Published in final edited form as:

Vaccine. 2008 June 6; 26(24): 3059–3071. doi:10.1016/j.vaccine.2007.12.004.

Identification of immunogenic HLA-B7 “Achilles' heel” epitopes within highly conserved regions of HIV

Anne S. De Groot^{a,d,e,*}, Daniel S. Rivera^{a,b}, Julie A. McMurry^a, Soren Buus^c, and William Martin^a

^a EpiVax, Inc., Providence, RI, United States

^b University of Edinburgh, UK

^c University of Copenhagen, Denmark

^d Brown University, Providence, RI, United States

^e University of Rhode Island, Providence, RI, United States

Summary

Genetic polymorphisms in class I human leukocyte antigen molecules (HLA) have been shown to determine susceptibility to HIV infection as well as the rate of progression to AIDS. In particular, the HLA-B7 supertype has been shown to be associated with high viral loads and rapid progression to disease. Using a multiplatform in silico/in vitro approach, we have prospectively identified 45 highly conserved, putative HLA-B7 restricted HIV CTL epitopes and evaluated them in HLA binding and ELISpot assays. All 45 epitopes (100%) bound to HLA-B7 in cell-based HLA binding assays: 28 (62%) bound with high affinity, 6 (13%) peptides bound with medium affinity and 11 (24%) bound with low affinity. Forty of the 45 peptides (88%) stimulated a IFN- γ response in PBMC from at least one subject. Eighteen of these 40 epitopes have not been previously described; an additional eight epitopes have not been previously described as restricted by B7. The HLA-B7 restricted epitopes discovered using this in silico screening approach are highly conserved across strains and clades of HIV as well as conserved in the HIV genome over the 20 years since HIV-1 isolates were first sequenced. This study demonstrates that it is possible to select a broad range of HLA-B7 restricted epitopes that comprise stable elements in the rapidly mutating HIV genome. The most immunogenic of these epitopes will be included in the GAIA multi-epitope vaccine.

Keywords

T cell epitope; HIV; HLA-B7; Vaccine

Introduction

The development of a safe and efficacious HIV vaccine is widely believed to be essential for stopping the AIDS pandemic [1,2]. Vaccines designed to generate an antibody response against the surface protein of HIV-1 such as AIDSVAXTM [3–6] and T-cell directed vaccines such as the Merck Ad5 vaccine [7] have been unable to protect high-risk HIV-negative individuals from infection. These failures are believed to be due in part to the high degree of variability in the HIV genome and to limitations on the breadth of epitope-specific antibody and T-cell

* Corresponding author at: EpiVax, 146 Clifford Street, Providence, RI 02903, United States. Tel.: +1 401 272 2123; fax: +1 401 272 7562. E-mail address: AnnieD@EpiVax.com (A.S. De Groot).

responses established following vaccination [8,9]. These obstacles have significantly hampered efforts to develop HIV vaccines that are focused on the generation of an antibody response.

HIV-1 specific T helper cells and cytotoxic T cells (CTL) play a central role in control of the virus following infection [10–18]. Thus, vaccines that are able to stimulate T cells will be able to boost humoral immune responses and may also delay or prevent the progression of HIV to AIDS in infected individuals. The most effective vaccine should elicit both broadly cross-reactive neutralizing antibodies as well as potent T-cell responses, in order to curtail the expansion of any virus that evades the humoral response.

However, the genetic variability of the HIV virus constitutes a significant challenge to efforts to design a globally relevant HIV vaccine driven by cellular immune response [19,20]. Clade-to-clade differences in T-cell epitopes have been well documented [21]. Sequencing studies have shown that isolates derived from many different clades circulate within each geographic area, making it more difficult to select candidate vaccine epitopes that are conserved in all strains of HIV. One way to address the variability of HIV is to develop vaccines from isolates representing the circulating (regional) clades. However, the same questions that had been raised about single clade vaccines are likely to be reiterated: will these vaccines protect against challenge by heterologous HIV-1 isolates? How many country-specific vaccines will we need?

In contrast to country-specific vaccine efforts, we believe that there is a real need for a globally relevant HIV vaccine, comprised of epitopes that are conserved across clades. We have proposed to develop an HIV vaccine that would be composed of those epitopes that comprise the most stable elements of the rapidly mutating HIV genome. These regions may represent the 'Achilles' heel' of the virus such as functional domains that cannot be mutated in response to immunological pressure [22]. The continued presence of these epitopes, over time, suggests that they lie in regions of the HIV genome that may be resistant to selective pressure because they insure viral fitness [23,15].

Recent studies have shown that the mutations observed in HIV-1 are associated with both host and population HLA class I alleles and were least likely to occur in functional domains [24]. Thus, we have embarked on a search for epitopes that are restricted by common HLA alleles and that are, in addition, conserved over time in the HIV genome, indicating that escape (from these epitopes) is less likely to occur. Studies by DeLisi [25] and Sette [26] have demonstrated that epitope-based vaccines containing epitopes restricted by six "supertype" HLA such as HLA-B7 can provide the broadest possible coverage of the human population. We have therefore proposed to develop a vaccine that will include epitopes restricted by these common HLA supertypes: HLA-A1, HLA-A2, HLA-A3, HLA-A24, HLA-B7 and HLA-B44. In this study, we describe the identification of highly conserved and immunogenic HLA-B7 epitopes for the GAIA HIV vaccine.

The development of a protective immune response to HIV in the context of the HLA-B7 is particularly important in order to reduce the susceptibility to HIV worldwide. HLA-B7 is an allele that is highly prevalent regardless of the geographic location of the population. Individuals with rare class I HLA have been shown to have a reduced susceptibility to HIV due to decreased selection pressure on CTL epitopes restricted by these rare alleles [27]. Conversely, HLA-B7 epitopes have been demonstrated to have been reduced due to the prevalence of HLA-B7 in human populations, leading to an association between HLA-B7 and susceptibility to infection and higher viral loads [28]. Thus, the induction of broad and dominant HLA-B7 restricted immune response targeted to the regions in the viral genome that are the least capable of mutation due to functional restraints is an important strategy for HIV vaccine design. In this study, we applied two advanced immunoinformatics tools, Conservatrix

and EpiMatrix, to identify highly conserved putative HLA-B7 restricted T cell epitopes and have validated these epitopes by confirming immune response to the epitopes in vitro assays using blood from subjects who have been chronically infected with HIV.

Materials and methods

Selecting a highly conserved HIV-1 sequence data set

The sequences of all HIV-1 strains published on GenBank between 1 January 1990 and 21 April 2004 (the most recent data available at the time of this analysis) were obtained. Sequences posted to GenBank prior to 31 December 1989 were excluded based on our observation that early sequences were more likely to be derived from HIV clade B. Sequences shorter than 80% and longer than 105% of a given protein's nominal length were also excluded. Short sequences were excluded because inclusion of these fragments skews the selection of conserved epitopes in favor of regions of particular interest to researchers, such as the CD4 binding domain, or the V3 loop of HIV (unpublished observation). Longer sequences were excluded because these sequences tend to cross protein boundaries confusing the categorization process. A second dataset was downloaded from the Los Alamos HIV Database using the same criteria and the two datasets were merged. The combined dataset contains 10,803 unique entries selected for the next phase of the analysis.

Conservatrix was then used to search the 10,803 protein sequences for segments that were highly conserved among the input sequences. Peptides conserved in at least 5% of the input sequences were retained for further analysis.

Epitope selection

The EpiMatrix algorithm was used to select peptides from the 5494 highly conserved nine-mer sequences that were predicted to have high HLA-B7 binding affinity [29,30]. Each amino acid was scored for predicted affinity to the binding pockets using the HLA-B7 matrix motif. Normalized scores were then compared to the scores of known HLA-B7 ligands. Nine-mers scoring higher than 1.64 on the EpiMatrix Z scale (the top 5% of all scores) were selected as they fell within the same Z score range as published HLA-B7 epitopes and were presumed to be likely to bind to HLA-B7. The final set of 45 epitopes was selected in April 2004; 14 had been published previously at that time (Table 1).

