LA-UR- 09-05921

Approved for public release; distribution is unlimited.

Title:	Inability to Induce Consistent T-Cell Responses Recognizing Conserved Regions Within HIV-1 Antigens: a Potential Mechanism for Lack of Vaccine Efficacy in the Step Study
Author(s):	Bette Korber, Z # 108817, T-6/ T-Division James Szinger, Z#183317, T-6/T-Division
Intended for:	Journal of Infectious Diseases

Form 836 (7/06)

Inability to Induce Consistent T-Cell Responses Recognizing Conserved Regions Within HIV-1 Antigens: a Potential Mechanism for Lack of Vaccine Efficacy in the Step Study

Fusheng Li[1], Sheri A. Dubey[2], Bette T. M. Korber[7,8], James Szinger[8], Michael N. Robertson[2], M. Juliana McElrath[1,3], John W. Shiver[2], Danilo R. Casimiro[2], Lawrence Corey[1,3,5,6], Steven G. Self[1,4]

[1]Vaccine and Infectious Disease Institute, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; [2]Merck Research Laboratories, North Wales, PA 19454, USA; Departments of [3]Medicine, [4]Biostatistics, [5]Laboratory Medicine, [6]Epidemiology, University of Washington, Seattle, WA 98195, USA; [7]Santa Fe Institute, Santa Fe, NM 87501; [8]Los Alamos Laboratory, Los Alamos, NM 87545, USA

ABSTRACT

T cell based vaccines are based upon the induction of CD8+ T cell memory responses that would be effective in inhibiting infection and subsequent replication of an infecting HIV-1 strain, a process that requires a high probability of matching the epitope induced by vaccination with the infecting viral strain. We compared the frequency and specificity of the CTL epitopes elicited by the replication defective Ad5 gag/pol/nef vaccine used in the STEP trial with the likelihood of encountering those epitopes among recently sequenced Clade B isolates of HIV-1. On average vaccination elicited only one epitope per gene. Importantly, the highly conserved epitopes in gag, pol, and nef (> 80% of strains in the current collection of the Los Alamos database [www.hiv.lanl.gov]) were rarely elicited by vaccination. Moreover there was a statistically significant skewing of the T cell response to relative variable epitopes of each gene; only 20% of persons possessed > 3 T cell responses to epitopes likely to be found in circulating strains in the CladeB populations in which the Step trial was conducted. This inability to elicit T cell responses likely to be found in circulating viral strains is a likely factor in the lack of efficacy of the vaccine utilized in the STEP trial. Modeling of the epitope specific responses elicited by vaccination, we project that a median of 8-10 CD8+ T cell epitopes are required to provide >80% likelihood of eliciting at least 3 CD8+ T cell epitopes that would be found on a circulating population of viruses.. Development of vaccine regimens which elicit either a greater breadth of responses or elicit responses to conserved regions of the HIV-1 genome are needed to fully evaluate the concept of whether induction of T cell immunity can alter HIV-1 *in vivo*.

Introduction:

In principle, an effective HIV-1 vaccine designed to elicit antiviral T cell immunity must be able to consistently direct those responses to T cell epitopes likely to be present in the diverse population of circulating viral strains. HIV-1 vaccines must either generate sufficiently broad responses to compensate for extensive sequence variability in HIV-1, or focus the response primarily on highly-conserved epitopes. The HIV-1 genes represented in most candidate HIV vaccines are typically full-length (or nearly full-length) genes derived from natural isolates. These can be based on a natural strain specifically selected to be relatively central to circulating strains in a given population, to maximize the potential cross-reactivity [1], or by selecting natural sequence that maximizes coverage of potential epitopes in a population of sequences [2, 3]. Alternatively, one can use synthetic sequences (e.g., consensus or inferred ancestral sequence) as the basis for antigen design [4-6]. No studies have determined experimentally whether conserved or variable epitopes are preferentially recognized by the human T-cells after vaccination with such immunogens and how this preference is related to the outcome of vaccine trials.

