that GEP Fvs were well folded, but had even higher T_m values than 4E10 Fv (52.8 ℃).

GEP neutralization potencies were dramatically reduced

The neutralization potency of GEPs was tested against clade A (Q461.d1, Q461.e2[51]) and B (SF162[52]) HIV-1 isolates using standard TZM-bl assays (**Fig. 3**)[53]. GEPs displayed only very weak neutralization potencies, with a trend of greater effect against the clade A isolates, particularly the neutralization-sensitive strain Q461.d1, and with GEP 1 and GEP 7 showing marginally better potencies across tested isolates. However, GEP Fv potencies were markedly reduced in comparison with 4E10 Fv.

GEPs bound Env proteins and engineered antigens, but with reduced affinities

In order to characterize the change in binding properties during 4E10 ontogeny, the binding of 4E10 and GEP Fvs to soluble Env gp140 trimers ($gp140_3$) from three HIV-1 isolates (SF162, $SF162^{K160N}$ [54] and Q461.e2) and four engineered 4E10 ES immunogens (T72, T93, T117 and T344[14,55]) was evaluated by surface plasmon resonance (SPR) interaction analysis on Biacore instrumentation. Isolates and ESs were selected to span a range of binding properties, where previous studies had shown 4E10 bound SF162 $gp140_3$ with a K_D of 135 nM[16] and ESs with K_Ds of either ~10 picomolar (T117) or ~100 picomolar (T72, T93 and T344)[14,55]. All seven GEPs showed unquantifiably weak, but detectable binding to chip-coupled clade A (Q461.e2) and clade B (SF162, $SF162^{K160N}$) $gp140_3$ (**Fig. 4**) and T72, T93, and T344 (*data not shown*) in qualitative SPR analyses, with ballparked K_Ds all estimated to be well above the ~micromolar B cell activation threshold. Quantitative SPR analyses of GEPs binding to T117 showed affinities ranging from the low nanomolar to low micromolar range (**Fig. 5**, **Table 1**). Five GEPs (1, 2, 3, 7 and 8) bind T117 with nearly identical behavior, including GEP 1 and GEP 7, which differ only by alternate J segment utilization, showing that the two incorporated mutations did not affect binding, so no further IGHJ1*01 GEPs were produced for study. GEP 4 and GEP 6, which both incorporate differences from G96H in 4E10 (A or V), showed approximately 10-fold (GEP 4) or 100-fold (GEP 6) reductions in affinity relative to the cluster of other GEPs. Kinetically, the five clustering GEPs (1, 2, 3, 7 and 8) showed affinity reductions for T117 relative to 4E10 overwhelmingly through increased off-rates (k_ds). GEPs 4 and 6, in addition to a comparable increase in k_d, also show progressive reductions in on-rates (k_as).

4E10- and GEP-ES complex structures show binding site conservation

In order to shed light on the structural differences accounting for reduced GEP binding affinities, crystal structures of GEP 1, 2 and 7 in complex with T117 were determined at resolution values

of 2.9 Å, 1.8 Å, and 3.1 Å respectively, rebuilt and refined with good statistics (**Table 2**), and compared to two reference 4E10/antigen complex structures: 4E10 bound to an epitope peptide (2FX7.pdb[12]) or a related ES, T88 (3LH2.pdb[14]). Superpositions showed that almost all direct contacts to the core epitope (NWFDIT) are conserved between 4E10 and GEP complexes to a remarkable degree (**Fig. 6**, **Tables 3**). Only six of the 35 predicted somatic mutations affect sequence positions making direct contacts to the core epitope: Y/K32L, S/Q93L, T/S31H, I/V51H, F/L54H and T/I56H. However, residues at two of these positions (93L, 51H) contacted the epitope solely through main-chain interactions, which were not affected by the side-chain substitutions. Replacements at three other positions (31H, 54H, 56H) nearly perfectly recapitulated contacts, conserving the interface. The 4E10 and GEP paratope/epitope interface is largely hydrophobic, thus binding is mainly mediated by Van der Waals contacts and entropy of desolvation, which is conserved through equivalent positioning of nonpolar groups. This was reflected in the close concordance between the surface areas buried in the complexes (SA) by core epitope and the corresponding shape complementarity (Sc[56]) between Ab and antigen: 4E10 Fab/epitope peptide: SA = 355 $Å^2$, Sc = 0.78; 4E10/T88: SA = 367 $Å^2$, Sc = 0.77; GEP 1/T117: SA = 342 $Å^2$, Sc = 0.69; GEP 2/T117: SA = 305 $Å^2$, Sc = 0.74; GEP 7/T117: SA = 343 $Å^2$, Sc = 0.66. Only two somatic mutations were predicted to contribute to a stronger binding interaction, possibly accounting for two logs of affinity gain. The Y/K32L mutation replaces a cramped hydrogen bond involving the tyrosine hydroxyl with a good salt bridge. The P/L95H mutation, involving non-contacting residues, restructured LCDR3 to reposition the side-chain of the conserved serine at 94L from a non-contacting position in GEPs to one contributing a hydrogen bond in 4E10 complex structures.

GEP structures revealed that structural plasticity was retained during affinity maturation

In order to determine whether the 4E10 combining site rigidified during affinity maturation, undergoing binding site preconfiguration, the crystal structures of GEP 7 and GEP 1 were determined in the unbound state at resolution values of 1.9 Å and 1.7 Å respectively, rebuilt and refined with good statistics (**Table 2**), and compared to the unbound structure of 4E10[17]. Superpositions of 4E10 and GEP bound and unbound structures as Fv units or separated V_H and V_L domains (**Fig. 7**) showed that interdomain movements were limited compared to CDR rearrangements. Comparisons of V_H and V_L structures in bound vs. unbound structures (**Fig. 8**) showed that 4E10 retained at least as much CDR flexibility as was present in GEPs, particularly in HCDR2 and HCDR3, with LCDRs relatively constrained. Calculated root mean square deviations (RMSDs) of V_H and V_L superpositions, with and without CDRs, confirmed that while the FWR regions of GEPs and 4E10 were nearly identical, the bulk of rearrangements observed during binding occurred in HCDRs, with 4E10 movements as large, or larger, than observed in GEPs (**Fig. 9**). LCDR movements, while much smaller in degree overall, contributed less to rearrangements in 4E10 than GEPs, suggesting some minimal degree of rigidification during maturation. These analyses need to be interpreted cognizant of the constraints imposed by crystallization, which involved varying non-physiological solution conditions and variable crystal contacts between structures. However, these caveats are reduced by comparing multiple structures with multiple molecules per asymmetric unit (AU).

Engineered ESs can make extensive contacts outside of the targeted epitope

The goal of reverse engineering an Ab is to scaffold the desired epitope to re-elicit Abs that solely recognize the epitope[6,57]. However, the 4E10 linear epitope, as currently defined, is smaller than typical Ab/antigen interfaces, which poses the design challenge of isolating humoral responses to the epitope and not contiguous scaffold surfaces. Previous crystal structures of 4E10/ES complexes showed that many ESs achieved this goal well[14,55]. However,

GEP/T117 complex structures reported here showed extensive GEP/scaffold contacts (**Fig. 10, Table 4**). A dominant feature of these interactions was the binding of the side-chain of the GEP-specific residue F54H in a deep hydrophobic cleft of the T117 scaffold protein, a putative phosphotransferase from *S. typhimurium*. These additional contacts raise the concern that Abs elicited by T117 immunizations may have off-target (non-Env) specificities. Nevertheless, the T117 scaffold is highly complementary to 4E10 and the GEPs, which may help explain the increased affinity of T117 for 4E10 and GEPs, and may be ideal for preferentially targeting 4E10 GEPs through the F54H interaction (F54H is present in the heavy chain gene used by all GEPs). The additional T117 contacts appear to have had the effect of increasing the affinity for T117 over other ESs by two orders of magnitude for 4E10 and by one to three orders of magnitude for the GEPs.

GEPs displayed a similar degree of polyspecificity to 4E10, but different autoreactivity

A validated phage-displayed library consisting of 413,611 overlapping 36-mer peptides spanning the entire human proteome combined with phage immunoprecipitation sequencing (PhIP-Seq) was used to assess the polyspecificity and autoantigen recognition profiles of GEPs in comparison to 4E10 (**Fig. 11A**, **Table 5**)[17,50]. GEP 2 and GEP 4 were selected to represent both the clustering (GEP 2) and 96H mutant (GEP 4) GEPs, and an affinity-matured, murine anti-canine CD28 Ab, 1C6[58], was included for comparison. 4E10, GEP and 1C6 Fvs were coupled to beads and analyzed in duplicate. Results were plotted as replicate #1 versus replicate #2 $-\text{Log}_{10}$ *P*-values[50]; highly discordant, and therefore spurious, hits falling near the axes were excluded from analysis. 241 peptides were also discarded because they displayed nonspecific binding to all Abs tested[17]. The top hit in the 4E10 Fv PhIP-Seq analysis reported here, a peptide derived from the type 2 inositol 1,4,5-trisphosphate receptor, matches the top hit from the previous PhIP-Seq analysis of IgG 4E10[17]. However, overall scores were considerably

reduced in the PhIP-Seq analysis of 4E10 as an Fv construct versus intact IgG (replicate average $-\text{Log}_{10}$ *P*-values for the top ten scoring peptides of 35.3 to 255 for IgG versus 4.35 to 12.3 for Fv), likely due to decreased accessibility of coupled Fv relative to IgG, an increased chance of inactivating Fv versus IgG during chemical coupling, and the inherent increase in local avidity of bivalent IgG versus univalent Fv on potentially sparsely-coupled beads. Given these caveats, overall scoring behavior was very similar across the Fvs tested (4E10, average = 0.32, σ = 0.35; GEP 2, average = 0.22, σ = 0.44; GEP 4, average = 0.25, σ = 0.52; 1C6, average = 0.27, σ = 0.50), with 4E10 and 1C6, the affinity-matured Abs, showing the highest averages. GEP 4 and 1C6 showed the largest number of significant peptides hits, with 61 and 194 peptides with replicate-averaged $-\text{Log}_{10}$ *P*-values ≥4.0 respectively; 4E10 and GEP2 had 12 and 20 peptides, respectively, scoring ≥4.0. Qualitatively, the results were not dramatically different, but with GEP 4 and 1C6 showing nominally greater spreads of top-scoring peptides rising above the bulk responses. None of the top-scoring three-dozen 4E10 peptides appeared in the top three-dozen hits from either GEP; however two of the top-scoring dozen GEP peptides (derived from zinc finger Ran-binding domain-containing protein 3 or hyaluronidase-3 isoform 1 precursor) were in common between GEP 2 and GEP 4, scoring 1 and 12 (GEP 2) and 4 and 2 (GEP 4), respectively. While 1C6 showed the highest scoring spread of top hits (replicate average $-\text{Log}_{10}$ *P*-values for the top ten scoring peptides of 12.14 to 30.9), the top ten scoring peptides displayed a considerably higher average hydrophobic character (Φ; as calculated with the Sigma-Aldrich PEPscreen Library Design Tool), with average Φ values of: 4E10 = 0.37; GEP 2 = 0.42; GEP 4 = 0.45; 1C6 = 0.71 (higher values are more hydrophobic). Using relative hydrophobicity of the top-scoring peptides determined by PhIP-Seq as a surrogate measure of the overall hydrophobicity of the combining site is consistent with the structures of the Abs (**Fig. 11B**), where the 1C6 combining site is structured as a large, broad, hydrophobic concavity and 4E10 and its GEPs sporting smaller, convex hydrophobic surfaces.

Discussion

Reverse engineered ESs are ideal reagents for studying the binding properties of GEPs. Where peptides and Env do not display sufficient affinities, the best ESs display strong affinities across the GEP ensemble, allowing for the biophysical characterization of 4E10 ontogeny through binding kinetics and complex crystal structures. Consistent with current theory and previous results comparing GEPs with matured anti-HIV bnAbs or other Abs, 4E10 displays orders-of-magnitude better affinities than candidate GEPs for both Env $gp140_3$ proteins and engineered ESs. This gain in affinity is potentially sufficient to account for the concurrent gain in neutralization potency over GEPs. Previous results showing minimal contributions from FWR mutations [23] and comparisons of complex structures of 4E10 with GEP/T117 complex structures showing near-complete conservation of the recognition interface (**Fig. 6**), argue that affinity maturation is likely due, at least in part, to the small number of somatic mutations, *e. g.* 32L and 95H, that add or improve direct contacts to epitope. This limited number of mutations is presumably readily achievable during conventional vaccination. Based on PhIP-Seq peptidome binding results (**Fig. 11**), 4E10 and GEPs have distinct autoantigen profiles, suggesting that 4E10 autoreactivity was acquired during ontogeny and not inherent in GEPs, consistent with recognizable GEPs populating naïve repertoires of mature B cells at relatively high frequencies[15]. It has also been previously argued that, when considering vaccination strategies, 4E10 autospecificity is not an insurmountable hurdle[17]. Among ESs, T117 in particular interacts with GEPs sufficiently strongly to drive B cell activation and maturation (**Fig. 5**) and selectively interacts with GEP-specific structural features, *e.g.* F54H (**Fig. 10**). These aspects of 4E10 combine to seemingly argue that 4E10 ontogeny follows a relatively short path and that directed re-elicitation of 4E10, and perhaps other anti-HIV bnAbs, may be an achievable goal.