Ranking the conserved epitopes

Since the selection of these epitopes, a new algorithm, Aggregatrix, was recently developed at EpiVax; Aggregatrix iteratively searches for the combination of epitopes that achieves maximal cross-clade representation. In this study, we used Aggregatrix to compute the percent coverage of our HIV epitopes for the non-redundant set of sequences from our internal HIV sequence dataset (described above). In 2005, we evaluated each peptide individually and the set in aggregate for coverage of strains by year 1995–2005, strains by country of origin, and strains by clade. The percent coverage reflects the proportion of ORFs in our non-redundant 1995–2005 isolate set that contain the peptide in question. The aggregate coverage reflects the proportion of HIV isolates that contains at least one peptide from our selected set (all of our HLA-B7 peptides). In aggregate, the B7 peptide set covered between 54 and 86% of strains in a given year, between 33 and 100% of strains in a given country and between 0 and 100% of strains in a given clade. Collectively, as shown at the bottom of the graph in Fig. 1, these highly conserved HLA-B7 peptides cover 85, 78, 78, and 80% of sequences representing clades A, B, C, and D, respectively. On average, the chimeric clade sequences were 76% covered by our HLA-B7-restricted epitopes.

To visualize Aggregatrix, we designed a matrix of color-gradient squares (Fig. 1) to denote percent coverage. As can be seen in this figure, the HLA-B7 epitopes are not only well conserved within years (see last row at the bottom of the figure for the aggregate coverage, per year) but also conserved across years. This is remarkable breadth of coverage for a limited set of HLA-B7 epitopes, given the well-known ability of HIV to mutate away from HLA [31, 32].

Peptide synthesis

Peptides corresponding to the epitope selections were prepared by 9-fluoronylmethoxycarbonyl (Fmoc) synthesis on an automated Rainen Symphony/Protein Technologies synthesizer (Synpep, Dublin, CA). The peptides were delivered 90% pure as ascertained by HPLC, Mass Spec, and UV scan (insuring purity, mass and spectrum, respectively).

Purified HLA class I binding assay

MHC I binding assays were performed as previously described [33]. Briefly, the HLA class I molecule was incubated at an active concentration of 2 nM together with 25 nM human beta2 microglobulin (b2m) and an increasing concentration of the test peptide at 18 °C for 48 h. The HLA molecules were then captured on an ELISA plate coated with the pan-specific anti-HLA antibody, W6/32, and HLA-peptide complexes were detected with an anti-b2m specific polyclonal serum conjugated with Horse Radish Peroxidase (Dako P0174) followed by a signal enhancer (Dako Envision). The plates were developed, and the colorimetric reaction was read at 450 nm using a Victor2 Multilabel ELISA reader. Using a standard, these readings were converted to the concentration of HLA-peptide complexes generated, and plotted against the concentration of test peptide offered. The concentration of peptide required to half-saturate (EC50) the HLA was determined. At the limited HLA concentration used here, the EC50 approximates the equilibrium dissociation constant, K_d . Relative affinities, high binders ($K_d < 50$ nM), medium binders (50 nM $< K_d < 500$ nM), low binders (500 nM $< K_d < 50,000$ nM), and non-binder ($K_d > 50,000$ nM) were based on comparison with known HLA-B7 ligands.

Cell-based HLA class I binding assay

In vitro evaluation of MHC binding was performed by measuring the ability of exogenously added peptides to bind to the surface of B-LCL expressing the class I MHC allele, B7, as described [34]. This assay was done in high throughput by using the PCA 96 micro-flow cytometer (Guava Technologies). Each test peptide, at 15 mM concentration, was added to 4×10^5 B*07-expressing B-LCL cells in a 96-well plate, and allowed to incubate for 1 h. A FITC-labeled reference peptide was added (150 nM) and the plate was placed at 4 °C for 12–18 h. After incubation, the cells were washed in PBS containing 1.0% BSA, fixed in 0.5% paraformaldehyde, and analyzed by flow cytometry using the micro-flow cytometer. The ability of each test peptide to inhibit the binding of the FITC-labeled reference peptide was measured and compared to a standard curve generated by adding varying concentrations (0.0375, 0.075, 0.15, 0.3, 0.6, and 1.2 μ M) of unlabeled reference peptide. Data analysis was performed using the EasyCyte software package.

Blood samples

ELISpot assays were performed using peripheral blood mononuclear cells (PBMC) separated by Ficoll centrifugation whole blood. HIV-seropositive subjects living in Providence were recruited in accordance with all federal guidelines and institutional policies. The study subjects belonged to a cohort of long-term slow- or non-progressors (CD4 > 350 for >10 years with minimal or no treatment), or from healthy, chronically HIV infected patients (CD4 > 350 ; on or off treatment, viral load $<50,000$), recruited at an HIV clinic situated in Providence Rhode

Island (Miriam Hospital). A heparinized blood sample was drawn after obtaining informed consent. HIV-seronegative discarded blood specimens were obtained from the Rhode Island Blood Center (RIBC). RIBC donor blood was also obtained in accordance with all federal guidelines and institutional policies.

ELISpot assays

The frequency of epitope-specific T lymphocytes was determined using Mabtech® IFN- γ ELISpot kits according to manufacturer's instructions (Mabtech, Sweden). Washed PBMC from each donor were added at 2×10^5 cells per well to each of two ELISpot plates pre-coated with anti-IFN- γ antibody. To the ELISpot plates were added individual peptides at 10 μ g/ml in triplicate, positive controls PHA (10 mg/ml) in triplicate, and CEF (10 μ g/ml) in triplicate. Twelve to 15 wells per plate of no stimulus were used as background. Biotinylated anti-IFN- γ was then added followed by streptavidin-HRP. The ELISpot plates were incubated overnight and then washed. Following the washes, ELISpot plates were developed by addition of TMB substrate. The frequency of antigen-specific cells was calculated as the number of spots per 10^6 PBMC seeded. Responses were considered positive if the number of spots was at least twice background and was also greater than 20 spots per one million cells over background (one response over background per 50,000 PBMC). This relative number of spots can be expected when stimulating with peptide as compared to the larger responses expected with whole protein. Results considered positive by these dual criteria are generally consistent with but more stringent than statistical significance (at $p < 0.05$) using the two-tailed non-parametric Mann-Whitney *U*-test to compare the number of spots in the peptide wells with that of the control wells.

Results

Epitope mapping and selection

Forty-five epitopes were chosen based on the criteria described in the methods. The Conservatrix algorithm was used to search a database of 10,803 unique HIV-1 protein sequences for highly conserved sequences. These sequences were parsed into overlapping 9-mers. The EpiMatrix algorithm was used to analyze these conserved peptides for predicted binding affinity to the class I MHC allele, B7. Normalized scores were compared to the scores of known B7 ligands. Nine-mers scoring higher than 1.64 on the EpiMatrix Z-scale represent the top 5% of all scores and are likely to bind. The 45 peptides chosen for this study had EpiMatrix Z scores between 1.92 and 4.45. All 45 of the peptides were compared to the HIV Molecular Immunology Database (<http://www.hiv.lanl.gov/content/immunology/index.html>). Of these 45, this study confirms 40 as epitopes (ELISpot positive). Of these 40 epitopes, 13 were published prior to the selection of the peptides and an additional nine were published subsequently. This publication thus describes 18 epitopes for the first time and an additional eight epitopes that were previously published but for alleles other than HLA-B7.

Epitope conservation

The conservation of these epitopes across years, clades, and countries is illustrated in Fig. 1. Each column of the matrix represents the set of HIV proteins that falls into a given category (year isolated, clade, or country). Each row of the matrix represents a single 9-mer or 10-mer that was selected as a B7 epitope. The color of a given box in the matrix represents the percentage of isolates that contain the peptide referenced on that row. The isolates are limited to the origin protein of the chosen peptide. For example, if the peptide on a row comes from POL, each square shows what percent of POL proteins from each year (clade, country) that contains this peptide. The bottom row of the matrix shows the percentage of each protein set that contains at least one peptide from our pool. Black boxes indicate that no isolates of the protein are available for that year, clade, or country.