Recently, the results of the STEP trial, a test of concept study of a T cell based vaccine, reported no protection against either HIV-1 acquisition of infection or post-infection control of viremia [7, 8]. The vaccine candidate, made by Merck Research Laboratories, consisted of a mixture of replication defective Adenovirus type 5 constructs containing the HIV-1 gag, pol and nef genes which were inserted into the E-1 region of the Ad5 backbone. Gag and pol were selected because of their relatively high

conservation both interclade and intraclade among all HIV gene products. Specifically, the *gag* coding sequence was derived from the CAM-1 strain of HIV-1 (GenBank Locus BAA00992)[9], because its *gag* gene most closely resembled the clade B consensus amino acid sequence. The *pol* construct included only the reverse transcriptase and integrase gene products from IIIb, and *nef* from JRFL strains. All three transgenes were "optimized" by using frequently utilized codons to improve protein expression in mammalian cells [9-11]. The *pol* transgene segment was inactivated by substituting alanine codons for amino acids at enzymatically active sites [10, 12-18], and the *nef* transgene segment was inactivated through substitutions, which prevents attachment to the cytoplasmic membrane and retrotrafficking into endosomes[10, 19]. The vaccine was formulated as a 1:1:1 mixture of each gene product and 3 separate doses of 10^{10} viral particles of the mixture were administered. The vaccine was immunogenic in most individuals: >80% of subjects who received the vaccine in the Step trial had detectable HIV-specific CD8+ T lymphocytes that were capable of producing multiple cytokines, including IFN-γ and TNF-α; these responses were long-lasting, persisting for more than one year[8]. Despite these immune responses, no reduction in acquisition or effect on reducing post acquisition viral load was seen. The reasons behind the lack of efficacy despite being highly immunogenic remain unclear and are central to the future development of an effective HIV vaccine. As such we undertook an analysis of whether the T-cell epitope specificities elicited by vaccination were within conserved or variable regions of the HIV proteins and whether the vaccine elicited the type of T cell coverage likely to be effective after encountering the diversity of HIV-1 circulating in the regions where the vaccine trial was conducted.

We analyzed the epitope mapping data on 72 subjects who participated in an earlier Phase I trial of the same vaccine utilized in the STEP study. 43 were Ad5 seronegative and 29 Ad5 seropositive at enrollment. Their median age was 35; 61 were Caucasian; 68% were male and 32% female. Details of the T cell and antibody responses to the vaccine among these subjects have been described [10]; their responses to vaccination were similar if not higher in magnitude to those enrolled in the STEP trial. Moreover, the subjects we analyzed were enrolled from many US study sites in which the ensuing STEP trial was conducted. The incidence of HIV-1 acquisition in STEP was highest in the U.S. As such, the epitope mapping data from these samples from the earlier Phase I trial were of relevance to immune responses seen in STEP.

Class I HLA typing was performed on all subjects [**Sheri, ref for method**] Using an IFN-γ ELISpot assay, epitopes recognized among T-cells from positive 15mer peptide pool responders were first mapped to a 16 amino acid region with peptide mini-pools, each containing eight 9-mer peptides that matched the vaccine inserts. Positive mini-pools were deconvoluted to identify the individual 9-mer CTL epitopes, where sample volumes were sufficient to do so. **[Sheri, how many did not have sufficient volume to deconvolute and how were they handled?]** We analyzed the frequencies by which the identified epitopes are induced among the vaccine recipients and the frequencies by which they are found among sequences of isolates from clade B regions annotated in the Los Alamos National Laboratory (LANL) HIV-1 sequence database. For comparison, theoretical enumerations of all the potential T cell epitopes ("PTEs") and previously identified CTL epitopes ("known epitopes") present in the immunogens were conducted. In the first approach, the vaccine immunogen is decomposed into its constitutive 9-mer

PTEs each of which is characterized by its frequency of occurrence among circulating viral strains[20, 21]. The PTE measurement provides characterization of strain variability that is more directly relevant to T-cell responses than other metrics for protein variation that do not explicitly refer to the basic peptide units of T-cell recognition (e.g., protein distance).

From 72 individuals, 285 mini-pools (66 Gag, 56 Nef, and 163 Pol) were found positive by ELISpot assay. On average each vaccine recipient recognized only one minipool/gene. Of the 72 vaccinees evaluated, 58 (81%) had at least one response to any HIV gene but only 29 (40%) had $\geq$4 responses with a median (mean) response breadth of 3 (3.9) epitopes per subject (**Table 1A**). Among all the positive mini-pools, 57 mini-pools did not perform deconvolution study due to either limited sample or relative low T-cell responses (**Supp. Table 1**). Another 61 mini-pools failed to show 9-mer epitopes in deconvolution study. In the rest of the 167 mini-pools, a single 9-mer epitope was identified from each mini-pool most of the time. In rare cases, (can we get a numerator here??) two epitopes, or three epitopes, were determined. Deconvolution study identified a total of 108 unique epitopes from the positive mini-pools, with 31, 21 and 57 in Gag, Nef, and Pol, respectively (**Table 1B**). Surprisingly, most (62 of 109) of the gene-specific CTL responses elicited by the vaccine were directed at epitopes unique to an individual vaccinee; for unique epitopes, 22 of 31 (71%) of the Gag epitopes, 9 of 21 (43%) of the Nef epitopes, and 31 of 57 (54%) of the Pol epitopes mapped after vaccination were seen in only a single individual. Only 4 Gag, 4 Nef, and 3 Pol epitope was recognized by more than 3 of the 72 individuals. The individual diversity and uniqueness of the recognized

epitope were irrespective of the HLA typing, i.e., vaccinated subjects with similar HLA types exhibited markedly different epitope specific responses.