In many aspects, however, 4E10 ontogeny appears more convoluted. Contrasting with the consensus model of Ab ontogeny and results from other systems (e.g.[37]), 4E10 is considerably less thermostable than GEPs, which is difficult to account for only from static views of structures. If neutralization by 4E10 requires all somatic mutations *in toto*, perhaps to enable flexibility, then there may be a significant hurdle to re-elicitation of 4E10, because it lowers the headroom available to tolerate mutations that increase affinity or potency that otherwise might degrade stability. The conformer ensemble sampled by GEP combining sites echoes the startling plasticity of 4E10, with structural comparisons of bound versus unbound states (**Figs. 7**, **8**, and **9**) showing perhaps a small degree of 4E10 LCDR rigidification, but an increase in the conformer ensemble sampled by the HCDRs. This directly contradicts the consensus model, and specific examples (e.g.[39]) of Ab maturation, showing alternately that preconfiguration of 4E10 does not occur during or contribute to affinity maturation. HCDR flexibility is highlighted by F29H, a canonical residue responsible for defining an HCDR1 canonical loop structure[49], which is able to flip out into solvent in the unbound 4E10[17] and GEP 1 structures, allowing HCDR1 to dynamically sample non-canonical conformers. A high number of conserved glycine residues may contribute to 4E10 and GEP HCDR mobility. The reduced affinities due to decreased k_as (**Fig. 5**), relative to the clustering GEPs, of GEP 4 and 6 with non-glycine residues at position 96H, predicted to restrict HCDR3 conformer sampling, suggests that HCDR3 mobility is needed to achieve the bound-state conformation with optimal kinetics by destabilizing non-optimal conformers. Retention of significant combining site plasticity also strongly argues for a functional role other than polyspecificity, which is unremarkable in comparison to other Abs based on PhIP-Seq peptidome binding results (**Fig. 11**, [17]). Little apparent change in polyspecificity also contradicts the consensus model of maturation and specific prior examples[59]. Also unexpectedly, the ~100-fold increase in affinity of 4E10 over GEPs to T117 is through a decreased k_d, suggesting that improvements in binding affinity arose from a more favorable bound state and not a decrease in entropic barriers imposed by a more flexible GEP.

Thermodynamic studies of 4E10 and GEPs were not possible because no single antigen, needed for valid comparisons, bound to both 4E10 and GEPs with parameters accessible to measurement. However, while crystal structures show a highly-conserved bound state in both 4E10 and GEPs, the small number of improved contacts could theoretically account for the observed improvement in affinity. Complicating this analysis, and its use as a vaccine immunogen, are the extensive contacts between GEPs, and presumably also 4E10, and the scaffold of T117 outside of the stabilized MPER epitope (**Fig. 10**). Unless the MPER epitope forms an isolated structure extended away from the rest of Env, the interaction with T117 also highlights the possibility that 4E10 makes similar contacts to elements of Env outside of the linear 4E10 epitope. HCDR conformer dynamics may therefore be understood as enabling such interactions, which may foster conformational changes in Env leading to neutralization by any of several mechanisms.

T117 was the product of a cutting-edge design effort to generate a structurally-stabilized MPER epitope with optimized binding properties for use as a vaccine immunogen to drive 4E10 ontogeny based on the current paradigm of affinity maturation. In many regards, that effort was highly successful and may lead to the use of ESs as the basis of an efficacious HIV vaccine. However, 4E10 appears to take the affinity maturation pathway less traveled, one that contrasts, in many fundamental ways, the current paradigm. The concern is that conventional immunization protocols, based on the current paradigm of structural stabilization of optimal binding conformers, will not successfully re-elicit 4E10 or other unconventional Abs. Understanding unconventional maturation pathways then becomes paramount for the future of molecular vaccinology.

Methods

Protein prediction, expression, purification and characterization

The 4E10 heavy and light chain nucleotide sequences were analyzed using *JoinSolver*[60], IMGT/V-QUEST[61,62], Ab-Origin[63], SoDA[64] and iHMMune[65] to compositely identify segments with the fewest nucleotide mismatches, generating a combinatorial ensemble of 12 GEPs (**Fig. 1**). 4E10, GEPs, and 1C6 Fvs were engineered, expressed as inclusion bodies in *E. coli* BL21(DE3) RIL cells (Invitrogen), refolded and purified as previously described[16]. Circular dichroism (CD) spectra of 4E10 Fv and GEPs were measured on a J-815 spectrometer (Jasco) at a concentration of 10 μM in 10 mM Na_2HPO_4/KH_2PO_4 (pH = 7.4). Temperature melts were performed at 210 nM with a temperature ramp from 25 ℃ to 95 ℃ at a slope of 2 ℃/minute, data pitch of 2 ℃, and delay time of 10 s (**Fig. 2**). T_ms were determined by nonlinear least-squares analysis using a linear extrapolation model with Spectra Analysis software (Jasco).

Neutralization Assays

Neutralization assays (**Fig. 3**) were performed using single-round entry-competent viruses and TZM-bl cells as previously described[53]. IC_{50} values and percent neutralizations at concentrations of 0.96 μM (4E10 and GEP Fvs) or 0.17 μM (4E10 IgG) and were calculated as previously described[53 16].

SPR Interaction Analyses

All SPR experiments were performed at 25 ℃ in a ru nning buffer of HBS-EP+ (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% P-20; GE Healthcare) plus 0.1 mg/ml bovine serum albumin on a Biacore T100 instrument with the T200 sensitivity enhancement (GE Healthcare). For analyses of the binding of 4E10 and GEP Fv analytes to chip-coupled ES ligands (**Fig. 5**),

ESs at 1 μg/mL in running buffer were captured on a SA sensor chip (GE Healthcare) following either carboxy (T117) or amine (T72, T93, T344) biotinylation following the manufacturer's recommended protocol (EZ-Link, Thermo Scientific) at a ligand density of approximately 30 response units (RU). For analyses of the binding of 4E10 and GEP Fv analytes to $gp140_3$ (**Fig. 4**), $gp140_3$ at 1 μg/mL in 10 mM sodium acetate (pH = 5.0) were immobilized at a density of approximately 250 RUs, in HBS-EP+ on a CM5 sensor chip (GE Healthcare) by direct amine coupling chemistry; $gp140_3$ were prepared as previously described [16]. Reference flow cells in both analyses were left blank. 4E10 and GEP Fv analytes were injected in randomized duplicate runs at a flow rate of 50 μl/minute. For kinetic measurements, GEP analytes (concentration range duplicates) were injected for 7 min followed by a 10 min dissociation phase; 4E10 was injected for 7 min followed by a 28 min dissociation phase. Regeneration was achieved by injection of 10 mM glycine (pH = 1.5) at a flow rate of 50 μl/min for 10 s followed by HBS-EP buffer stabilization for 2 min. For equilibrium measurements, analytes (concentration range duplicates) were injected for 6 min followed by a 3 min dissociation phase. No regeneration was used. For equilibrium measurements with $gp140_3$, analytes (300 nM duplicates) were injected for 2 min followed by a 3 min dissociation phase. Regeneration was achieved by injection with 10 mM glycine, pH 1.5 at a flow rate of 50 μl/min for 5 s followed by a 6 min stabilization period. Sensorgrams obtained from SPR measurements were double-reference subtracted[66] and analyzed with BIAevaluation 2.0.3 software (GE Healthcare) following previously described methodology[16].

Crystallization and Crystallography

Crystals of GEPs and 4E10 Fv were grown by the hanging-drop vapor diffusion method at 25 ℃ with the following well solutions:

GEP 1: 25% w/v PEG 3350, 0.2 M NaCl, 0.1 M Tris (pH = 8.5)

GEP 1/T117: unbuffered 12% w/v PEG 8000, 0.1 M KCl, 5% v/v glycerol

GEP 2/T117: 1.6 M Li_2SO_4, 0.1 M sodium acetate (pH = 5.0)

GEP 7: 20% w/v PEG 2000, 0.2 M Trimethylamine N-oxide, 0.1 M Tris (pH = 9.0)

GEP 7/T117: 10% w/v PEG 10000, 0.1 M NH_4OAc, 0.1 M Bis-Tris (pH = 5.5)

1C6: 3.5 M sodium acetate, 0.1 M Bis-Tris Propane (pH = 7.0)

Crystals were cryopreserved in mother liquor containing 15% v/v glycerol (GEP 1, GEP 1/T117, GEP 7/T117), mother liquor containing 10% (1C6) or 20% (GEP 7) v/v glycerol, or 20% w/w sucrose (GEP 2/T117). Diffraction data for GEP 1, and 4E10 Fv were collected at the Advanced Light Source beamline 5.0.2 (Lawrence Berkeley National Laboratory, Berkeley, CA) and reduced using HKL-2000[67] or d*TREK (Rigaku). Diffraction data for GEP 7, GEP 1/T117, GEP 7/T117 and GEP 2/T117 were collected in house with CuKα radiation on a R-AXIS IV++ image plate detector with HR optics (Rigaku) at -170 ℃. Initial phase information for all data sets was determined by molecular replacement, as implemented in the CCP4i program suite[68,69]. Phases were improved by subsequent rounds of model building and refinement using COOT[70] and REFMAC[71]. Structure validation was carried out with Procheck[72], the MolProbity server[72], and the RCSB ADIT validation server. Data collection and structure refinement statistics are shown in **Table 2**.

PhIP-Seq analysis

For each Fv analyzed, 3 mg of magnetic beads (Invitrogen M-270 Epoxy Dynabeads) were resuspended in 60 μL 0.1 M $NaPO_4$ (pH = 7.4). Beads were rocked at ambient temperature for 24 hrs with 60 μg of each Fv in 1 M $(NH_4)_2SO_4$ and then washed with 10 mM glycine in PBS to cap unreacted epoxy groups. Activity of Fv coupled beads was confirmed by epitope-scaffold binding prior to PhIP-Seq analyses. PhIP-Seq analyses were performed as previously described[17,50].

Acknowledgements

I would like to thank Roland K. Strong[1] for his assistance with data analysis and manuscript preparation; Della Friend[1] for her assistance with SPR experiments and analysis; James Jaffee[1], Meg Holmes[1], and Mesfin Gewe[1] for their contributions to crystallography data for GEP 7, GEP 2, and 1C6, respectively; H. Benjamin Larman[2] and Stephen J. Elledge[2] for their work on PhIP-Seq experiments; Andrew Stuart[3] and Leonidas Stamatatos[3,7] for their contribution on the neutralization studies; Kevin Larimore[4] and Philip D. Greenberg[4,5,6] for their work on the GEP phylogenetic analysis; and Tim Vanden Bos and Madeline LaPatra for their assistance with protein production.

[1]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

[2]Department of Genetics, Harvard University Medical School, and Division of Genetics, Howard Hughes Medical Institute, Brigham and Women's Hospital, Boston, MA 02115

[3]Seattle Biomedical Research Institute, Seattle, Washington 98109

[4]Department of Immunology, University of Washington, Seattle, WA 98195

[5]Department of Medicine, University of Washington, Seattle, WA 98195

[6]Program in Immunology, Cancer Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

[7]Department of Global Health, University of Washington, Seattle, Washington 98109

Figures

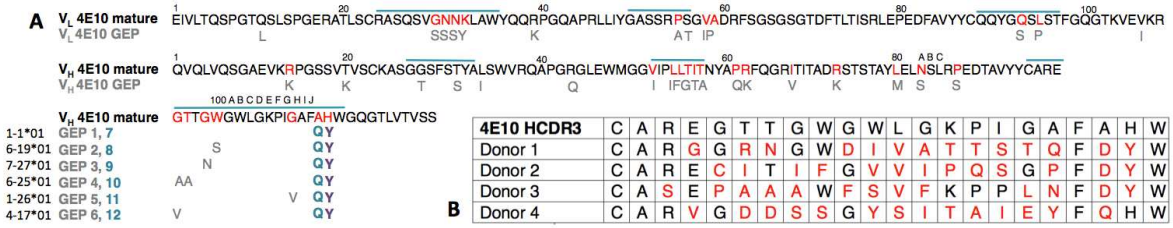

Figure 1: **(A)** VL and VH sequences of 4E10 and GEP Fvs are shown; CDRs are indicated by a blue overscore. 4E10 affinity matured residues are colored red; the corresponding GEP residues are shown below in grey. Each GEP VL domain contains the IGKV3-20*01 gene. Each GEP VH domain contains the IGHV1-69*06 V gene segment, one of six D gene segments (shown to the left of the corresponding GEP) and either the IGHJ4*02 (purple, bold font) or IGHJ1*01 (blue, bold font) J gene segment, yielding an ensemble of 12 GEPs. GEP shorthand numbering is shown in grey (GEP 1-6) and blue (GEP 7-12) beside the corresponding D plus J gene segment combination. **(B)** HCDR3s from candidate 4E10 GEPs, determined through deep sequencing of naïve B cell germline IgH genes from four uninfected individuals, shows the degree of variability seen in 4E10 GEPs [15]. Each germline rearrangement uses the VH1-69 and JH1, or JH4 gene segments. Amino acids in red designate somatic mutations between GEP and 4E10. The number of nucleotide changes needed to achieve these somatic mutations is 16 for donor 1, 17 for donor 2 and 18 for donors 3 and 4.