In vitro peptide binding to soluble HLA-B7

The 45 peptides selected for in vitro evaluation were evaluated for binding to HLA-B7 using a soluble HLA binding assay (Table 2). Seven of the 45 peptides bound with high affinity (16%), 3 bound at medium affinity (7%), 20 bound at low affinity (47%), and 13 exhibited no detectable binding (30%). Two of the peptides, nef-1124 and gag-1156, were not tested in the soluble assay. Due to discrepancies between soluble binding assay and ELISpot assay, we evaluated the epitopes in a second, cell-based binding assay. Using this assay, 28 of the peptides were considered high affinity ligands (62%), 6 peptides bound with medium affinity (13%) and 11 with low affinity (24%).

Subjects

Seven HIV-infected subjects were recruited at the Miriam Hospital Immunology Center. Informed consent was provided. The subjects are listed in Table 3 and their CD4 counts and viral loads are shown. A criterion for entry in the study was detectable viral load but under 10,000 because we have observed that subjects with undetectable viral loads also have very low CTL responses. These subjects were on antiretroviral therapy as indicated. Information on resistance, clinical course, and further details on the stage of disease is not known. In general, all of the subjects were healthy at the time they were recruited for the study.

For this limited cohort of chronically HIV-infected subjects there was no clear association between viral load, CD4 T-cell count, and years of known HIV-1 infection and the number of responses to HLA-B7 epitopes. Nor was there an association between having more than one HLA allele that could present the peptide and immune response to these epitopes.

ELISpot assays

PBMC from seven HLA-B7 subjects (Table 3) were evaluated for IFN- γ secretion in response to each peptide of the 45 B7 peptides. Forty of the 45 peptides (88%) stimulated a IFN- γ response in at least one subject.

PBMC from seven study subjects were evaluated for IFN- γ secretion in response to each member of the set of 45 B7-binding peptides. Forty of the 45 peptides (89%) stimulated a IFN- γ response in at least one subject (Table 4). PMBC from all seven subjects responded to the peptides. The number of positive peptide responses per subject ranged from 5/45 to 29/45, with an average of 13.4/45.

The number of responding subjects per peptide ranged from 0/7 subjects to 5/7 subjects, with an average of 2.1/7 subjects. Many of the peptides were novel when first identified using our methodology (only 16 had been published at the time of selection) however an additional eight of the epitopes were subsequently confirmed by others even though the HLA restriction was either not confirmed (positive T-cell response was defined by inclusion in a pool of peptides reactive in ELISpot assays) or did not match our prediction and binding assay results (HLA-B7). Twenty-two of the 24 published epitopes were re-confirmed in this study.

There were seven peptides that generated a positive response in the PMBC of five of the seven subjects. Of these, five peptides had previously been published: NEF-1133 TPGPGIRYPL [35], GAG-1164 GPKEPFRDY [36], VIF-1136 HPRISSEVHI [37], NEF-1124 FPVRPQVPL [38] and ENV-1137 IPRRIRQGL [39]. GAG-1150 GPSHKARVL [28] and NEF-1139 GPGTRFPLTF [40] were novel at the time of selection but subsequently validated though HLA type was not defined.

NEF-1163 VPLRPMTYKA, one of two peptides that generated a positive response in the PMBC of four subjects, was previously published by Gahery-Segard [41] the HLA type was

not defined as B7 in that publication. However, POL-1134 LPGRWKPKMI had not been published previously for any allele; recognition by four of only seven subjects in this study suggests that this novel epitope may be immunodominant.

There were four peptides that stimulated responses in PBMC from three patients: GAG-1166 IPMFSALSEG and VPR-1125 FPRPWLHGL were published prior to 2004 [42,43] POL-1151 TPVNIIGRNL and POL-1132 VPRRKAKII have been described subsequently [44,45]. HLA restrictions identified for these four epitopes did not include HLA-B7.

Twelve peptides were positive in ELISpot assays with PBMC from two subjects. Two of these, GAG-1156 HPVHAGPVA and ENV-1146 RPNNTTRKSI [46] were published B7 epitopes at the time of selection. POL-1152 FVNTPLVLK [47] and TAT-1140 EPVDPNLEPW [28] were subsequently published for other alleles. POL-1130 LPPIVAKEI was subsequently published as B7 restricted [48]. The remaining seven epitopes recognized by two subjects are novel; these are: POL-1143 DPIWKGPALK, ENV-1148 TPLCVTLNCT; ENV-1158 GPCTNVSTV; GAG-1147 APPAESFRF, GAG-1159 RPEPTAPPA, POL-1128 IPYNPQSQGV, VIF-1135 IPLGDARLVI.

Fifteen peptides, including 10 novel peptides, were positive in ELISpot assays for one of the seven subjects. The novel epitopes are: ENV-1129 VPTDPNPQEI, ENV-1131 SPLSFQTRL, GAG-1162 EPTAPPAESF, POL-1153 YPGIKVKQL, POL-1157 TPGIRYQYNV, POL-1165 IPSTNNETPG, TAT-1160 QPKTACTNCY, VIF-1138 TPKKIKPPL, VIF-1154 IPLGEAKLV, VIF-1155 PPLPSVRKL. Two additional epitopes NEF-1127 YPLTFGWCF and POL-1167 were novel at the time of selection but subsequently published [48,49]. GAG-1144 NPPIPVGEI [50] POL-1145 LPEKDSWTV [50] and ENV-1161 IPIHYCAPA [51] were published prior to selection in 2004.

Five of the 45 peptides stimulated no responses in this limited cohort of seven subjects. Of these, POL-1142 has been subsequently confirmed since the time of selection, though HLA restriction was not specified [48]. POL-1141 EPFRKQNPDI was previously reported as an A11 epitope [52]. The remaining epitopes ENV-1168 (FDITNWLWYI), GAG-1149 (IPVGDIYKRW), POL-1126 (FPISPIETV) have yet to be described as epitopes.

Only one of the four HIV seronegative subjects responded to any of the 45 epitopes; the one peptide was POL-1153 (data not shown).

Discussion

The high variability in HIV-1 has complicated efforts at generating an effective and globally relevant vaccine. Viral adaptation in response to immune pressure, both humoral and cellular, not only alters established immune targets, but may also introduce irrelevant “dummy” targets with lower immunogenicity or located in non-functional sites [53]. The GAIA HIV vaccine development program is focused on identifying T cell epitopes least susceptible to mutation due to functional or structural constraints. We believe that a robust and focused immune response targeted to the “Achilles' heel” of HIV will prevent the generation of CTL escape mutants and effectively neutralize the virus. In this study we sought to identify highly conserved and immunogenic HLA-B7 epitopes for incorporation into the GAIA vaccine.

Prior to the development of bioinformatics tools for T-cell epitope selection, the cost and effort required to identify T-cell epitopes from protein sequences was a significant barrier to the development of novel epitope-based vaccines. EpiMatrix has been successfully applied to the analysis of published epitopes [54] to the prospective selection of CTL epitopes from an HIV-infected patient's HIV quasi-species sequences [55] to the identification of novel Mtb SOD protein B7-restricted epitopes [56] to the derivation of novel T helper epitopes from the genome

of *S. gingivalis* [57], *F. tularensis* [58] and *M. tuberculosis* [58,59] to the discovery of new HIV clade E A11-restricted epitopes [55] and to the selection of novel HIV CTL epitopes that are conserved across clades [60].

Using the immunoinformatics tools EpiMatrix and Conservatrix, we were able to efficiently scan over 10,000 HIV sequences for putative HLA-B7 epitopes located in highly conserved regions. In some cases, the HLA-B7 restricted epitopes defined in this study were conserved in as many as 93% of comparison sequences. The average conservation of an individual epitope sequence was 42%. One sign that the epitopes are stable in the face of rapidly evolving virus quasispecies is that a number of them have been conserved in HIV since the beginning of the epidemic. These sequences containing these epitopes may be essential for viral fitness. As the epitopes are associated with functional immune responses, they do not appear to represent stable escape mutations that develop during viral evolution and become fixed in the sequence at the population level [61].

In a standard consensus sequence approach, amino acids at each position are selected based upon the frequencies of occurrence in the viral sequences present in the dataset. By contrast, these peptides also provide an advantage over this approach because the intact sequence of each epitope is exactly conserved. A definitive study has shown that sets of peptides that are identical to epitopes present in the infecting virus are better inducers of immune response from infected individuals than the clade-specific consensus peptides [62].