Despite individual variability in the epitope specificity of the vaccine induced T cell response, the location of the epitope responses after vaccination along the protein sequence tracked well with the distribution of responses reported in the literature after initial infection (**Figure 1**). There was a very strong correlation between the number of epitopes that span each position in the protein among the vaccinees and the reported human HIV epitopes in the Los Alamos database (Kendall's rank correlation = **[Bette,the value of tau here]**, $p < 2 \times 10^{-16}$). This consistency suggests that although the number of responses to the vaccine was low, the localization of responses to the vaccine follows that of natural infection, a likely desirable characteristic of a T cell base vaccine. Given the relative length of the Nef protein (216 amino acids, versus 850 in Pol and 500 in Gag), there was a significantly greater density of responses to Nef than to the other proteins (Chi Square test, $p = 0.03$ comparing Nef and Gag, $p = 0.007$ comparing Nef and Pol).

We next performed a series of analyses to evaluate the coverage elicited by vaccinations. We define vaccine coverage here to mean the fractions of viral sequences in the target population containing $\geq 1, \geq 2, \geq 3\ldots$ of the reactive epitopes for a given vaccinee; the average of each of these quantities over a sample of vaccinees provides the expected vaccine coverage at a given degree (e.g., expected coverage by ≥ 3 epitopes). We note that for a genetically homogeneous viral population, the expected coverage corresponds to response breadth (e.g., fraction of vaccinees with ≥ 3 reactive epitopes). Thus, the expected coverage is a combination of both the breadth of response and the pattern of population frequencies for the sets of reactive epitopes among vaccinees.

Direct estimates of expected vaccine coverage are presented in **Table 2**. The average vaccinee would cover with ≥1 epitopes in any gene 57% of a clade B viral population and this coverage is only 20% for coverage with ≥3 epitopes. Restricting consideration to Gag-specific epitopes reduces the expected coverage by ≥1 and ≥3 epitopes to 31% and 1%, respectively.

To augment this analysis of expected vaccine coverage, we explicitly examined epitope-specific coverage by comparing the overall conservation levels of observed epitopes to 1) known epitopes in the LANL database and 2) all the potential 9 amino acid T cell epitopes (potential T cell epitopes [PTEs]) within the three immunogens in subtype B strains (**Figure 2**). The frequencies of the observed epitopes among vaccinees to each of the three proteins, Gag, Pol, and Nef are significantly lower than those of the known epitopes (Kolmogorov-Smirnov test; $p < 0.001$ Gag, $p < 0.001$ Nef, $p < 0.01$ Pol), although this result is likely due to an artifact of the similarity in the reagents to detect CTL responses. ("founder effect") [22, 23]. This effect is particularly pronounced for the Nef protein. Epitope frequencies to the three HIV proteins elicited by vaccination were also lower than 9-mer PTEs in the immunogens (Kolmogorov-Smirnov test; $p < 0.05$ Gag, $p < 0.05$ Nef, $p < 0.01$ Pol), suggesting a preference for recognition of variable epitopes. Although statistically significant for Gag and Pol, the magnitude of these differences is relatively modest (**Figure 2**). Nonetheless, detailed examination of the differences for Gag and Pol suggest that relatively variable epitopes (with frequency 0.1-0.4) were over-represented in the observed epitopes (**Figure 2, Table 3**) and highly conserved epitopes (frequency > 0.8) in Gag and Pol were under-represented. A similar difference in the distribution of recognition frequencies of highly conserved versus less

conserved potential epitopes after wild-type infection with HIV-1 was noted based on the Los Alamos Immunology database (Wilcoxon rank test $p = 0.0006$), suggesting some form of "antigenic masking" of these regions. The mechanism for these observations are unclear; potential explanations include that such highly conserved but relatively immunologically silent regions could be the consequence of retroviruses evolving to evade epitope processing in regions where mutations have a very high fitness cost, or could be due to very conserved domains also being conserved in HERVs, causing such regions to be seen as "self". This lack of eliciting conserved epitopes is a disconcerting finding that may be a major factor for the future of the STEP trial and helps to define future goals from preliminary studies of subsequent candidate T-cell vaccines.