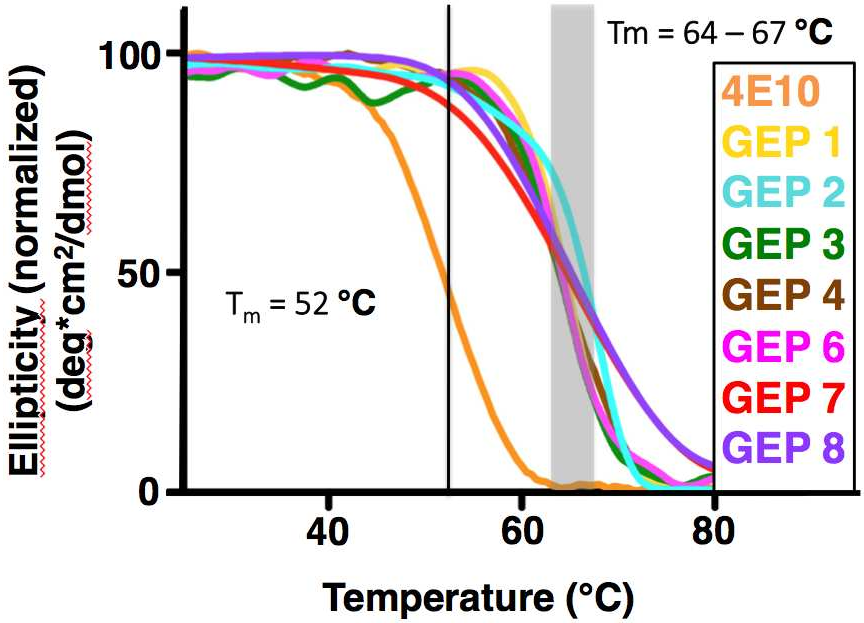

Figure 2: CD melting curves are shown for 4E10 and GEPs. Melting temperatures, given by the inflection point of the sigmoidal melting curve, are indicated by a black line for 4E10 or a shaded grey box for GEPs.

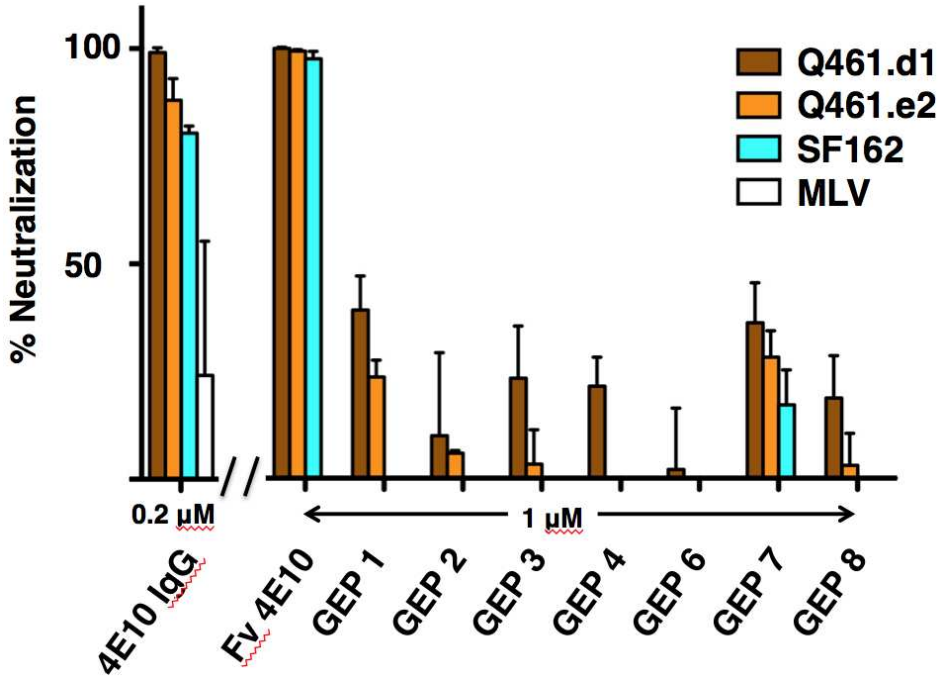

Figure 3: Neutralization potencies are shown for 4E10 IgG (0.2 uM), 4E10 Fv and GEPs (1uM) against clade A (Q461.d1, Q461.e2 [51]) and B (SF162 [52]) HIV-1 isolates using standard TZM-bl assays.

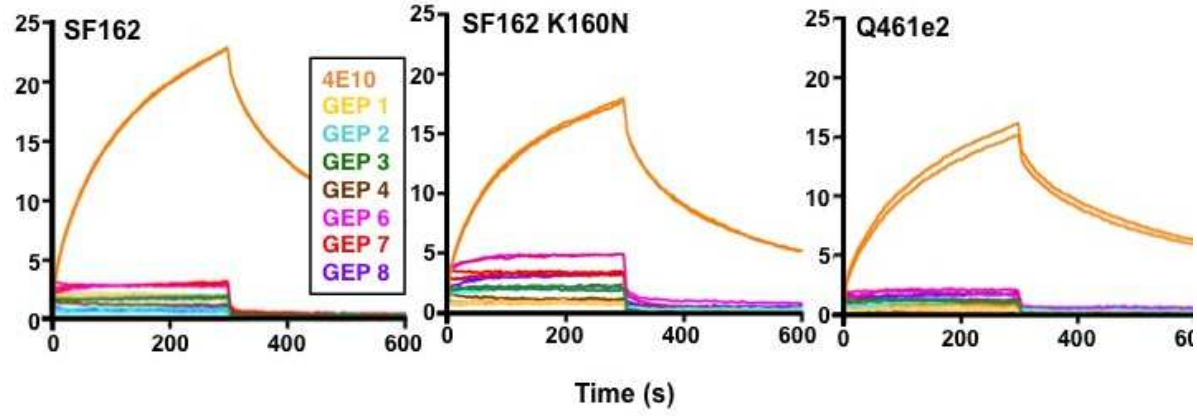

Figure 4: Double reference subtracted SPR sensorgrams are shown for the binding of 4E10 and GEPs (300 nM duplicates) interacting with chip-coupled $gp140_3$s.

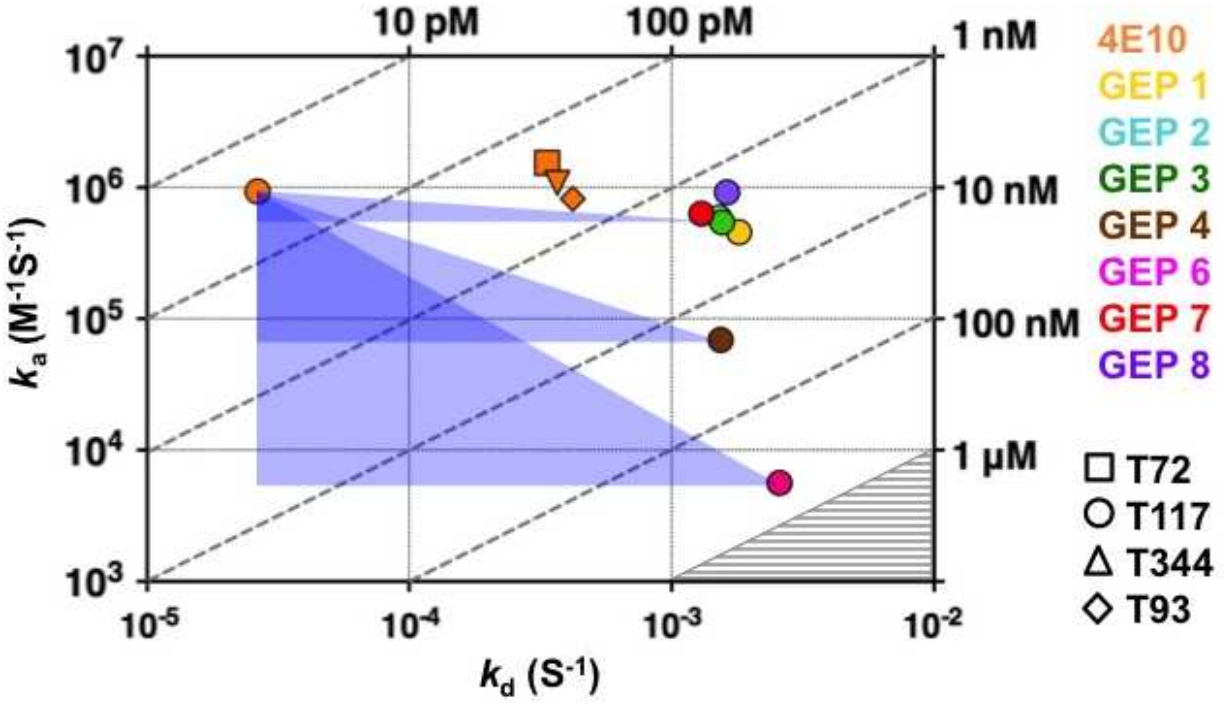

Figure 5: K_Ds for the interaction of 4E10 or GEPs with ES T117 are plotted as on-rate (k_a) vs. off-rate (k_d), with K_D isotherms indicated by dashed lines. Due to weak binding and fast kinetics, KDs between GEPs and ESs T72, T344, and T93 could only be analyzed by equilibrium measurements; their values range from 1 to 10 uM and are depicted by the grey shaded region. The purple triangles show the T117 affinity shift between 4E10 and GEPs, with the sides parallel to the X and Y axes of each triangle highlighting the on-rate and off-rate components, respectively.

Table 1: SPR analyses of 4E10 and GEP binding to ESs.

T117	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_D (M)
4E10	9.35(2) x 10^5	2.65(1) x 10^{-5}	2.83(1) x 10^{-11}
GEP 1	4.56(2) x 10^5	1.80(1) x 10^{-3}	3.96(2) x 10^{-9}
GEP 2	5.90(1) x 10^5	1.52(1) x 10^{-3}	2.57(1) x 10^{-9}
GEP 3	5.43(1) x 10^5	1.56(1) x 10^{-3}	2.88(1) x 10^{-9}
GEP 4	6.90(4) x 10^4	1.54(1) x 10^{-3}	2.23(1) x 10^{-8}
GEP 6	5.61(3) x 10^3	2.58(1) x 10^{-3}	4.60(3) x 10^{-7}
GEP 7	6.31(2) x 10^5	1.30(1) x 10^{-3}	2.06(1) x 10^{-9}
GEP 8	9.18(2) x 10^5	1.63(1) x 10^{-3}	1.78(1) x 10^{-9}

T93	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_D (M)
4E10	8.22(1) x 10^5	4.21(1) x 10^{-4}	5.13(1) x 10^{-10}
GEP 1	--	--	9.3(1) x 10^{-6}
GEP 2	--	--	5.1(7) x 10^{-6}
GEP 3	--	--	2.2(1) x 10^{-6}
GEP 4*	--	--	--
GEP 6*	--	--	--
GEP 7	--	--	5.23(6) x 10^{-6}
GEP 8	--	--	3.1(3) x 10^{-6}

T344	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_D (M)
4E10	1.08(1) x 10^6	3.67(1) x 10^{-4}	3.41(1) x 10^{-10}
GEP 1	--	--	6.4(6) x 10^{-6}
GEP 2	--	--	8(1) x 10^{-6}
GEP 3	--	--	4.8(3) x 10^{-6}
GEP 4*	--	--	--
GEP 6*	--	--	--
GEP 7	--	--	5.7(4) x 10^{-6}
GEP 8	--	--	5.2(4) x 10^{-6}

T72	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_D (M)
4E10	$1.54(1) \times 10^{6}$	$3.36(1) \times 10^{-4}$	$2.19(1) \times 10^{-10}$
GEP 1	--	--	$1.4(1) \times 10^{-6}$
GEP 2	--	--	$1.25(6) \times 10^{-6}$
GEP 3	--	--	$1.78(4) \times 10^{-6}$
GEP 4	--	--	$7.9(5) \times 10^{-6}$
GEP 6*	--	--	--
GEP 7	--	--	$1.34(5) \times 10^{-6}$
GEP 8	--	--	$1.26(6) \times 10^{-6}$

*binding too weak or not detected

Table 2: Data collection and refinement statistics

	GEP 1	GEP 1/T117	GEP 2/T117	GEP 7	GEP 7/T117	IC6
PDB accession code	4LRN	4M8Q	4M62	**	**	4LCI
Data collection*						
Space group	*P*222	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_1$	$P2_12_12_1$
Cell dimensions						
a, *b*, *c* (Å)	35.54, 48.77, 110.7	72.56, 77.67, 77.82	69.88, 103.5, 109.3	35.54, 48.95, 117.27	72.91, 78.00, 78.40	44.60, 71.62, 73.23
α, β, γ (°)	90, 90, 90	90, 112.03, 90	90, 90, 90	90, 90, 90	90, 112.29, 90	90, 90, 90
Resolution (Å)	50.00-1.90 (1.93-1.90)	24.61-2.89 (2.95-2.89)	69.88-1.80 (1.86-1.80)	50.00-1.70 (1.73-1.70)	62.73-3.10 (3.21-3.10)	44.60–1.90 (1.97–1.90)
R_{merge}	0.091 (0.434)	0.092 (0.428)	0.035 (0.149)	0.062 (0.232)	0.126 (0.421)	0.104 (0.255)
$I / \sigma I$	43.64 (6.46)	25.24 (2.82)	20.2 (5.00)	53.16 (8.63)	6.9 (1.9)	10.0 (4.1)
Completeness (%)	100.0 (100.0)	99.6 (97.8)	98.0 (91.7)	99.9 (100.0)	98.6 (97.6)	99.4 (95.0)
Redundancy	14.1 (14.3)	3.5 (3.0)	3.44 (2.37)	6.7 (6.6)	2.89 (2.85)	6.63 (4.38)
Refinement						
No. reflections (unique)	15,973 (753)	17,070 (861)	72,714 (6,712)	23,229 (1,127)	14,742 (1,437)	18024 (1,736)
R_{work} / R_{free}	0.190/0.233	0.261/0.302	0.189/0.216	0.178/0.207	0.280/0.304	0.179/0.230
No. atoms						
Protein	1,643	5,426	6,052	1,764	5,687	1854
Ligand/ion	--	--	45	23	--	25
Water	106	--	696	141	--	166
B-factors (Å^2)						
Protein	25.34	59.85	19.41	26.09	56.94	28.86
Ligand/ion	--	--	42.97	32.81	--	44.58
Water	33.90	--	27.26	35.61	--	36.47
R.m.s. deviations						
Bond lengths (Å)	0.004	0.004	0.003	0.014	0.004	0.0154
Bond angles (°)	0.765	0.733	0.708	1.636	0.865	1.584
MolProbity score	1.19	1.01	0.83	1.34	1.48	1.16
Residues in	95.6	96.4	98.3	97.79	95.8	95.71

most favored regions (%)						
Res. in disallowed regions (%)	0	0	0	0	0	0
Est. coordinate error (max. likelihood ESUc) (Å)	0.108	0.428	0.079	0.113	0.555	0.107

*One crystal was used per data set. Values in parentheses

are for highest-resolution shell.