Not surprisingly, the highest ELISpot responses, in both magnitude and number of subjects responding, were observed in previously published epitopes. Three of the novel peptides were also shown to be broadly recognized, as confirmed in ELISpot assays. However, with one exception, the 19 previously unpublished epitopes were shown to be positive in only one or two of the seven subjects. Whether this is due to the small sample size remains to be evaluated. Three of the novel peptides were unable to stimulate an IFN- γ response by memory T cells and may not represent a functional T-cell epitope, even though all of the putative epitopes were shown to bind to HLA B in cell-based binding assays. Further studies with a larger cohort would need to be performed in order to determine whether the peptides are truly HLA-B7 restricted epitopes.

This study also afforded an opportunity to compare two different peptide binding assay formats. In general, the two assays yielded similar results, however, some discrepancies were noted. These may result from the potential for peptide degradation and/or further processing in the cellular binding assay, which also bears upon the observed difference in the ability to predict an epitope that is able to stimulate IFN- γ secretion. For each of the HLA binding methods, T-cell activation was shown to be concordant with MHC binding affinity. Concordance between the ELISpot assay and the soluble HLA binding assay (100% at the high- and medium-affinity binders and 85% for the low affinity binders) suggests that the soluble assay may be more precise with regard to higher affinity peptides as compared to the cell-based assay which had lower concordance with the ELISpot assay (high, medium, and low affinity binders having 85%, 83% and 72% respectively). However, at the lowest affinities, the cell-based assay results are more compatible with previously published confirmations of the peptides as B7 epitopes, suggesting that the cell-based assay may be more sensitive at measuring lower affinity peptide-HLA interactions. Alternatively, the cell-based assay allows further peptide processing identifying binding and T cell recognition of processed epitope.

Of the 24 previously published epitopes, 15 previously described have also been confirmed to be restricted by HLA alleles other than B7. We observed this phenomenon in our 2003 paper that first described this approach to HIV Vaccine design. Epitopes that are versatile enough

share HLA restriction may be of particular relevance for vaccine design as a greater portion of the population can be covered with fewer epitopes.

Several of the subject responses to the peptides described in this study could have occurred in the context of other HLA alleles. For example, ENV-1137 IPRRIRQGL and VPR-1125 FPRPWLHGL could have been presented in the context of HLA-A2, as has been reported by other laboratories [39]. However, these peptides had very high HLA-B7 binding affinity, and were recognized by subjects who had HLA-B7 but did not have HLA-A2, therefore, presentation by HLA-B7 is implicated. Likewise, peptide VPR-1125 FPRPWLHGL has been shown by others to be restricted by HLA A*0201 and A*0204. Two subjects, H-945 and H-962 responded with 222 and 1015 SFC/10⁶, respectively. Because B7 was the only allele these two subjects had in common, it is likely that the epitope was presented in the context of HLA-B7. GAG-1144 NPPIPVGEL has been shown to be restricted by HLA-B8 but was recognized by a subject (H-0980) that did not possess the HLA-B8 allele; we have observed similarities in binding affinities and presentation in the context of HLA-B7/B8 previously [63]. The promiscuous epitope NEF-1124 FPVRPQVPL has been defined as being restricted by a number of HLA alleles in addition to HLA-B7 [38]. Only two (H-0955 and H-0980) of the five subjects responding to NEF-1124 possessed an allele for which the peptide was previously described as restricted. The response by subject H0980 is particularly robust (1230 SFC/10⁶), possibly owing to presentation by both B7 and A29. Peptides NEF-1127 YPLTFGWCF and POL-1130 VPRRKAKII were previously described as restricted by A*2402 and B*4201, respectively. The responding subjects did not possess the alleles in question, implicating presentation by B7.

Thus, many of the peptides selected for this study appear to be restricted by multiple HLA. The binding versatility of these peptides is supported by subsequent analysis by “ClustiMer” [16]. Future HIV genome analyses will apply Conservatrix and EpiMatrix (with the clustering function ClustiMer) to select peptides for supertype and/or promiscuous characteristics. Epitopes identified using this method are highly conserved in isolates derived from a wide range of countries. Since the Conservatrix analysis was performed independently of any starting sequence (whether clade B, or other subtype), the novel epitopes discovered in this study appear to be conserved in isolates from many different countries, across continents, and appear to be independent of clade or subtype grouping. Conservation of the novel these novel epitopes across countries is illustrated in Fig. 1.

It is entirely possible that this analysis of HIV sequences has uncovered regions of HIV-1 that are essential, in some way, to the survival of the virus (the Achilles' heel, or heels, of the HIV-1 genome). For example, these regions may be relevant for binding to cellular receptors, to the function of certain proteins, or may be related to the three-dimensional configuration of one of the virus' proteins. The degree of conservation for each of the epitopes would expand markedly if the amino acids were allowed to vary at “non-anchor” positions for B7. Indeed, this was the case for the HIV A2 epitopes selected by the same methods [63]. Any given peptide in the B7 set covers a minimum of 5% and a maximum of 93% of isolates; the average peptide covers 42% of isolates.

In this study, we have not evaluated responses to the peptides using blood samples obtained from individuals known to be infected with other clades of HIV-1. However, there is (1) evidence from our conservation analysis that these epitopes are conserved in a wide range of isolates from Africa, Asia, South America, and the Pacific and (2) additional proof comes from the literature—many of the published epitopes were defined by other authors in far-flung regions of the world.

Several aspects of this study deserve additional scrutiny. For example, all of the study subjects were recruited in the Northeastern United States (Rhode Island and Massachusetts). As their HIV-1 isolates were not sequenced, it would be impossible to determine whether the subjects were infected with Clade B or non Clade B isolates of HIV. Additional studies of these peptides are currently underway in Mali where the circulating HIV strains have greater diversity than those in the USA. Furthermore, whether or not the sequences of the HIV virus species infecting the subjects corresponded exactly to the epitopes selected for this study is unknown. However, the data presented here are consistent with the hypothesis that the responses observed are due to T memory cells responding to epitopes that are processed and presented in the course of natural HIV infection.

In other studies, Rowland Jones, Walker, and other colleagues have evaluated highly conserved *peptide* epitopes (instead of whole genes cloned into vaccinia) for cross-clade T cell recognition (using cells from subjects infected with various clades of HIV) [58,59,66]. These studies provide exciting proof that non-identical sequences representing non-clade B versions of an epitope can stimulate a single T-cell clone. Thus, the promiscuous nature of TCR interaction with MHC-epitope complexes will further expand the potential for “cross-clade” recognition of conserved epitopes. Accordingly, our estimates of cross-clade conservation (which are based on absolute conservation of every amino acid in the sequence, may significantly underestimate the number of isolates to which T cells recognizing these epitopes may be able to cross-react.

Some caveats are worth mentioning when reviewing this data. For example, CD8+ or CD4+ depletion was not performed prior to ELISpot assays, and although confirmed binding to HLA-B7 was observed, we cannot rule out the possibility that the peptides are binding to HLA Class II molecules. However, for most of the selected peptides, responding subjects were only matched for the HLA-B7 allele. Since it is unlikely that the responding subjects happened to be matched at their class II alleles, the ELISpot responses observed here are likely be due to CD8+ restricted responses.

Despite the disappointing results of several recent “T cell-directed” vaccines [64], we believe that T cell-directed vaccines will work if properly constructed. Several laboratories have published data supporting the hypothesis that protective immune response to a number of pathogens requires the development of broad T cell responses to an ensemble of different epitopes [65,66]. Following exposure to a pathogen, epitope-specific memory T cell clones are established [67]. These clones respond rapidly and efficiently upon any subsequent infection, secreting cytokines, killing infected host cells, and marshalling other cellular defenses against the pathogen. This link between broad epitope response and protection from disease has been confirmed for HIV, hepatitis B (HBV), hepatitis C (HCV), LCMV, and malaria [68,69].

Studies of acutely infected subjects and long-term non-progressors (LTNP) have now convincingly shown that broad CTL and Th responses to both dominant and subdominant epitopes, restricted by multiple HLA alleles, are associated with better control of HIV-1 infection [57] accentuating the importance of broadly conserved CTL epitopes. Furthermore, many of the protective T cells in this broad repertoire may target subdominant epitopes. Using epitope-mapping tools such as those described in this study, it is now possible to map entire genome databases for the ensemble of epitopes that may induce a protective immune response to HIV. This study provides an illustration of the bioinformatics approach to selecting broadly conserved T-cell epitopes for HIV-1 vaccine development. The epitopes described in this report are to be included in a DNA-vectored, peptide-boost vaccine that is similar in concept but different in approach from other epitope-based HIV vaccines [12,20,24].