Our analysis also revealed a surprisingly low frequency of recognized epitopes in the circulating strains that is not obvious from protein similarity (or distance). For example, the Nef insert has an average protein similarity of 78% to subtype B strains while the average epitope frequency in the same viruses is only 9%; this is direct consequence of CD8 epitopes typically being around 9 amino acids long, if 25% of the positions vary between two strains is highly likely that an epitope will vary. This disparity between actual epitope mapping data and sequence/protein homology is an important issue in the consideration of candidate HIV-1 vaccines types of inserts that should be designed [2, 20, 21].

We then performed a simple statistical analysis of the T cell breadth required to achieve a high probability that vaccine elicited epitopes will be of the type that match the epitope frequency in circulating strains in a population. We plotted the individual-level vaccine coverage probabilities vs breadth of response (number of defined epitopes) and

fit non-linear quantile regression models to the resulting scatterplot in order to project the response breadth required for different levels of expected vaccine coverage (**Figure 3**). The results of these analyses suggest that a T cell based vaccine containing gag/pol/nef to expect 90% coverage by 1 or more epitopes would require four or more epitopes to be recognized. A same coverage by 3 or more epitopes would require recognizing at least 8 to 9 epitopes. This prediction assumes the same degree of conservation of epitopes as observed in the current data. However, if more highly conserved epitopes are elicited by vaccination, the number of reactive epitopes elicited by vaccination required to achieve coverage may be markedly reduced. Development of a more detailed predictive model of vaccine coverage as a function of both breadth and epitope-level frequency would be useful for guiding future insert designs and for assessment of coverage by future candidate vaccines.

The above analysis of epitope coverage implicitly assumes that T-cell epitopes have equal cell-killing activity. It is clear now that all CTL responses are not created equal. First, CTL responses of different antigen specificities demonstrated markedly differences in antiviral efficacy. CTLs that recognize early proteins (*Nef*, *Tat*, and *Rev*) appear to have higher antiviral efficacy than late proteins (i.e. *Pol*) [24]. Protein expression level may also account for the antiviral efficacy differences among different antigens [25, 26]. For example, although *Pol* protein is frequently used as the vaccine immunogen for its high conservation level, considering its low expression level (5% of *Gag* protein) and low epitope density on the infected cells it may be not be an optimal T cell immunogen. Beyond the effects of protein expression levels and kinetics on epitope presentation for CTL recognition, there may be other overriding HLA allele-specific and

T-cell-specific factors. For example, recent studies suggest that epitopes restricted by some HLA alleles are more likely to be associated with reduced viral load than other epitopes [27, 28].

In summary our analyses provide an important series of observations about the outcome of the STEP trial and future immunogen design. It is clear that while immunogenic in most individuals, the Merck gag/pol/nef vaccine elicited a narrow epitope breadth, and importantly the epitopes induced were not to the most highly conserved regions of the genome. Overall, the conservation level of the recognized epitopes by the vaccinees was lower than what one might expect for the given vaccine strains. In particular, a strong preference for recognition of relative variable epitopes was observed.

Our analysis suggests several new avenues for eliciting effective T cell-based immunity for HIV-1 vaccines. One approach to achieving the immune focusing required to elicit responses to conserved regions of the genome is to construct artificial immunogens by linking together highly-conserved epitopes or regions [29]. However, such artificial immunogens usually lack strong immunogenicity in human trials, for reasons that are neither obvious nor fully explored. An interesting alternative approach is to use heterologous inserts for prime and boost vaccination. This strategy, especially if utilized with a viral vector, may generate a broadened response to conserved epitopes with each successive inoculation. Computational algorithms to design optimal heterologous pairs that selectively boost the conserved epitope determinants can be designed and this strategy has the potential to overcome the serious limitations of the immune system observed in this study. Alternatively, one could attempt to develop

immunogens with variable regions masked against immune recognition, although if he most conserved epitopes are not presented in natural infection a vaccine stimulated response to these epitopes may be of limited value. A better understanding of antigen processing mechanisms is needed to define how to better elicit T cell responses to conserved regions—the basis of an effective T cell vaccine.