** Awaiting deposition in protein data bank.

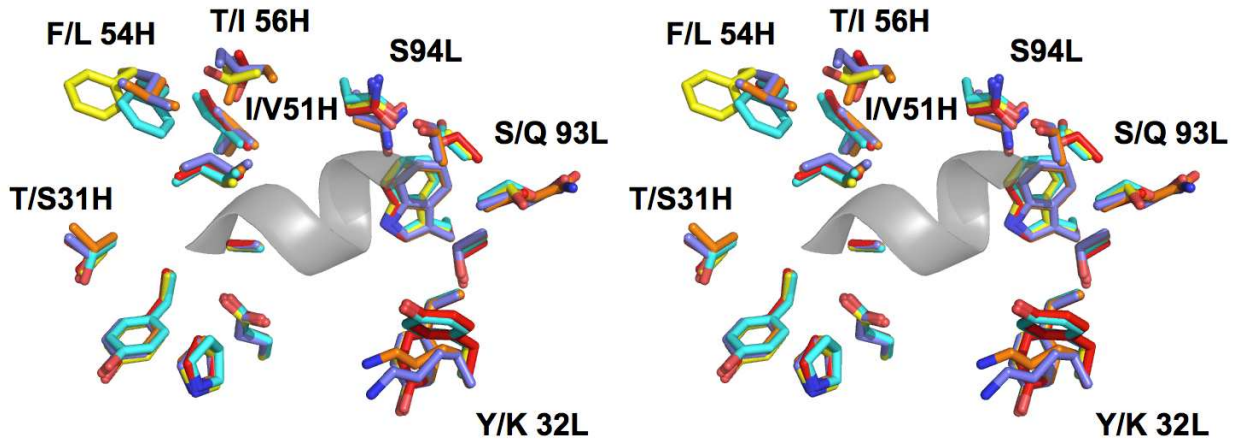

Figure 6: The combining site of 4E10 and GEPs is shown in stereo view. Superposition of 4E10 and GEP contacts to the NWFDIT epitope (cartoon representation colored grey) are shown in licorice stick representation and colored as follows: 4E10 Fv (purple), 4E10 Fab (orange), GEP 1 (yellow), GEP 2 (cyan), GEP 7 (red). The superposition consists of the following structures: 4E10 Fv in complex with ES T88 (purple); 3LH2.peb, 4E10 Fab in complex with 16mer NWFDIT peptide (orange); 2FX7.pdb, GEP 1 in complex with ES T117; 4M8Q.pdb, GEP 2 in complex with ES T117; 4M62.pdb, and GEP 7 in complex with ES T117; awaiting deposition.

Table 3: Epitope contacts for 4E10 and GEPs determined by MolProbity[72] from complex crystal structures. van der Waals contacts are shown in black, electrostatic interactions in red. Asterisks indicate residues that have been somatically mutated.

NWFDIT epitope	4E10 Fab (16mer peptide)	4E10 Fv (epitope-scaffold T88)	GEP 1 (epitope-scaffold T117)	GEP 2 (epitope-scaffold T117)	GEP 7 (epitope-scaffold T117)
Asn	-- L 91 TYR L 91 TYR L 92 GLY L 92 GLY L 93 GLN* L 94 SER L 94 SER	-- -- -- L 92 GLY L 93 GLN* L 94 SER L 94 SER	L 32 TYR L 91 TYR L 91 TYR L 92 GLY L 92 GLY L 93 SER* -- --	L 32 TYR L 91 TYR L 91 TYR L 92 GLY L 92 GLY L 93 SER* -- --	L 32 TYR L 91 TYR L 91 TYR L 92 GLY L 92 GLY L 93 SER* -- --
Trp	H 33 ALA H 50 GLY H 51 VAL* H 52 ILE H 56 ILE* H 56 ILE* -- H 58 ASN L 94 SER L 94 SER	H 33 ALA H 50 GLY H 51 VAL* H 52 ILE -- H 56 ILE* H 58 ASN H 58 ASN L 94 SER L 94 SER	H 33 ALA H 50 GLY H 51 ILE* H 52 ILE H 56 THR H 56 THR* -- H 58 ASN -- --	H 33 ALA H 50 GLY H 51 ILE* H 52 ILE -- -- -- H 58 ASN -- --	H 33 ALA H 50 GLY H 51 ILE* H 52 ILE -- H 56 THR* -- H 58 ASN -- --
Phe	H 47 TRP H 100J PHE L 91 TYR L 93 GLN* L 94 SER L 96 SER	H 47 TRP H 100J PHE L 91 TYR L 93 GLN* L 94 SER L 96 SER	H 47 TRP H 100J PHE L 91 TYR L 93 SER* L 94 SER L 96 SER	H 47 TRP H 100J PHE L 91 TYR L 93 SER* L 94 SER L 96 SER	H 47 TRP H 100J PHE L 91 TYR L 93 SER* L 94 SER L 96 SER
Asp	L 32 LYS* L 32 LYS*	L 32 LYS* L 32 LYS*	L 32 TYR* L 32 TYR*	L 32 TYR* L 32 TYR*	L 32 TYR* L 32 TYR*
Ile	H 52 ILE H 54 LEU* H 56 ILE*	H 52 ILE H 54 LEU* H 56 ILE*	H 52 ILE H 54 PHE* H 56 THR*	H 52 ILE H 54 PHE* --	H 52 ILE H 54 PHE* H 56 THR*

Thr	H 31 THR* H 32 TYR H 33 ALA H 52 ILE H 95 GLU H 95 GLU H 100F PRO	H 31 THR* H 32 TYR H 33 ALA H 52 ILE H 95 GLU H 95 GLU H 100F PRO	H 31 SER* H 32 TYR H 33 ALA H 52 ILE H 95 GLU H 95 GLU H 100F PRO	H 31 SER* H 32 TYR H 33 ALA H 52 ILE H 95 GLU H 95 GLU H 100F PRO	H 31 SER* H 32 TYR H 33 ALA H 52 ILE H 95 GLU H 95 GLU H 100F PRO

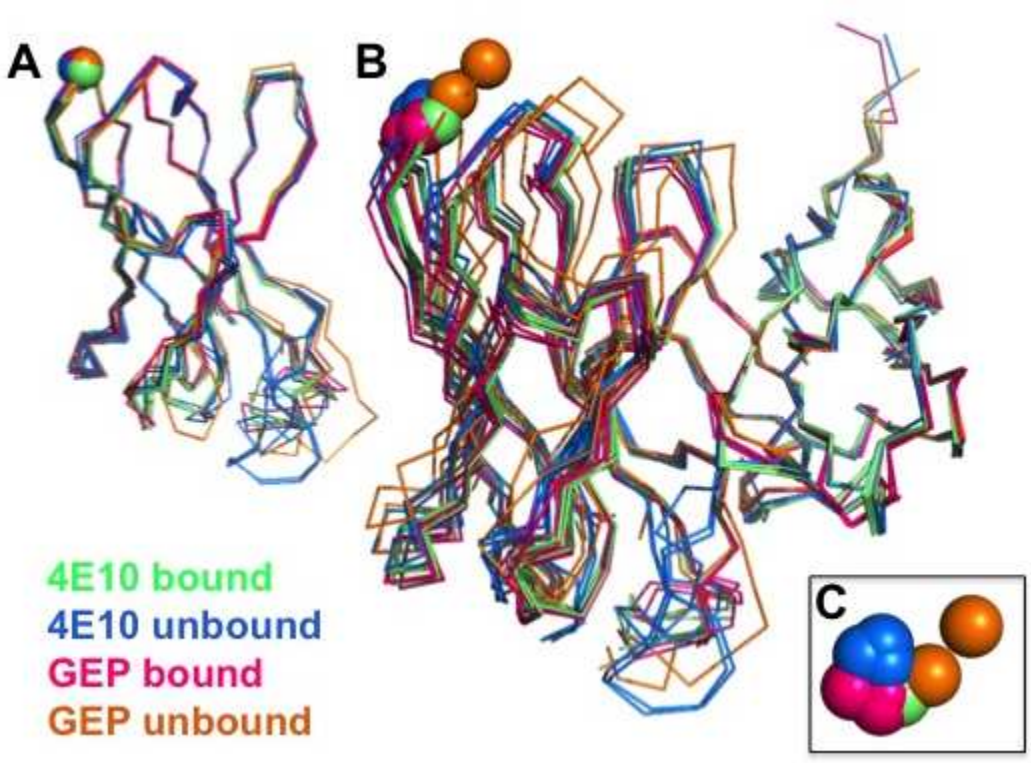

Figure 7: **(A)** Superpostion of VH from 4E10 bound (2FX7.pdb, 3LH2.pdb), unbound (4LLV.pdb), GEP 1 bound (4M8Q.pdb), GEP 1 unbound (4LRN.pdb), GEP 2 bound (4M62.pdb), GEP 7 bound (awaiting deposition), and GEP 7 unbound (awaiting deposition) is shown in ribbon representation. The Cα in residue P14H in each Fv, chosen to illustrate interdomain movement upon binding between 4E10 and GEPs, is shown as a sphere and colored as in A. **(B)** Superposition of VL from 4E10 and GEPs colored as in A. **(C)** Top view of spheres as shown in A and B illustrating interdomain movement.

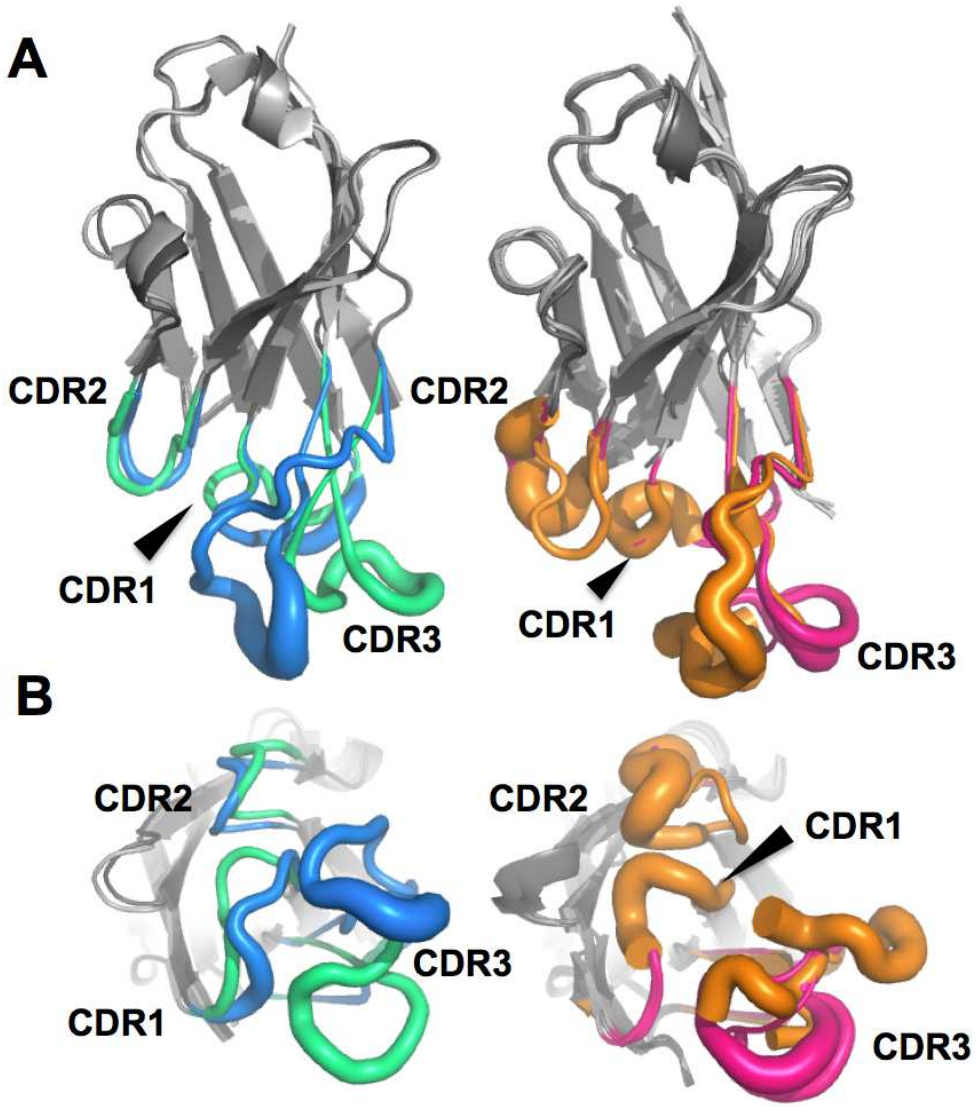

Figure 8: **(A)** Superpositions of the VH domains are shown for 4E10 bound (green/light grey; 3LH2.pdb) versus unbound (blue/ dark grey; 4LLV.pdb) and GEPs 1 (4M8Q.pdb), 2 (4M62), and 7 (awaiting deposition) bound (pink/ light grey) versus GEPs 1 (4LRN.pdb) and 7 unbound (awaiting deposition) (orange/ dark grey). CDRs are shown in a b-factor putty representation while the FWR are shown in cartoon representation. **(B)** View of HCDRs from below, shown as in A.