Acknowledgements

The authors gratefully acknowledge Ken Mayer and Michelle Lally for providing the donor subjects for the study, to Miriam Goldberg who provided assistance with the development of Aggregatrix and to Betty Bishop who performed the ELISpot assays. We are collectively grateful to the HIV-infected study subjects who were willing to provide their blood samples for this study and hope that this publication will accelerate the search for vaccine or a cure for HIV/AIDS.

Funding sources: the HIV vaccine research program was initiated using funds provided by a local Rhode Island “seed fund” (Slater Biotechnology Phase I Award) followed by funding from the following NIH grants: R43 AI 46212, R21 AI 45416, R03 TW006306. This particular work was supported by NIH R01 AI050528.

References

1. Graham BS. Clinical trials of HIV vaccines. *Annu Rev Med* 2002;53:207–21. [PubMed: 11818471]
2. Ho DD, Huang Y. The HIV-1 vaccine race. *Cell* 2002;110(2):135–8. [PubMed: 12150921]
3. McCarthy M. HIV vaccine fails in phase 3 trial. *Lancet* 2003;361(9359):755–6. [PubMed: 12620743]
4. Cohen J. Public health. AIDS vaccine trial produces disappointment and confusion. *Science* 2003;299:1290–1. [PubMed: 12610260]
5. Hoag H. Mixed results win HIV vaccine a guarded response. *Nature* 2003;421:877. [PubMed: 12606957]
6. Vaccination and enrollment are discontinued in phase II trials of Merck's investigational HIV vaccine candidate. http://www.merck.com/newsroom/press_releases/.
7. Kresge, JK. What next? As data analysis for the STEP trial gets underway, some other trials are placed in a temporary holding pattern. IAVI Report. <http://www.iavireport.org/>
8. Catanzaro AT, Koup RA, Roederer M, Bailer RT, Enama ME, Moodie Z, et al. Phase 1 safety and immunogenicity evaluation of a multiclade HIV-1 candidate vaccine delivered by a replication-defective recombinant adenovirus vector. *J Infect Dis* 2006;194(12):1638–49. [PubMed: 17109335]
9. Dubey S, Clair J, Fu TM, Guan L, Long R, Mogg R, et al. Detection of HIV vaccine-induced cell-mediated immunity in HIV-seronegative clinical trial participants using an optimized and validated enzyme-linked immunospot assay. *J Acquir Immune Defic Syndr* 2007;45(1):20–7. [PubMed: 17310936]
10. Letvin NL. Strategies for an HIV vaccine. *J Clin Invest* 2002;110(1):15–20. [PubMed: 12093882]
11. Borrow P, Evans CF, Oldstone MB. Virus-induced immuno-suppression: immune system-mediated destruction of virus-infected dendritic cells results in generalized immune suppression. *J Virol* 1995;69(2):1059–70. [PubMed: 7815484]
12. Musey L, Hughes J, Schacker T, Shea T, Corey L, McElrath MJ. Cytotoxic-T-cell responses, viral load, and disease progression in early human immunodeficiency virus type 1 infection. *N Engl J Med* 1997;337(18):1267–74. [PubMed: 9345075]
13. Rosenberg ES, Billingsley JM, Caliendo AM, Boswell SL, Sax PE, Kalams SA, et al. Vigorous HIV-1-specific CD4+ T cell responses associated with control of viremia. *Science* 1997;278:1447–50. [PubMed: 9367954]
14. Schmitz JE, Kuroda MJ, Santra S, Sasseville VG, Simon MA, Lifton MA, et al. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science* 1999;283:857–60. [PubMed: 9933172]
15. Goulder PJ, Brander C, Tang Y, Tremblay C, Colbert RA, Addo MM, et al. Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* 2001;412:334–8. [PubMed: 11460164]
16. Shedlock DJ, Shen H. Requirement for CD4 T cell help in generating functional CD8 T cell memory. *Science* 2003;300:337–9. [PubMed: 12690201]
17. Koup RA, Safrin JT, Cao Y, Andrews CA, McLeod G, Borkowsky W, et al. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J Virol* 1994;68(7):4650–5. [PubMed: 8207839]
18. Langford SE, Ananworanich J, Cooper DA. Predictors of disease progression in HIV infection: a review. *AIDS Res Ther* 2007;4:11. [PubMed: 17502001]

19. De Groot AS, Bosma A, Chinai N, Frost J, Jesdale BM, Gonzalez MA, et al. From genome to vaccine: in silico predictions, ex vivo verification. *Vaccine* 2001;19(31):4385–95. [PubMed: 11483263]
20. Sbai H, Mehta A, DeGroot AS. Use of T cell epitopes for vaccine development. *Curr Drug Targets Infect Disord* 2001;1(3):303–13. [PubMed: 12455403]
21. Allen TM, Altfeld M, Geer SC, Kalife ET, Moore C, O'sullivan KM, et al. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. *J Virol* 2005;79(21):13239–49. [PubMed: 16227247]
22. Peyerl FW, Bazick HS, Newberg MH, Barouch DH, Sodroski J, Letvin NL. Fitness costs limit viral escape from cytotoxic T lymphocytes at a structurally constrained epitope. *J Virol* 2004;78(24):13901–10. [PubMed: 15564498]
23. Frahm N, Korber BT, Adams CM, Szinger JJ, Draenert R, Addo MM, et al. Brander C Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. *J Virol* 2004;78(5):2187–200. [PubMed: 14963115]
24. Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* 2002;296:1439–43. [PubMed: 12029127]
25. Zhang C, Cornette JL, Berzofsky JA, DeLisi C. The organization of human leucocyte antigen class I epitopes in HIV genome products: implications for HIV evolution and vaccine design. *Vaccine* 1997;15(1213):1291–302. [PubMed: 9302734]
26. Sette A, Sidney J. HLA supertypes and supermotifs: a functional perspective on HLA polymorphism. *Curr Opin Immunol* 1998;10(4):478–82. [PubMed: 9722926]
27. Trachtenberg E, Korber B, Sollars C, Kepler TB, Hraber PT, Hayes E, et al. Advantage of rare HLA supertype in HIV disease progression. *Nat Med* 2003;9(7):928–35. [PubMed: 12819779]
28. Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature* 2004;432:769–75. [PubMed: 15592417]
29. De Groot AS, Jesdale BM, Szu E, Schafer JR. An interactive web site providing MHC ligand predictions: application to HIV research. *AIDS Res Human Retroviruses* 1997;13:539–41.
30. Schafer JA, Jesdale BM, George JA, Kouttab NM, De Groot AS. Prediction of well-conserved HIV-1 ligands using a matrix-based algorithm, EpiMatrix. *Vaccine* 1998;16(19):1880–4. [PubMed: 9795396]
31. Iversen AK, Stewart-Jones G, Learn GH, Christie N, Sylvester-Hviid C, Armitage AE, et al. Conflicting selective forces affect T cell receptor contacts in an immunodominant human immunodeficiency virus epitope. *Nat Immunol* 2006;7(2):179–89. [PubMed: 16388312]
32. Nguyen L, Chaowanachan T, Vanichseni S, McNicholl JM, Mock PA, Nelson R, et al. Frequent human leukocyte antigen class I alleles are associated with higher viral load among HIV type 1 seroconverters in Thailand. *J Acquir Immune Defic Syndr* 2004;37(2):1318–23. [PubMed: 15385741]
33. Sylvester-Hvid C, Kristensen N, Blicher T, Ferre H, Lauemoller SL, Wolf XA, et al. Establishment of a quantitative ELISA capable of determining peptide—MHC class I interaction. *Tissue Antigens* 2002;59(4):251–8. [PubMed: 12135423]
34. Kessler JH, Mommaas B, Mutis T, Huijbers I, Vissers D, Benckhuijsen WE, et al. Competition-based cellular peptide binding assays for 13 prevalent HLA class I alleles using fluorescein-labeled synthetic peptides. *Human Immunol* 2003;64:245–55. [PubMed: 12559627]
35. Jassoy C, Johnson RP, Navia BA, Worth J, Walker BD. Detection of a vigorous HIV-1-specific cytotoxic T lymphocyte response in cerebrospinal fluid from infected persons with AIDS dementia complex. *J Immunol* 1992;149(9):3113–9. [PubMed: 1383338]
36. McAdam S, Kaleebu P, Krausa P, Goulder P, French N, Collin B, et al. Cross-clade recognition of p55 by cytotoxic T lymphocytes in HIV-1 infection. *AIDS* 1998;12(6):571–9. [PubMed: 9583596]
37. Altfeld M, Allen TM, Yu XG, Johnston MN, Agrawal D, Korber BT, et al. HIV-1 superinfection despite broad CD8+ T-cell responses containing replication of the primary virus. *Nature* 2002;420:434–9. [PubMed: 12459786]