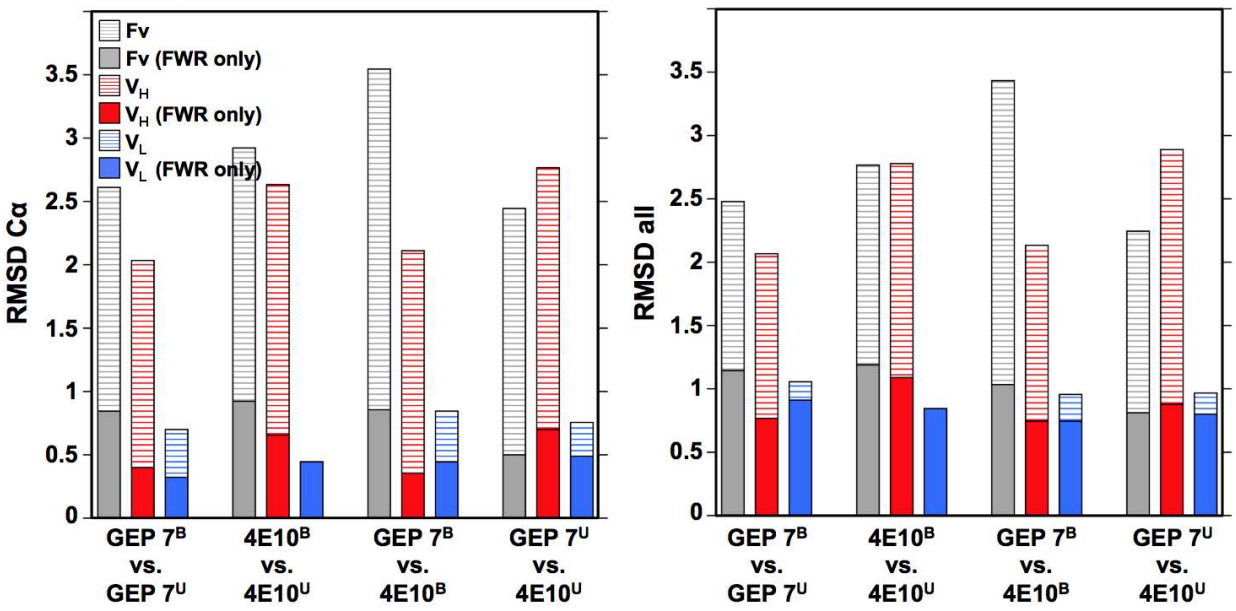

Figure 9: RMSDs for V_H, V_L, or Fv superpositions with and without CDRs are given for Cα (left) and all atoms (right).

Table 4: Scaffold contacts for GEPs determined by MolProbity[72] from complex crystal structures. van der Waals contacts are shown in black, electrostatic interactions in red.

Epitope-scaffold T117	GEP 1	GEP 2	GEP 7
8 ALA	H 100E LYS	H 100E LYS	H 100E LYS
56 ARG	H 54 PHE H 55 GLY --* --	H 54 PHE H 55 GLY H 56 THR H56 THR	H 54 PHE H 55 GLY H 56 THR H56 THR
60 PHE	H 54 PHE H 55 GLY --	H 54 PHE H 55 GLY H 56 THR	H 54 PHE H 55 GLY --
62 THR	H 54 PHE	H 54 PHE	H 54 PHE
64 LEU	H 53 ILE	H 53 ILE	H 53 ILE
73 ILE	H 54 PHE	H 54 PHE	H 54 PHE
75 HIS	H 54 PHE	H 54 PHE	H 54 PHE
100 ALA	H 53 ILE H 54 PHE A 73 LYS	H 53 ILE H 54 PHE H 73 LYS	H 53 ILE H 54 PHE H 73 LYS
101 GLY	H 73 LYS	H 73 LYS	H 73 LYS
125 ASN	H 100F PRO H 100E LYS	H 100F PRO H 100E LYS	H 100F PRO H 100E LYS
127 LEU	H 53 ILE H 54 PHE	H 53 ILE H 54 PHE	H 53 ILE H 54 PHE
128 TRP	-- -- H 100D GLY H 100E LYS H 100F PRO	H 100C LEU H 100C LEU H 100D GLY H 100E LYS H 100F PRO	-- -- H 100D GLY H 100E LYS H 100F PRO

***56 Arg from T117 is not built in the GEP 1 complex structure

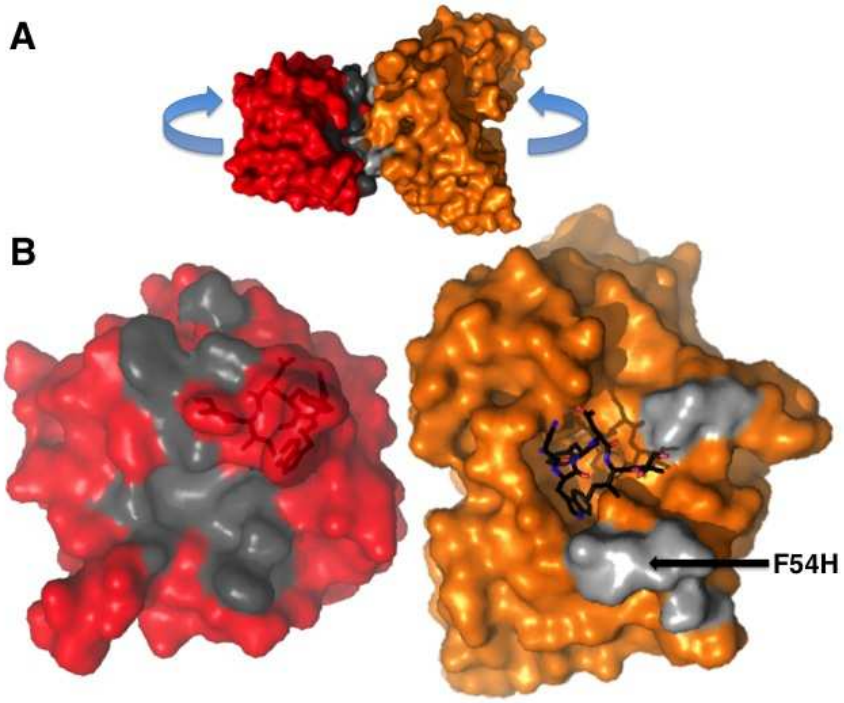

Figure 10: **(A)** The complex of GEP 7 (orange) bound to ES T117 (red) shown in surface representation. Non-epitope-scaffold contacts are shown in grey. **(B)** T117, on the left, is shown in a red semi transparent surface representation with the NWFDIT epitope portion of the scaffold indicated by a black licorice representation. GEP 7 is shown on the right, colored as in A. The NWFDIT epitope from T117 is shown bound for reference. The footprint of GEP 7 (non-epitope contacts) on T117 is colored dark grey; the footprint of T117 (non-epitope) on GEP 7 is colored light grey.

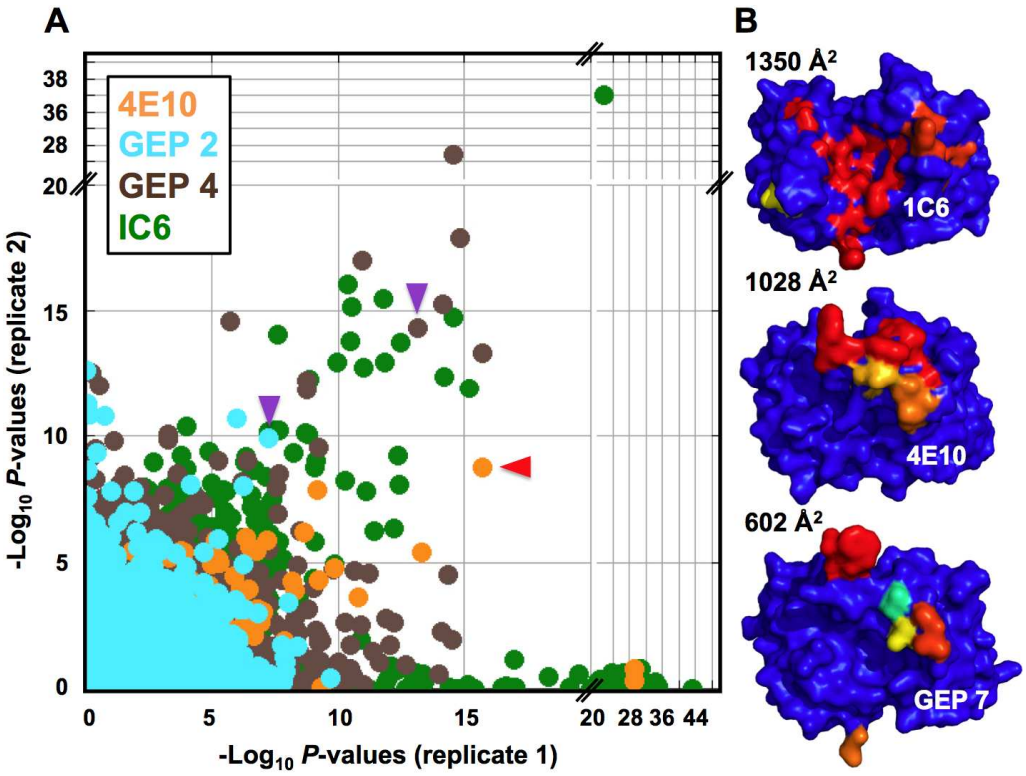

Figure 11: (A) PhIP-Seq results are plotted as -Log_{10} P values, one replicate on the abscissa, the other on the ordinate; note the discontinuity in axis scales. The top scoring 4E10 peptide derived from IP_3R is highlighted with a red arrow; one peptide, derived from the zinc finger Ran-binding domain-containing protein 3, which bound to both GEP 2 and GEP 4, is highlighted with purple arrows. Proximity to the diagonal indicates good replicate concordance; peptides with highly discordant replicate values, falling along the axes, were discarded from the analysis. **(B)** The molecular surfaces of the Fv domains of IC6 (4LCI.pdb), 4E10 (unbound; 4LLV.pdb) and GEP 7 (unbound; awaiting deposition) are shown, oriented with VH domains at left and the VL domains at right. The surface is colored to show hydrophobic patches, defined by the program HotPatch [reference]; patches are colored in descending order of total area (red, orange, yellow ...). The total surface area for red and orange hydrophobic patches is given. The crystal structure of GEP 7 is partially disordered in HCDR1 and 3 and so patch area is underrepresented in the calculation.

Table 5: PhIP-Seq results for 4E10 Fv

Rank	Avg. -log*P*	Peptide sequence	Peptide source	□	Charge
1	12.25	SKCRVFNTTERDEQGSKV NDFFQQTEDLYNEMKWQK	inositol 1,4,5-trisphosphate receptor type 2	0.17	-1
2	9.37	PPPRCISTNKCTAPEVE NAIRVPGNRSFFSLTEIVR	complement receptor type 1 isoform S precursor	0.48	+1.9
3	7.32	KAFNYRSYLTTHQRSHTG ERPYKCEECGKAFNSRSY	zinc finger protein 267	0.18	+4.1
4	7.22	PPHELTEEEKQQILHSE EFLSFFDHSTRIVERALSE	cytoplasmic dynein 1 intermediate chain 2	0.28	-5.7
5	6.22	AHTGEKPYVCRECGRGFR QHSHLVRHKRTHSGEKPY	Krueppel-related zinc finger protein 1	0.10	+5.4
6	5.21	STPIEWYPDYEVEAYRRR HHNSSLNFFNWFSDHNFA	XP_373395.2	0.49	-1.7
7	4.72	EQGIVGPRWWVFPSLRFAA VSRPFCGAWVLSWGQAT	XP_498650.1	0.71	+2
8	4.49	GAQPPFDAQSPLDSQPQPS GQPWNFHASTSWYWRQS	nuclear fragile X mental retardation-interacting protein 1	0.42	-0.9
9	4.36	LYEEISMPLLADVRLNYLG GLVGASPWAVFPNYFGG	inter-□-trypsin inhibitor heavy chain H6 precursor	0.74	-2
10	4.25	YKVDVTWTRARGASRGWR SRHQLKGRNGWRNSRVYK	TP53-target gene 5 protein	0.12	+10.1
11	4.15	AKWREVSHTFSNYPPGVR YIWFQHGGVDTHYWAGWY	F-box only protein 44 isoform 1	0.58	+1.3
12	4.07	MNPQIRNPMKAMYPGTFY FQFKNLWEANDRNETWLC	APOBEC-3C	0.61	+1
13	3.78	SGGYGSGGYGGSATPSGR ICAGVGGGYRGVSRGGFR	ATP-dep. RNA helicase A	0.22	+4
14	3.77	LMMKKRVRLEEAFEFVK QRRSIISPNFSFMGQLLQF	dual specificity protein phosphatase 4 isoform 2	0.48	+4

15	3.59	DGMYQRFLRQHVHPEETG GSDRYCNLMMQRRKMTLY	ribonuclease 4 precursor	0.28	+2.1
16	3.47	TSLEVEPFASLTEAVRSS VPRLLINRDLVGPLAWHP	NAD-dependent deacetylase sirtuin-3, mitochondrial isoform a	0.56	-0.9
17	3.40	SEQFYDRSLGIMRRVLPP GNRRYPNHRHRARINTDF	secreted phosphoprotein 24	0.26	+5.2
18	3.38	MKPGFSPRGGGFGGRGGF GDRGGRGGRGGFGGGRGR	rRNA 2'-O-methyltransferase fibrillarin	0.03	+7
19	3.31	NARRFSAGQWEARRGWRL FNCSASLDWPRMVESCLG	ribosomal biogenesis protein LAS1L isoform 1	0.43	+2.9
20	3.23	ESWLSRFSYAWLAPLLAR GACGELRQPQDICRLPHR	multidrug resistance-associated protein 7 isoform MRP7A	0.51	+2

Table 5 continued: PhIP-Seq results for GEP 2 Fv

Rank	Avg. -log*P*	Peptide sequence	Peptide source		Charge
1	8.35	EGWQCSLCTYINNSELPY CEMCETPQGSAVMQIDSL	**zinc finger Ran-binding domain-containing protein 3**	0.61	-5.2
2	6.11	PVFSFSKTSEYHDIMYPA WTFWEGGPAVWPIYPTGL	protein O-glucosyltransferase 1 precursor	0.72	-1.9
3	5.74	KVTDTKPRVAEWRYGPAR LWYDMLGVSEDGSGFDYG	transcription initiation factor TFIID subunit 1-like	0.29	-1
4	5.58	MTTFTEPEVVFLQSRGNE VCRKIWLGLFDARTSLVP	arf-GAP domain and FG repeat-containing protein 2	0.56	0
5	5.57	SGWGSRSQAPYGTLGAVS GGEQVLLHEEAGDSGFVS	peroxisome proliferator-activated receptor □coactivator-related protein 1	0.32	-2.9
6	5.05	MGLSRRNPSYPWLWEDG SPLMPHLFRVRGAVSQTYP	oxidized low-density lipoprotein receptor 1 isoform 1	0.53	+2.1
7	4.90	ASTLGSMPSFTARLTRGQL QHLGTRGSNTSWRPGTG	RING finger protein 37 isoform a	0.33	+4.1
8	4.72	QPNPHGNMMYTGPSHHSY MNAAGVPKQSLNGPYMRR	histone acetyltransferase KAT6A	0.40	+3.3
9	4.62	KLMGKDESTSRNRRSLSP ILPGRHSPAPPPDPGFPA	rho guanine nucleotide exchange factor 18 isoform a	0.25	+3.1
10	4.47	IQRYCNCNSSMPRPVKVA AVGGQSYLSSILRFFVKS	phosphofurin acidic cluster sorting protein 1	0.54	+4.9
11	4.44	RYDVQERHPKGKMIPVLH NTDLEQKKPRRKDTPALH	protein phosphatase 1 regulatory subunit 17 isoform 1	0.09	+4.3
12	4.39	RPGFAGPAVLDWEEWCPL WAGNWGRRRAYQAASWAW	**hyaluronidase-3 isoform 1 precursor**	0.58	+1
13	4.39	PIPSGSYYAPYGMPYTSM PMMNLGYYGQYPAPLYLS	SET-binding protein isoform a	0.71	0

14	4.28	NQSQGCLPARTCHSPAHS PCHIHAHCLFERNGAVSC	thrombospondin-3 isoform 1 precursor	0.50	+1.2
15	4.27	LEREVTDVDSVVGRSSVG ERRTLAQSRRLSERALSL	dedicator of cytokinesis protein 8 isoform 1	0.11	+1
16	4.23	RSRRKQHLLPPCVDEPEL LFSEGPSTSRWAAELPFP	forkhead box protein M1 isoform 1	0.38	+0.1
17	4.22	TQAFDFYSRYFAPWVGVA EDPVTGSAHAVLSSYWSQ	phenazine biosynthesis-like domain-containing protein, isoform a	0.53	-1.9
18	4.17	TKWMNMKAVFGHPFSLGW ASPFATPDQGKADPYQYV	palmitoyltransferase ZDHHC3 isoform 2	0.52	+1.1
19	4.11	AGSWHPRSYAAYALKTW AHCMSYSNSALNPLLYAFL	kiSS-1 receptor	0.68	+2.1
20	4.07	EETAGEPWEDGFEAELSP VEQKLSALRSPLAQRPFF	NP_955358.1	0.31	-5

Table 5 continued: PhIP-Seq results for GEP 4 Fv

Rank	Avg. -log*P*	Peptide sequence	Peptide source	□	Charge
1	16.36	PKGKCLGSQDYLELANRF PQQAWEEARQFFLKKEKK	S phase cyclin A-associated protein in the ER isoform a	0.21	+3
2	14.52	RPGFAGPAVLDWEEWCPL WAGNWGRRRAYQAASWAW	**hyaluronidase-3 isoform 1 precursor**	0.58	+1
3	13.98	ISHIQPFSFLDLESLRSL HLDSNRLPSLGEDTLRGL	leucine-rich repeat and fibronectin type-III domain-containing protein 2 precursor	0.50	-1.8
4	13.73	EGWQCSLCTYINNSELPY CEMCETPQGSAVMQIDSL	**zinc finger Ran-binding domain-containing protein 3**	0.61	-5.2
5	9.43	LEQNLAAAEEGPLEPAVV DAFNQAWHLFAHECPNYF	NP_957704.1	0.59	-5.9
6	9.40	QAAPAPQPVFVGPAVPQG AVMLVLPQGALPPPAPCA	Krueppel-like factor 11 isoform a	0.68	0
7	8.16	EETAGEPWEDGFEAELSP VEQKLSALRSPLAQRPFF	NP_955358.1	0.31	-5
8	7.40	SEALSRDPETLVGYSMVG CQRAMMANRLSFFFDFRG	fatty acid synthase	0.45	0
9	7.37	RKREKCWGRSSVMAEYG LYGAMVRHSIPLPESILKS	visual system homeobox 2	0.34	+4
10	6.65	LIKTKQRKESRFQLFMQ EAESHPQCRRLQLRDLIIS	rho guanine nucleotide exchange factor 11 isoform 1	0.27	+4
11	6.06	SDSSWQQQPGQPPPHSTW NCHSLSLYSATKGSPHPG	uncharacterized protein C14orf43	0.38	+0.2
12	5.98	NGKIKYECNVCAKTFGQL SNLKVHLRVHSGERPFKC	PR domain zinc finger protein 1 isoform 2	0.40	+5
13	5.97	KGKAGTPSGSSADEDTF	suppressor of cytokine	0.11	+2

		SSSSAPIVFKDVRAQRPIR	signaling 6		
14	5.75	DEMKEIQERQRDKLYERR GGGSGGGEESEGEEVDED	transcriptional activator protein Pur-	-0.25	-8
15	5.64	QPPFFSKEQPQALNFGGI GMVIGHEITHGFDDNGRN	membrane metallo-endopeptidase-like 1	0.44	-1.8
16	5.60	RTTENPTLERKPYSSPR DSSLPALTSSALAFLKSRS	zinc finger FYVE domain-containing protein 26	0.23	+3
17	5.60	NQSQGCLPARTCHSPAHS PCHIHAHCLFERNGAVSC	thrombospondin-3 isoform 1 precursor	0.50	+1.2
18	5.29	MAKKGCRHLVCSSGGNA GIAAAYAARKLGIPATIVL	serine dehydratase-like	0.48	+5
19	5.28	TTPFTLEGRPRGELHEQY RRNTVLTTMHAFPYIKTR	dedicator of cytokinesis protein 8 isoform 1	0.33	+3.2
20	5.26	MSAQSLPAATPPTQKPPR IIRPRPPSRSRAAQSPGP	rho guanine nucleotide exchange factor 15	0.25	+6

Table 5 continued: PhIP-Seq results for 1C6 Fv

Rank	Avg. -log*P*	Peptide sequence	Peptide source	□	Charge
1	30.87	AVLKYENNVMNIRQFNCS PHPYWLPNFMDVFTWSLP	Ser/Thr-protein phosphatase 2B catalytic subunit □ isoform 1	0.81	0
2	14.67	PTWDQVPPFQWSTSPFSG LLHMGQPDLWKFAPVKVS	putative phospholipase B-like 2 isoform 1	0.64	+0.1
3	13.63	QGLVLNWGLMTTRGQGLM SSWGLGAHWGLPVNLGTG	*Record removed* (XP_495971.1)	0.70	+0.1
4	13.54	AGNHFINVMLSHPNHTGN NACFPSPNYNPFTWMFFL	trans-2,3-enoyl-CoA reductase-like	0.86	+0.2
5	13.30	PIPSGSYYAPYGMPYTSM PMMNLGYYGQYPAPLYLS	SET-binding protein isoform a	0.71	0
6	13.21	MASNSSSCPTPGGGHLNG YPVPPYAFFFPPMLGGLS	retinoic acid receptor □ isoform 1	0.69	0
7	13.11	MNPQIRNPMKAMYPGTFY FQFKNLWEANDRNETWLC	APOBEC-3C	0.61	+1
8	12.85	PENRGGFQGFGFGDGGF QMSFGIGAFPFGIFATAFN	E3 ubiquitin-protein ligase RNF185 isoform 1	0.61	-1
9	12.41	MNHMLPDPGTWEEYFETFI NGKVVWGSWFDHVKGWW	sulfotransferase 1C2 isoform a	0.70	-2.8
10	12.14	LQPPPSRFKQFFCLSLPS SWDYSLPQLPWMVNSSSF	WD repeat-containing protein 27 isoform 1	0.73	+1
11	11.87	TDRYMWSDASGLQECTKA GTKPPSLQWAWVSDWFVD	tectonin □-propeller repeat-containing protein 1	0.45	-2
12	11.46	EYGPPRKQPKQQHGPGFW FQPPVCSNWGCWGGPWRP	uncharacterized protein C12orf12	0.52	+3
13	10.84	KVTDTKPRVAEWRYGPAR LWYDMLGVSEDGSGFDYG	transcription init. factor TFIID subunit 1-like	0.29	-1
14	10.80	EQGKEPWMVVREETGRWC	zinc finger protein 461	0.45	-1.9

		PGTWKTWGFHNNFLDNNE			
15	10.57	TYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGTPRQ	dnaJ homolog subfamily B member 11 precursor	0.33	-2.7
16	10.26	FWGTGLSLPSLPVSFPLQ AFCPKFRWGRTAFFSWDT	metallothionein-1E	0.75	+2
17	9.46	PVFSFSKTSEYHDIMYPA WTFWEGGPAVWPIYPTGL	protein O-glucosyltransferase 1	0.72	-1.9
18	9.45	NTTWYSNDTWYGNDTWYG NEMWYGNDSWYANDTWNS	sodium channel protein type 4 subunit □	0.60	-5
19	9.41	STRLPSEYIYGFGEVEHTA FKRDLNWNTWGMFTRDQ	sucrase-isomaltase, intestinal	0.42	-0.9
20	9.27	SQWGQWSQVYGNPQQYGQ YMANGWQVPPYGVYGQPW	nucleolysin TIAR isoform 1	0.59	0

References

1. Stephenson, K. E. & Barouch, D. H. A global approach to HIV-1 vaccine development. *Immunological Reviews* **254,** 295–304 (2013).
2. Haynes, B. F. & McElrath, M. J. Progress in HIV-1 vaccine development. *Current Opinion in HIV and AIDS* 1 (2013). doi:10.1097/COH.0b013e328361d178
3. Bonsignori, M. *et al.* HIV-1 antibodies from infection and vaccination: insights for guiding vaccine design. *Trends Microbiol.* **20,** 532–539 (2012).
4. Klein, F. *et al.* Antibodies in HIV-1 Vaccine Development and Therapy. *Science* **341,** 1199–1204 (2013).
5. Stamatatos, L. HIV vaccine design: the neutralizing antibody conundrum. *Current Opinion in Immunology* **24,** 316–323 (2012).
6. Burton, D. R. Antibodies, viruses and vaccines. *Nat Rev Immunol* **2,** 706–713 (2002).
7. Lang, J. *et al.* B cells are exquisitely sensitive to central tolerance and receptor editing induced by ultralow affinity, membrane-bound antigen. *J. Exp. Med.* **184,** 1685–1697 (1996).
8. Qi, H. Extrafollicular Activation of Lymph Node B Cells by Antigen-Bearing Dendritic Cells. *Science* **312,** 1672–1676 (2006).
9. Batista, F. D. & Neuberger, M. S. Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* **8,** 751–759 (1998).
10. Kim, M. *et al.* Antibody mechanics on a membrane-bound HIV segment essential for GP41-targeted viral neutralization. *Nat Struct Mol Biol* **18,** 1235–1243 (2011).
11. Zwick, M. B. *et al.* Broadly Neutralizing Antibodies Targeted to the Membrane-Proximal External Region of Human Immunodeficiency Virus Type 1 Glycoprotein gp41. *Journal of Virology* **75,** 10892–10905 (2001).
12. Cardoso, R. M. F. *et al.* Broadly Neutralizing Anti-HIV Antibody 4E10 Recognizes a

Helical Conformation of a Highly Conserved Fusion-Associated Motif in gp41. *Immunity* **22,** 163–173 (2005).