38. Novitsky V, Cao H, Rybak N, Gilbert P, McLane MF, Gaolekwe S, et al. Magnitude and frequency of cytotoxic T-lymphocyte responses: identification of immunodominant regions of human immunodeficiency virus type 1 subtype C. *J Virol* 2002;76(20):10155–68. [PubMed: 12239290]
39. Price P, Johnson RP, Scadden DT, Jassoy C, Rosenthal T, Kalams S, et al. Cytotoxic CD8+ T lymphocytes reactive with human immunodeficiency virus-1 produce granulocyte/macrophage colony-stimulating factor and variable amounts of interleukins 2, 3, and 4 following stimulation with the cognate epitope. *Clin Immunol Immunopathol* 1995;74(1):100–6. [PubMed: 7527747]
40. Li F, Malhotra U, Gilbert PB, Hawkins NR, Duerr AC, McElrath JM, et al. Peptide selection for human immunodeficiency virus type 1 CTL-based vaccine evaluation. *Vaccine* 2006;24(4748): 6893–904. [PubMed: 16890329]
41. Gahery-Segard H, Pialoux G, Charmeteau B, Sermet S, Poncelet H, Raux M, et al. Multiepitopic B- and T-cell responses induced in humans by a human immunodeficiency virus type 1 lipopeptide vaccine. *J Virol* 2000;74(4):1694–703. [PubMed: 10644339]
42. Musey L, Hu Y, Eckert L, Christensen M, Karchmer T, McElrath MJ. HIV-1 induces cytotoxic T lymphocytes in the cervix of infected women. *J Exp Med* 1997;185(2):293–303. [PubMed: 9016878]
43. Corbet S, Nielsen HV, Vinner L, Lauemoller S, Therrien D, Tang S, et al. Optimization and immune recognition of multiple novel conserved HLA-A2, human immunodeficiency virus type 1-specific CTL epitopes. *J Gen Virol* 2003;84:2409–21. [PubMed: 12917462]
44. Frahm N, Adams S, Kiepiela P, Linde CH, Hewitt HS, Lichterfeld M, et al. HLA-B63 presents HLA-B57/B58-restricted cytotoxic T-lymphocyte epitopes and is associated with low human immunodeficiency virus load. *J Virol* 2005;79(16):10218–25. [PubMed: 16051815]
45. Frahm, N.; Linde, C.; Brander, C. Identification of HIV-derived, HLA class I restricted CTL epitopes: insights into TCR repertoire, CTL escape and viral fitness. In: Korber, BTM.; Brander, C.; Haynes, BF.; Koup, R.; Moore, JP.; Walker, BD.; Watkins, DI., editors. *HIV molecular immunology*. Los Alamos, New Mexico: Los Alamos National Laboratory, Theoretical Biology and Biophysics; 20062007. p. 3-28.
46. Maksiutov AZ, Bachinskii AG, Bazhan SI. Searching for local similarities between HIV-1 and human proteins: application to vaccines. *Mol Biol (Mosk)* 2002;36(3):447–59. [PubMed: 12068630]
47. Wilson CC, Palmer B, Southwood S, Sidney J, Higashimoto Y, Appella E, et al. Identification and antigenicity of broadly cross-reactive and conserved human immunodeficiency virus type 1-derived helper T-lymphocyte epitopes. *J Virol* 2001;75(9):4195–207. [PubMed: 11287569]
48. Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med* 2007;13(1):46–53. [PubMed: 17173051]
49. Bailey JR, Williams TM, Siliciano RF, Blankson JN. Maintenance of viral suppression in HIV-1-infected HLA-B*57+ elite suppressors despite CTL escape mutations. *J Exp Med* 2006;203(5):1357–69. [PubMed: 16682496]
50. Addo MM, Yu XG, Rathod A, Cohen D, Eldridge RL, Strick D, et al. Comprehensive epitope analysis of human immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. *J Virol* 2003;77(3):2081–92. [PubMed: 12525643]
51. Lieberman J, Fabry JA, Shankar P, Beckett L, Skolnik PR. Ex vivo expansion of HIV type 1-specific cytolytic T cells from HIV type 1-seropositive subjects. *AIDS Res Hum Retroviruses* 1995;11(2): 257–71. [PubMed: 7742040]
52. Jassoy C, Harrer T, Rosenthal T, Navia BA, Worth J, Johnson RP, et al. Human immunodeficiency virus type 1-specific cytotoxic T lymphocytes release gamma interferon, tumor necrosis factor alpha (TNF-alpha), and TNF-beta when they encounter their target antigens. *J Virol* 1993;67(5):2844–52. [PubMed: 7682629]
53. Chang KM, Rehmann B, McHutchison JG, Pasquinelli C, Southwood S, Sette A, et al. Immunological significance of cytotoxic T lymphocyte epitope variants in patients chronically infected by the hepatitis C virus. *J Clin Invest* 1997;100(9):2376–85. [PubMed: 9410918]
54. Meister GE, Roberts CG, Berzofsky JA, De Groot AS. Two novel T cell epitope prediction algorithms based on MHC-binding motifs; comparison of predicted and published epitopes from *Mycobacterium tuberculosis* and HIV protein sequences. *Vaccine* 1995;13(6):581–91. [PubMed: 7483779]

55. Bond KB, Sriwanthana B, Hodge TW, De Groot AS, Mastro TD, Young NL, et al. An HLA-directed molecular and bioinformatics approach identifies new HLA-A11 HIV-1 subtype E cytotoxic T lymphocyte epitopes in HIV-1-infected Thais. *AIDS Res Hum Retroviruses* 2001;17(8):703–17. [PubMed: 11429111]
56. Dong Y, Demaria S, Sun X, Santori FR, Jesdale BM, De Groot AS, et al. HLA-A2-restricted CD8+-cytotoxic-T-cell responses to novel epitopes in mycobacterium tuberculosis superoxide dismutase, alanine dehydrogenase, and glutamine synthetase. *Infect Immun* 2004;72(4):2412–5. [PubMed: 15039371]
57. Smith DJ, King WF, Barnes LA, Peacock Z, Taubman MA. Immunogenicity and protective immunity induced by synthetic peptides associated with putative immunodominant regions of *Streptococcus mutans* glucan-binding protein B. *Infect Immun* 2003;71(3):1179–84. [PubMed: 12595430]
58. McMurry JA, Gregory SH, Moise L, Rivera D, Buus S, De Groot AS. Diversity of *Francisella tularensis* Schu4 antigens recognized by T lymphocytes after natural infections in humans: identification of candidate epitopes for inclusion in a rationally designed tularemia vaccine. *Vaccine* 2007;25(16):3179–91. [PubMed: 17291638]
59. De Groot AS, McMurry J, Marcon L, Franco J, Rivera D, Kutzler M, et al. Developing an epitope-driven tuberculosis (TB) vaccine. *Vaccine* 2005;23(1718):2121–31. [PubMed: 15755582]
60. De Groot AS, Bishop EA, Khan B, Lally M, Marcon L, Franco J, et al. Engineering immunogenic consensus T helper epitopes for a cross-clade HIV vaccine. *Methods* 2004;34(4):476–87. [PubMed: 15542374]
61. Leslie A, Kavanagh D, Honeyborne I, Pfafferoth K, Edwards C, Pillay T, et al. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. *J Exp Med* 2005;201(6):891–902. [PubMed: 15781581]
62. Altfeld M, Addo MM, Shankarappa R, Lee PK, Allen TM, Yu XG, et al. Enhanced detection of human immunodeficiency virus type 1-specific T-cell responses to highly variable regions by using peptides based on autologous virus sequences. *J Virol* 2003;77(13):7330–40. [PubMed: 12805431]
63. De Groot AS, Jesdale B, Martin W, Saint Aubin C, Sbai H, Bosma A, et al. Mapping cross-clade HIV-1 vaccine epitopes using a bioinformatics approach. *Vaccine* 2003;21(27–30):4486–504. [PubMed: 14505932]
64. http://www.kaisernetwork.org/daily_reports/rep_index.cfm?DR_ID=47688.
65. Gianfrani C, Oseroff C, Sidney J, Chesnut RW, Sette A. Human memory CTL response specific for influenza A virus is broad and multispecific. *Hum Immunol* 2000;61:438–52. [PubMed: 10773346]
66. Gillespie GM, Wills MR, Appay V, O'Callaghan C, Murphy M, Smith N, et al. Functional heterogeneity and high frequencies of cytomegalovirus-specific CD8 (+) T lymphocytes in healthy seropositive donors. *J Virol* 2000;74:8140–50. [PubMed: 10933725]
67. Blattman JN, Sourdive DJ, Murali-Krishna K, Ahmed R, Altman JD. Evolution of the T cell repertoire during primary, memory, and recall responses to viral infection. *J Immunol* 2000;165(11):6081–90. [PubMed: 11086040]
68. Harrer T, Harrer E, Kalams SA, Barbosa P, Trocha A, Johnson RP, et al. Abstract Cytotoxic T lymphocytes in asymptomatic long-term nonprogressing HIV-1 infection. Breadth and specificity of the response and relation to in vivo viral quasispecies in a person with prolonged infection and low viral load. *J Immunol* 1996;156(7):2616–23. [PubMed: 8786327]
69. Doolan DL, Hoffman SL, Southwood S, Wentworth PA, Sidney J, Chesnut RW, et al. Degenerate cytotoxic T cell epitopes from *P. falciparum* restricted by multiple HLA-A and HLA-B supertype alleles. *Immunity* 1997;7(1):97–112. [PubMed: 9252123]