13. Binley, J. M. *et al.* Comprehensive Cross-Clade Neutralization Analysis of a Panel of Anti-Human Immunodeficiency Virus Type 1 Monoclonal Antibodies. *Journal of Virology* **78,** 13232–13252 (2004).
14. Correia, B. E. *et al.* Computational Design of Epitope-Scaffolds Allows Induction of Antibodies Specific for a Poorly Immunogenic HIV Vaccine Epitope. *Structure* **18,** 1116–1126 (2010).
15. Larimore, K., McCormick, M. W., Robins, H. S. & Greenberg, P. D. Shaping of Human Germline IgH Repertoires Revealed by Deep Sequencing. *The Journal of Immunology* (2012). doi:10.4049/jimmunol.1201303
16. Xu, H. *et al.* Interactions between lipids and human anti-HIV antibody 4E10 can be reduced without ablating neutralizing activity. *Journal of Virology* **84,** 1076–1088 (2010).
17. Finton, K. A. K. *et al.* Autoreactivity and Exceptional CDR Plasticity (but Not Unusual Polyspecificity) Hinder Elicitation of the Anti-HIV Antibody 4E10. *PLoS Pathog* **9,** e1003639 (2013).
18. Doyle-Cooper, C. *et al.* Immune Tolerance Negatively Regulates B Cells in Knock-In Mice Expressing Broadly Neutralizing HIV Antibody 4E10. *The Journal of Immunology* **191,** 3186–3191 (2013).
19. Chen, Y. *et al.* Common Tolerance Mechanisms, but Distinct Cross-Reactivities Associated with gp41 and Lipids, Limit Production of HIV-1 Broad Neutralizing Antibodies 2F5 and 4E10. *The Journal of Immunology* **191,** 1260–1275 (2013).
20. Ruprecht, C. R. *et al.* MPER-specific antibodies induce gp120 shedding and irreversibly neutralize HIV-1. *Journal of Experimental Medicine* **208,** 439–454 (2011).
21. Amzel, L. M. & Poljak, R. J. Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* **48,** 961–997 (1979).

22. Wu, T. T. & Kabat, E. A. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* **132,** 211–250 (1970).

23. Klein, F. *et al.* Somatic Mutations of the Immunoglobulin Framework Are Generally Requiredfor Broad and Potent HIV-1 Neutralization. *Cell* **153,** 126–138 (2013).

24. Breden, F. *et al.* Comparison of Antibody Repertoires Produced by HIV-1 Infection, Other Chronic and Acute Infections, and Systemic Autoimmune Disease. *PLoS ONE* **6,** e16857 (2011).

25. Xiao, X. *et al.* Germline-like predecessors of broadly neutralizing antibodies lack measurable binding to HIV-1 envelope glycoproteins: Implications for evasion of immune responses and design of vaccine immunogens. *Biochemical and Biophysical Research Communications* **390,** 404–409 (2009).

26. Scheid, J. F. *et al.* Sequence and Structural Convergence of Broad and Potent HIV Antibodies That Mimic CD4 Binding. *Science* **333,** 1633–1637 (2011).

27. Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329,** 811–817 (2010).

28. McGuire, A. T. *et al.* Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies. *Journal of Experimental Medicine* **210,** 655–663 (2013).

29. Bates, J. T. *et al.* Reversion of Somatic Mutations of the Respiratory Syncytial Virus-Specific Human Monoclonal Antibody Fab19 Reveal a Direct Relationship between Association Rate and Neutralizing Potency. *The Journal of Immunology* **190,** 3732–3739 (2013).

30. Wu, X. *et al.* Focused Evolution of HIV-1 Neutralizing Antibodies Revealed by Structures and Deep Sequencing. *Science* **333,** 1593–1602 (2011).

31. Johnson, G. & Wu, T. T. Kabat database and its applications: 30 years after the first

variability plot. *Nucleic Acids Research* **28,** 214–218 (2000).

32. Johnson, G. & Wu, T. T. Preferred CDRH3 lengths for antibodies with defined specificities. *International Immunology* **10,** 1801–1805 (1998).

33. Ma, B.-J. *et al.* Envelope Deglycosylation Enhances Antigenicity of HIV-1 gp41 Epitopes for Both Broad Neutralizing Antibodies and Their Unmutated Ancestor Antibodies. *PLoS Pathog* **7,** e1002200 (2011).

34. Foote, J. & Eisen, H. N. Kinetic and affinity limits on antibodies produced during immune responses. *Proceedings of the National Academy of Sciences of the United States of America* **92,** 1254–1256 (1995).

35. Thomson, C. A. *et al.* Germline V-genes sculpt the binding site of a family of antibodies neutralizing human cytomegalovirus. *EMBO J* **27,** 2592–2602 (2008).

36. Scharf, L. *et al.* Structural basis for HIV-1 gp120 recognition by a germ-line version of a broadly neutralizing antibody. *Proceedings of the National Academy of Sciences* **110,** 6049–6054 (2013).

37. Wang, F. *et al.* Somatic hypermutation maintains antibody thermodynamic stability during affinity maturation. *Proceedings of the National Academy of Sciences* **110,** 4261–4266 (2013).

38. Lingwood, D. *et al.* Structural and genetic basis for development of broadly neutralizing influenza antibodies. *Nature* **489,** 566–570 (2013).

39. Schmidt, A. G. *et al.* Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. *Proceedings of the National Academy of Sciences* **110,** 264–269 (2012).

40. Wedemayer, G. J. Structural Insights into the Evolution of an Antibody Combining Site. *Science* **276,** 1665–1669 (1997).

41. Manivel, V., Sahoo, N. C., Salunke, D. M. & Rao, K. V. Maturation of an antibody response is governed by modulations in flexibility of the antigen-combining site. *Immunity*

13, 611–620 (2000).

42. Babor, M. & Kortemme, T. Multi-constraint computational design suggests that native sequences of germline antibody H3 loops are nearly optimal for conformational flexibility. *Proteins* **75,** 846–858 (2009).

43. Sethi, D. K., Agarwal, A., Manivel, V., Rao, K. V. S. & Salunke, D. M. Differential Epitope Positioning within the Germline Antibody Paratope Enhances Promiscuity in the Primary Immune Response. *Immunity* **24,** 429–438 (2006).

44. Yin, J., Beuscher, A. E., IV, Andryski, S. E., Stevens, R. C. & Schultz, P. G. Structural Plasticity and the Evolution of Antibody Affinity and Specificity. *Journal of Molecular Biology* **330,** 651–656 (2003).

45. Yin, J. *et al.* A Comparative Analysis of the Immunological Evolution of Antibody 28B4. *Biochemistry* **40,** 10764–10773 (2001).

46. Nguyen, H. P. *et al.* Germline antibody recognition of distinct carbohydrate epitopes. *Nat Struct Biol* **10,** 1019–1025 (2003).

47. Pauling, L. A Theory of the Structure and Process of Formation of Antibodies. *J Am Chem Soc* **62,** 2643–2657 (1940).

48. Changeux, J.-P. & Edelstein, S. Conformational selection or induced-fit? 50 years of debate resolved. *F1000 Biol Rep* **3,** (2011).

49. Chothia, C. & Lesk, A. M. Canonical structures for the hypervariable regions of immunoglobulins. *Journal of Molecular Biology* **196,** 901–917 (1987).

50. Larman, H. B. *et al.* Autoantigen discovery with a synthetic human peptidome. *Nature Biotechnology* **29,** 535–541 (2011).

51. Blish, C. A., Nedellec, R., Mandaliya, K., Mosier, D. E. & Overbaugh, J. HIV-1 subtype A envelope variants from early in infection have variable sensitivity to neutralization and to inhibitors of viral entry. *AIDS* **21,** 693–702 (2007).

52. Cheng-Mayer, C. & Levy, J. A. Distinct biological and serological properties of human

immunodeficiency viruses from the brain. *Ann Neurol* **23 Suppl,** S58–61 (1988).

53. Derby, N. R. *et al.* Antibody Responses Elicited in Macaques Immunized with Human Immunodeficiency Virus Type 1 (HIV-1) SF162-Derived gp140 Envelope Immunogens: Comparison with Those Elicited during Homologous Simian/Human Immunodeficiency Virus SHIVSF162P4 and Heterologous HIV-1 Infection. *Journal of Virology* **80,** 8745–8762 (2006).

54. Davenport, T. M. *et al.* Binding Interactions between Soluble HIV Envelope Glycoproteins and Quaternary-Structure-Specific Monoclonal Antibodies PG9 and PG16. *Journal of Virology* **85,** 7095–7107 (2011).

55. Correia, B. E. *et al.* Computational Protein Design Using Flexible Backbone Remodeling and Resurfacing: Case Studies in Structure-Based Antigen Design. *Journal of Molecular Biology* **405,** 284–297 (2011).

56. Lawrence, M. C. & Colman, P. M. Shape complementarity at protein/protein interfaces. *Journal of Molecular Biology* **234,** 946–950 (1993).

57. Burton, D. R. Scaffolding to build a rational vaccine design strategy. *Proceedings of the National Academy of Sciences* **107,** 17859–17860 (2010).

58. Graves, S. S. *et al.* Antagonistic and Agonistic Anti-canine CD28 Monoclonal Antibodies: Tools for Allogeneic Transplantation. *Transplantation* **91,** 833–840 (2011).

59. Lavoie, T. B. *et al.* Structural differences among monoclonal antibodies with distinct fine specificities and kinetic properties. *Molecular immunology* **36,** 1189–1205 (1999).

60. Souto-Carneiro, M. M., Longo, N. S., Russ, D. E., Sun, H.-W. & Lipsky, P. E. Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. *J. Immunol.* **172,** 6790–6802 (2004).

61. Brochet, X., Lefranc, M.-P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic*

Acids Research **36,** W503–8 (2008).

62. Giudicelli, V., Chaume, D. & Lefranc, M.-P. IMGT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis. *Nucleic Acids Research* **32,** W435–40 (2004).
63. Wang, X. *et al.* Ab-origin: an enhanced tool to identify the sourcing gene segments in germline for rearranged antibodies. *BMC Bioinformatics* **9 Suppl 12,** S20 (2008).
64. Volpe, J. M., Cowell, L. G. & Kepler, T. B. SoDA: implementation of a 3D alignment algorithm for inference of antigen receptor recombinations. *Bioinformatics* **22,** 438–444 (2006).
65. Gaeta, B. A. *et al.* iHMMune-align: hidden Markov model-based alignment and identification of germline genes in rearranged immunoglobulin gene sequences. *Bioinformatics* **23,** 1580–1587 (2007).
66. Myszka, D. G. Improving biosensor analysis. *J. Mol. Recognit.* **12,** 279–284 (1999).
67. Otwinowski, Z. & Minor, W. in *Macromolecular Crystallography Part A* (Charles W Carter, J.) **276,** 307–326 (Academic Press, 1997).
68. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst (2007). 40, 658-674 [doi:10.1107/S0021889807021206]* 1–17 (2007). doi:10.1107/S0021889807021206
69. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59,** 1131–1137 (2003).
70. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst (2004). D60, 2126-2132 [doi:10.1107/S0907444904019158]* 1–7 (2004). doi:10.1107/S0907444904019158
71. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **53,** 240–255 (1997).
72. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and

nucleic acids. *Nucleic Acids Research* **35,** W375–W383 (2007).

Chapter 5: Future Directions

Natural killer cell receptors KIR3DS1 and KIR3DL1

Despite the positive correlation of KIR3DS1 expression with improved clinical outcome in HIV-1 infection, the fundamental question of KIR3DS1 ligand specificity has not been answered by structural or functional studies. The question is particularly vexing due to the extremely high sequence identity between activating and inhibitory KIRs. To address this question, I have been collaborating with the Geraghty lab (FHCRC), providing recombinant, soluble KIR3DS1 and KIR3DL1 molecules, to perform solution and cell-based binding studies, and molecular modeling analysis of KIR/HLA complexes in order to identify candidate ligands. We have tested the binding affinity of KIR3DS1 and KIR3DL1 using surface plasmon resonance (SPR) against a panel of chip coupled HLA-A and -B molecules containing diverse peptides in order to parse out ligand preferences (Table 1). The integrity of HLA surfaces was demonstrated by binding of monoclonal antibodies W6/32, specific for fully assembled MHC class I molecules, and mAb HC10 and HCA2, specific for free HLA heavy chains. We find that KIR3DL1 shows a binding preference for HLA molecules with the BW4 epitope with an isoleucine at position 80 and is intolerant to certain peptide substitutions at p8 as supported by the literature[1]. We also show that, as predicted by structural modeling in chapter two, KIR3DS1 is able to bind HLA free heavy chain only in the absence of peptide and does not show a preference for the BW4 epitope. (Manuscript in preparation, co-first authored by Aura Burien and Kathryn Finton)

4E10 recognition of IP_3R

A fundamental problem in eliciting 4E10-like antibodies (Abs) through vaccinations, is that

mature 4E10 is autoreactive (chapter three), and thus deleted by immune tolerance mechanisms. One question that will need to be addressed when pursuing this Ab as a vaccine target, is whether or not recognition of the proposed 4E10 autoantigen, the inositol trisphosphate receptor (IP_3R), is coupled with recognition of the 4E10 epitope: an obstacle that cannot merely be solved by immunogen design. The first step in answering this question will be to characterize the 4E10 paratope in the binding interaction with IP_3R through X-ray crystallographic structure determination of the 4E10/IP_3R complex. Since IP_3R is a multi-pass transmembrane protein, and therefore not amenable to expression or crystallization, a single stable domain of the protein identified by trypsin cleavage, including the 4E10 epitope, was expressed as a fusion protein to murine siderocalin (muScn) in an HEK293 lentiviral expression system[2]. The complex will be isolated by size exclusion chromatography and screened for crystallizability using sub-microliter robotics and commercial factorial screens. Preliminary crystals will be optimized and diffraction data will be collected in house or at the Advanced Light Source, Berkeley. Data will be phased by molecular replacement using muScn (3S26.pdb) and 4E10 (4LLV.pdb, 3LH2.pdb) crystal structures.

4E10 ontogeny

NMR studies of flexibility

The magnitude of complementary determining region (CDR) loop flexibility seen in the mature 4E10 Ab, as determined by comparison of CDR conformations in the bound versus unbound state, is extraordinary, both in its implications for binding and neutralization mechanisms and as it challenges the canon that Abs rigidify during maturation in order to gain affinity. Since crystallographic structures provide static views of a single protein conformation with only limited

information regarding protein dynamics via B-factors, NMR will be used to quantitate differences in CDR conformational plasticity between 4E10 germline encoded precursors (GEPs) and mature 4E10. These results will be correlated with thermostability, binding affinity, and polyspecificity and neutralization profiles determined in chapters three and four in order to help determine the impact of CDR flexibility on Ab specificity and affinity. GEP sequence variations will also be examined in order to shed light on mutations that affect CDR mobility. A preliminary ^{1}H-^{15}N HSQC-TROSY spectrum of GEP 1 was obtained through the Klevit lab (University of Washington) (figure 1), showing good spectral resolution of isotopically-labeled GEP 1. Both ^{15}N-labled 4E10 and GEP 1 Fvs will be analyzed with picosecond through microsecond time regimes where backbone motions are known to occur.

Binding affinity and neutralization potency of GEP and 4E10 mutants

The findings that 4E10 did not undergo preconfiguration of the binding site and that only two mutations (Y/K32L and P/L95H) resulted in improved epitope contacts for 4E10 over its GEPs, prompts the obvious question of whether or not substitution of these residues into a GEP could restore 4E10-like binding affinity and neutralization potency. Support for the hypothesis that 4E10 employs higher order mechanisms (other than just epitope binding) in order to neutralize HIV would be supported in the case that the mutations are able to restore binding affinity, but not neutralization potency. We would also like to answer whether or not the F54H mutation present in GEPs could inhibit neutralization, as this prominent residue may sterically hinder the approach of the Fv to the 4E10 epitope. Two mutants will be made in order to address these questions. The GEP 1 mutant will contain the Y/K32L mutation that will replace a strained hydrogen bond with a stronger salt bridge, and the P/L95H mutation that will prevent reordering of the LCDR3 loop by proline, restoring the hydrogen bond mediated by Ser94L. The other mutant will revert L54L in mature 4E10 to F54L. The GEP 1 mutant will be refolded from

bacterial inclusion bodies and the 4E10 mutant will be made in the Daedalus system (see materials and methods, chapters 3 and 4, respectively).

B cell activation in 4E10 and GEP retrogenic mice

In order to complete the characterization of 4E10 ontogeny and determine whether or not elicitation of 4E10-like bnAbs through vaccination is feasible, retrogenic and transgenic mice expressing mature or GEP BCR sequences have been made by the Greenberg lab (University of Washington). Retrogenic mice (generated as in (3), figure 2A) possess only B cells with the GEP immunoglobulin genes and a fluorescent protein marker. Immunizations in these mice using epitope-scaffold (ES) immunogens will allow us to determine, first, whether or not GEPs are autoreactive and, second, whether or not ESs are able to activate B-cells through *in vivo* vaccinations. Post-immunization, the proliferative responses of B cells expressing GEPs will be quantified by flow cytometric analysis in conjunction with phenotypic analysis of germinal center and plasmablast differentiation. Biophysical characterizations of GEP and ES interactions discussed in chapters three and four (affinity, kinetics, thermodynamics, thermostabilities) will be correlated with the extent and quality of B cell activation. ESs can then be ranked on their ability to activate GEP expressing B cells, and those showing the widest breadth and potency will be considered for prime boost immunization strategies.

Development of a prime-boost immunization strategy in GEP transgenic mice

In order to more closely recapitulate the human immune response following vaccination, transgenic mice were created which have Igκ and IgH variable regions inserted at the mouse Ig

loci (generated as in (4), figure 2B). Placement of GEP genes in this region allows for somatic hypermutation and class switching needed for the generation of mature, high affinity Abs. Naïve B cells purified from GEP knock-in mice (B6 background, CD45.1^{neg}/CD45.2^{+}) will then be transferred into B6-CD45.1^{+} congenic recipient mice in order to better reflect a normal naïve B cell frequency among a polyclonal B cell population and to allow for tracking of transgenic B cells and their progeny by flow cytometry of CD45 cell surface markers. Transgene recipient mice will be primed with high-ranking ESs, and serum responses to the prime will be evaluated by SPR and ELISA for binding to candidate boosting immunogens. Sera will also be tested for its neutralization potency against a panel of HIV-1 isolates in the event that parameters other than affinity, kinetics and thermodynamics are responsible for neutralization. Following boosting, serum responses will again be analyzed for ES binding and neutralization potency. We will also investigate, by flow cytometry, the effect of boosting on germinal center, memory and plasma B cells to determine if immunogens were able to induce expansion and plasma cell development. The identity and extent of somatic mutations in response to immunizations will also be explored through sequencing of the V-D-J region to determine if 4E10-like mutations were selected for by the ES immunogens. By identifying ES immunogens able to drive desired 4E10-like mutations (e.g. L32K, which adds an additional epitope-contacting salt bridge, chapter four) resulting in better affinities and neutralization potencies, we hope to develop a successful prime-boost immunization strategy.

Figures

Table 1. Interaction of KIR3DS1 and KIR3DL1 with pHLA molecules and empty HLA molecules as determined by SPR. (+) indicates that binding was detected; (-) indicates that binding was not observed. The PDB coordinates are given for the structures used in modeling the complexes.

Epitope	HLA allele	Peptide	PDB	3DL1	3DL1 (empty)	3DS1	3DS1 (empty)
---	HLA-A1	CMV UL 109/5 (unknown sequence)	3B08	-	-		+
---	HLA-A0301	RLRAEAQVK	2XPG	-	-	-	+
BW4	HLA-A2301	PYLFWLAAI	2BCK	+	-	-	+
BW4	HLA-A2402	SFHSLHLLF	4F7T	-	-	-	+
BW4	HLA-B5701	ISPRTLNAW	3VH8	+	-	-	+
BW4	HLA-B2705	KRWIILGLNK	2BSS 2BST 2BSR	+ (weak)	-	-	+
BW6	HLA-B0702	RPPIFIRRL	3VCL	-	-	-	+
Bw6	HLA-B3503	DPNPQEVVL	2H6P	-	-	-	+
Bw6	HLA-B4501	AEMKTDAA	---	-	-		+
Bw6	HLA-B5502	KPWDVIPMV	---	-	-	-	+
Bw6	HLA-B0801	FLKEKGGL	3SPV	-	-	-	+
---	HLA-Cw0304	GAVDPLLAL	---	-	-	-	+
---	HLA-Cw1203	IAVGLLLYCKA	---	-	-	-	+

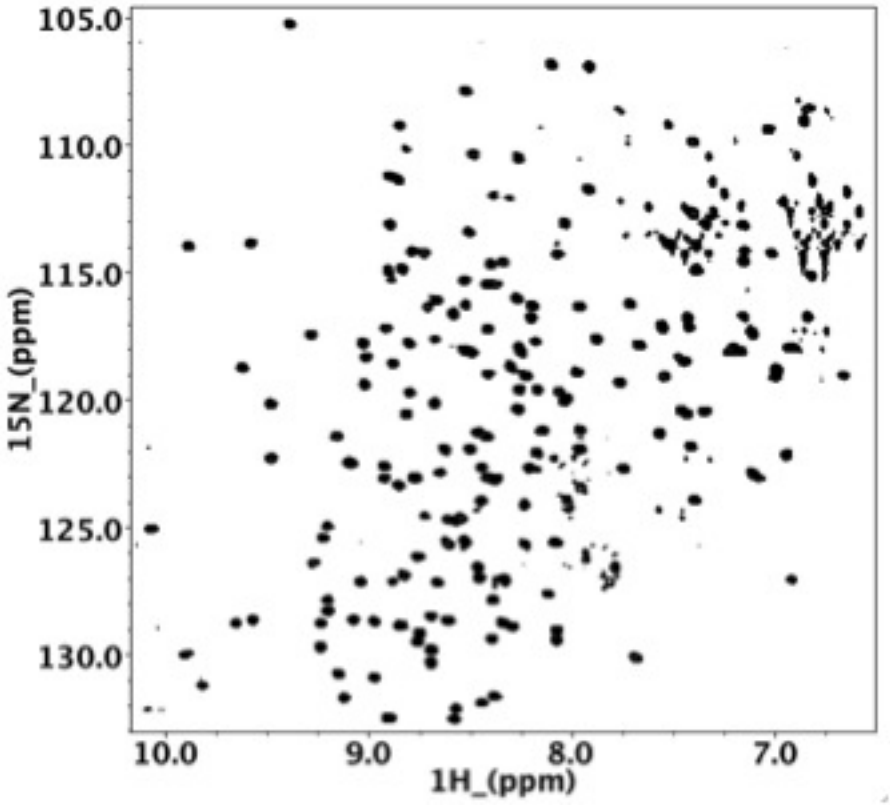

Figure 1. ^{1}H-^{15}N HSQC-TROSY spectrum of GEP #1 (100 μM, 22°) collected on a Bruker-AV III 800 with cryoprobe.

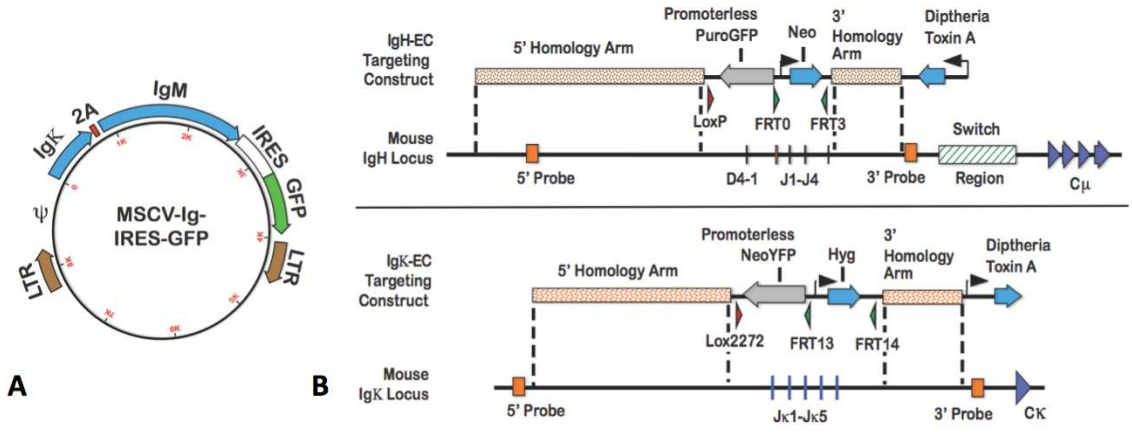

Figure 2. (A) Retroviral MSCV-Ig-IRES-GFP expression vector containing the Igκ and IgM of interest separated by the picaronavirus ribosomal skip peptide. (B) Constructs utilizing flippase recognition targeting (FRT) of GEP J segments into the mouse Ig loci. Cells are selected for by G418 (upper) or hygromycin (lower) resistance. Non-targeted integration is reduced by diphtheria toxin expression.

References

1. Vivian, J. P. *et al.* Killer cell immunoglobulin-like receptor 3DL1-mediated recognition of human leukocyte antigen B. *Nature* **479,** 401–405 (2011).
2. Bandaranayake, A. D. *et al.* 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 Research* **39,** e143–e143 (2011).
3. Holst, J. *et al.* Generation of T-cell receptor retrogenic mice. *Nat Protoc* **1,** 406–417 (2006).
4. Qiao, J., Oumard, A., Wegloehner, W. & Bode, J. Novel tag-and-exchange (RMCE) strategies generate master cell clones with predictable and stable transgene expression properties. *Journal of Molecular Biology* **390,** 579–594 (2009).