Figure 1.

GAIA HIV B7 peptides—individually and in aggregate—percent coverage of strains by year, country, and clade. Each row of the matrix denotes a specific peptide; the peptide's protein of origin is included within the peptide ID. Each column of the matrix denotes a specific year, country, or clade, grouped as indicated. The percentage coverage of strains is represented on a color gradient, with warm tones indicating values above 50% and cool tones indicating values below 50%. The bottom row of the matrix shows the percentage of each protein set that contains at least one peptide from our pool. Black boxes indicate that no isolates of the protein are available for that year, clade, or country. The bottom row represents the aggregate percent coverage for the epitope set. Each cell of the matrix represents the percent coverage per peptide, except for the bottom-row cells, which represent the aggregate percent coverage for the peptide set. Column headers are listed here for space considerations: left to right, the year columns are 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, and 2005; aggregate coverage of strains by year ranges from 51% (2005) to 65% (2003). The countries left to right are: Angola, Argentina, Australia, Belgium, Benin, Bolivia, Brazil, Botswana, Belarus, Canada, The democratic republic of the Congo (Zaire), Congo, Cote d'Ivoire, Chile, Cameroon, China, Colombia, Cuba, Cyprus, Germany, Djibouti, Dominican Republic, Ecuador, Estonia, Spain, Ethiopia, France, Gabon, United Kingdom, Georgia, Ghana, Gambia, Equatorial Guinea, Greece, Hong Kong, Israel, India, Italy, Japan, Kenya, Republic of Korea, Mali, Myanmar, Namibia, Niger, Nigeria, Netherlands, Norway, Portugal, Russian federation, Rwanda, Sweden, Senegal, Somalia, Chad, Thailand, Trinidad and Tobago, Taiwan, United Republic of Tanzania, Ukraine, Uganda, United States, Uruguay, Uzbekistan, Venezuela, Vietnam, Yemen, South Africa, Zambia, and Zimbabwe; aggregate coverage of strains by country ranges from 33% (Gabon) to 94% (Italy). The clade columns left to right are: 0206, 0209, 1819, A, AAD, AB, AC, ACD, ACDGKU, AD, ADF, ADGU, ADHK, AE, AF, AFG, AFU, AG, AGH,

AGHU, AGJ, AGU, AHJU, AU, B, BC, BF, BG, C, CD, CGU, CK, CPX, CU, D, DF, DK, DO, DU, F, G, GHJKU, GK, GKU, GU, H, J, JKU, JU, K, N, O, U. The HLA-B7 peptides together cover 85, 78, 78, and 80% of clades A, B, C, and D, respectively. On average, the chimeric clade sequences were 76% covered by our HLA-B7 restricted epitopes.

Table 1
Overview of HLA-B7 epitopes mapped for the GAIA HIV-Vax program

HIV-VAX-B7		Sequence	Conservation (%)	First year isolated	Countries covered	Main clade	Subjects responding by ELISpot	Previously published? (for B7) *prior to selection
Peptide	ID							
ENV	1161	IPIHYCAPA	75	1976	59	B	1/7	Y (B7) prior
ENV	ENV 1146	RPNNNTRKSI	30	1981	38	B	2/7	Y (B7) prior
ENV	1137	IPRRIRQGL	39	1976	44	B	5/7	Y (B7) prior
ENV	1129	VPTDPNPQEI	25	1983	48	C	1/7	N
ENV	1131	SPLSFQTRL	22	1981	14	B	1/7	N
ENV	1148	TPLCVTLNCT	51	1980	44	B	2/7	N
ENV	1158	GPCTNVSTV	35	1981	31	B	2/7	N
ENV	1168	FDITNWLWYI	18	1980	33	B	0/7	(N)
GAG	1144	NPPIPVGEI	44	1976	51	B	1/7	Y prior
GAG	1166	IPMFSAISEG	57	1976	51	B	3/7	Y prior
GAG	1164	GPKEPFRDY	93	1976	62	B	5/7	Y prior
GAG	1156	HPVHAGPVA	14	1983	23	C	2/7	Y (B7) prior
GAG	1150	GPSHKARVL	35	1976	44	C	5/7	Y
GAG	1162	EPTAPPAESF	41	1976	47	C	1/7	N
GAG	1147	APPAESFRF	21	1986	27	C	2/7	N
GAG	1159	RPEPTAPPA	48	1976	47	C	2/7	N
GAG	1149	IPVGDIIYKRW	26	1980	33	C	0/7	(N)
NEF	1163	VPLRPMITYKA	33	1983	40	B	4/7	Y prior
NEF	1133	TPGPGIRYPL	34	1981	39	B	5/7	Y prior
NEF	1124	FPVRPQVPL	76	1976	50	B	5/7	Y (B7) prior
NEF	1127	YPLTFGWCF	45	1981	38	B	1/7	Y
NEF	1139	GPCTRFPLTF	18	1983	31	B	5/7	Y
POL	1145	LPEKDSWTV	58	1983	43	C	1/7	Y prior
POL	1130	LPPIVAKKI	42	1976	33	C	2/7	Y(B7)
POL	1167	IPHPAGLKK	87	1981	47	C	1/7	Y
POL	1152	FVNTPLVKL	83	1981	46	C	2/7	Y
POL	1132	VPRRKAKII	64	1976	46	B	3/7	Y
POL	1151	TPVNIIGRNL	38	1983	29	B	3/7	Y
POL	1153	YPGIKVKQL	21	1976	29	B	1/7	N
POL	1157	TPGIRYQYNV	85	1976	46	C	1/7	N
POL	1165	IPSTNNETPG	21	1983	35	A1	1/7	N
POL	1128	IPYNPQSQGV	89	1976	47	B	2/7	N
POL	1143	DPIWKGPAKL	54	1976	36	C	2/7	N
POL	1134	LPGRWKPKMI	26	1983	26	B	4/7	N
POL	1142	VPVKLKPGM	87	1976	45	C	0/7	(Y)
POL	1141	EPFRKQNPDI	28	1983	22	B	0/7	(Y prior)
POL	1126	FPISPIETV	83	1976	49	C	0/7	(N)
TAT	1140	EPVDPNLEPW	19	1976	23	C	2/7	Y
TAT	1160	QPKTACTNCY	14	1981	21	B	1/7	N
VIF	1136	HPRISEVHI	21	1981	26	B	5/7	Y prior
VIF	1138	TPKKIKPPL	24	1981	19	B	1/7	N
VIF	1154	IPLGEAKLV	5	1983	16	B	1/7	N
VIF	1155	PPLPSVRKL	30	1983	38	B	1/7	N
VIF	1135	IPLGDARLVI	14	1981	21	B	2/7	N
VPR	1125	FPRPWLHGL	18	1983	40	C	3/7	Y prior

Forty-five peptides were selected as putative B7 epitopes for evaluation and incorporation in an HIV vaccine. The peptides have the characteristics of strong conservation and strong putative immunogenicity for B7. The peptide IDs and sequences are shown in columns 1 and 2, respectively. Peptides are grouped by source protein: there were 1550 sequences from ENV, 928 sequences from GAG, 2884 sequences from NEF, 563 sequences from POL, 766 sequences from TAT, 1311 sequences from VIF, and 1299 sequences from VPR. Percent conservation among these respective number of input strains is shown in column 3. Column 4 shows the earliest strain covered by each epitope; several epitopes cover strains isolated as early as the beginning of the epidemic and all of the peptides cover strains *at least* as recently as 2005. Column 5 shows the number of countries for which strains are covered by each of the peptides. Column 6 shows the main clade covered by each peptide. Column 7 summarizes the ELISpot assay results; details can be found in Table 4. Column 8 indicates whether the peptide was previously published and if so, whether the publication was prior to or after the selection of peptides and whether B7 was among the alleles for which the peptide was described as restricted.

Table 2
Comparison of soluble and cell-based HLA-B7 binding assays

HIV-VAX-B7	Peptide	ID	Sequence	Soluble assay		Cell-based assay		ELISPOT positive
				K_d	Predicted affinity	Mean fluorescence	Predicted affinity	
GAG		1156	HPVHAGPVA	Not tested	—	4	High	Y
NEF		1124	FPVRPQVPL	Not tested	—	4	High	Y
VPR		1125	FPRPWLHGL	3	High	3	High	Y
ENV		1137	IPRRIROGL	3	High	24	Medium	Y
VIF		1136	HPRISSEVHI	22	High	4	High	Y
ENV		1131	SPLSFQTRL	29	High	10	High	Y
NEF		1133	TPGPGIRYPL	32	High	4	High	Y
ENV		1146	RPNNTTRKSI	40	High	4	High	Y
GAG		1150	GFSHKARVL	50	High	4	High	Y
POL		1132	VPRRKAKII	55	Medium	4	High	Y
ENV		1161	IPHYCAPA	115	Medium	4	High	Y
VIF		1138	TPKKIKPPL	499	Medium	4	High	Y
VIF		1135	IPLGDARLVI	506	Low	4	High	Y
GAG		1159	RPEPTAPPA	533	Low	7	High	Y
POL		1153	YPGIKVKQL	600	Low	5	High	Y
POL		1126	FPISPIETV	641	Low	5	High	N
POL		1167	IPHPAGLKK	659	Low	5	High	Y
GAG		1162	EPTAPPAESF	1322	Low	14	Medium	Y
NEF		1127	YPLTFGWCF	1409	Low	4	High	Y
NEF		1139	GPGRFPLTF	2017	Low	4	High	Y
POL		1128	IPYNPQSQGV	2819	Low	4	High	Y
TAT		1140	EPVDPNLEPW	2969	Low	28	Low	Y
GAG		1166	IPMFSALESG	3118	Low	7	High	Y
ENV		1158	GPCTNVSTV	3125	Low	16	Medium	Y
POL		1157	TPGIRYQYNV	3131	Low	5	High	Y
POL		1151	TPVNIIGNRL	4046	Low	5	High	Y
POL		1134	LPGRWPKPMI	7334	Low	4	High	Y
POL		1142	VPVKLKPGM	9990	Low	5	High	N
GAG		1147	APPAESFRF	10,303	Low	5	High	Y
VIF		1155	PPLPSVRKL	10,587	Low	15	Medium	Y
POL		1130	LPPIVAKEL	10,928	Low	9	High	Y
VIF		1154	IPLGEAKLV	17,195	Low	11	Medium	Y
POL		1165	IPSTNNETPG	100,000	None	9	High	Y
NEF		1163	VPLRPMITYKA	100,000	None	10	High	Y
GAG		1149	IPVGDYKRW	100,000	None	17	Medium	N
POL		1145	LPEKDSWT	100,000	None	30	Low	Y
GAG		1144	NPPIPVGEI	100,000	None	32	Low	Y
POL		1141	EPFRKQNPDI	100,000	None	37	Low	N
POL		1152	FVNTPLVKL	100,000	None	38	Low	Y
TAT		1160	QPKTACTNCY	100,000	None	41	Low	Y
GAG		1164	GPKEPRDY	100,000	None	41	Low	Y
POL		1143	DPIWKGPAKL	100,000	None	43	Low	Y
ENV		1168	FDITNWLWYI	100,000	None	46	Low	N
ENV		1129	VPIDPNPQEI	100,000	None	47	Low	Y
ENV		1148	TPLCVTLNCT	100,000	None	47	Low	Y

The putative B7 epitopes were synthesized and analyzed for binding to HLA-B7 using two different binding assay formats.

Table 3

Donor clinical data at time of draw

	Most recent viral load	Most recent CD4 count	Years infected	On ARV treatment	Number of epitope responses	HLA-A	HLA-B
H_0943	550	522	17	Y	12	A01011, A02011	B07021, B0801
H_0945	870	359	2	N	29	A02011	B07021, B51011
H_0948	82	376	14	Y	5	A020101	B070201, B400101
H_0955	9397	204	1	N	13	A030101, A240201	B070201, B350101
H_0962	<50	472	18	Y	11	A030101, A3201	B070201
H_0976	7888	204	10.8	Y	8	A020101	B070201, B4901
H_0980	516	395	13	N	16	A290101, A680101	B070201, B520101

HIV-VAX-B7 Peptide ID	Sequence	Conservation	First Year Isolated	Countries covered	Main Clade	Available HLA restriction per [LANL] and IEDB																												
						A*01011 A*02011	A*02011	A*030101 A*240201	A*030101 A*3201	A*020101	A*020101	A*020101	A*020101	A*020101	A*020101	A*020101	A*290101 A*680101	First published																
POL 1126	FPISPIETV	83%	1976	49	C	B*07021 B*0801	B*07021 B*51011	-	-	-	-	B*070201 B*400101	B*070201 B*350101	B*070201 B*4901	B*070201 B*520101	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
TAT 1140	EPVDPNLEPW	19%	1976	23	C	-	251	-	-	-	-	-	-	-	157	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
TAT 1160	QPKTACTNCY	14%	1981	21	B	-	43	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
VIF 1136	HPRISSEVHI	21%	1981	26	B	237	605	-	-	-	60	-	-	-	35	-	3,293	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A
VIF 1138	TPKKIKPPL	24%	1981	19	B	-	58	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
VIF 1154	IPLGEAKLV	5%	1983	16	B	-	32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
VIF 1155	PPLPSVRKL	30%	1983	38	B	-	20	-	-	-	-	-	-	-	20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
VIF 1135	IPLGDARLVI	14%	1981	21	B	-	555	-	-	-	39	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
VPR 1125	FPRPWLHGL	18%	1983	40	C	-	1,106	-	-	-	222	-	-	-	-	-	1,015	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	[A*0204, A*0201]

The peptide IDs and sequences are shown in columns 1 and 2, respectively. The ID number for the PBMC donor is shown at the top of columns 3–9, directly under which, the corresponding HLA A and HLA B alleles for the subjects are noted. For simplicity the table shows only the ELISpot responses that meet a cutoff of at least twice the number of background spots and at least 20 spots per million above background. Forty of the 45 B7 were positive by ELISpot in at least one subject. Some epitopes have since been published by other groups (column 10), but many are novel. Column 11 shows the HLA restriction as determined from the references listed in LANL [in brackets] and as listed on the Immune Epitope Database (underlined).