METHODS SUMMARY

Epitope mapping. Epitope mapping after vaccination was performed on a study cohort based on previously determined positive ELISPOT responses to complete 15mer peptide pools for each protein [30]. We utilized a modified IFN-γ ELISPOT assay using peptide mini-pools matching the vaccine immunogen sequences to determine the breadth of epitope response to the vaccine. Each pool contained eight overlapping 9-amino-acid peptides spanning a 16-amino-acid region. Where sample was sufficient, positive minipool responses were resolved by subsequent ELISPOT assay using the individual 9mer peptides in the responding minipool. Cut-offs to establish a 5% false-positive rate were experimentally determined to be ≥4x mock and ≥80 (gag), 55(nef), 70 (pol pool-1) and 80 (pol pool-2) spots/10^6 PBMC.

Epitope frequency in circulating strains. 130 (one per patient), recently-isolated HIV-1 subtype B and C genome sequences were obtained from Los Alamos HIV Sequence Database (http://www.hiv.lanl.gov). Intact Gag, Pol and Nef proteins sequences were derived from the corresponding coding regions of the genomes. Frequencies of reactive, known and potential T-cell epitopes were calculated according to their occurrence in the viral sequences. Experimentally defined known HIV CTL epitopes

were obtained from the same source.

ACKNOWLEDGEMENTS

This work was supported by grants 5U1AI046703-05 and AI064061 from the National Institutes of Health. We thank David Wick for scientific proofreading.

REFERENCES

1. Burgers, W.A., et al., *Design and preclinical evaluation of a multigene human immunodeficiency virus type 1 subtype C DNA vaccine for clinical trial.* J Gen Virol, 2006. **87**(Pt 2): p. 399-410.
2. Fischer, W., et al., *Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants.* Nat Med, 2007. **13**(1): p. 100-6.
3. Kong, W.P., et al., *Expanded breadth of the T-cell response to mosaic human immunodeficiency virus type 1 envelope DNA vaccination.* J Virol, 2009. **83**(5): p. 2201-15.
4. Gaschen, B., et al., *Diversity considerations in HIV-1 vaccine selection.* Science, 2002. **296**(5577): p. 2354-60.
5. Santra, S., et al., *A centralized gene-based HIV-1 vaccine elicits broad cross-clade cellular immune responses in rhesus monkeys.* Proc Natl Acad Sci U S A, 2008. **105**(30): p. 10489-94.
6. Doria-Rose, N.A., et al., *Human immunodeficiency virus type 1 subtype B ancestral envelope protein is functional and elicits neutralizing antibodies in rabbits similar to those elicited by a circulating subtype B envelope.* J Virol, 2005. **79**(17): p. 11214-24.
7. 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, 2008. **372**(9653): p. 1881-93.
8. McElrath, M.J., et al., *HIV-1 vaccine-induced immunity in the test-of-concept Step Study: a case-cohort analysis.* Lancet, 2008. **372**(9653): p. 1894-905.
9. Myers, G., Human retroviruses and AIDS 1995: a compilation and analysis of nucleic acid and amino acid sequences.
10. Priddy, F.H., et al., *Safety and immunogenicity of a replication-incompetent adenovirus type 5 HIV-1 clade B gag/pol/nef vaccine in healthy adults.* Clin Infect Dis, 2008. **46**(11): p. 1769-81.
11. Lathe, R., *Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations.* J Mol Biol, 1985. **183**(1): p. 1-12.
12. Davies, J.F., 2nd, et al., *Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase.* Science, 1991. **252**(5002): p. 88-95.
13. Larder, B.A., S.D. Kemp, and D.J. Purifoy, *Infectious potential of human immunodeficiency virus type 1 reverse transcriptase mutants with altered inhibitor sensitivity.* Proc Natl Acad Sci U S A, 1989. **86**(13): p. 4803-7.
14. Larder, B.A., et al., *Site-specific mutagenesis of AIDS virus reverse transcriptase.* Nature, 1987. **327**(6124): p. 716-7.
15. Leavitt, A.D., L. Shiue, and H.E. Varmus, *Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro.* J Biol Chem, 1993. **268**(3): p. 2113-9.
16. Mizrahi, V., et al., *Site-directed mutagenesis of the conserved Asp-443 and Asp-498 carboxy-terminal residues of HIV-1 reverse transcriptase.* Nucleic Acids Res, 1990. **18**(18): p. 5359-63.

17. Schatz, O., et al., *Point mutations in conserved amino acid residues within the C-terminal domain of HIV-1 reverse transcriptase specifically repress RNase H function.* FEBS Lett, 1989. **257**(2): p. 311-4.
18. Wiskerchen, M. and M.A. Muesing, *Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells.* J Virol, 1995. **69**(1): p. 376-86.
19. Liang, X., et al., *Development of HIV-1 Nef vaccine components: immunogenicity study of Nef mutants lacking myristoylation and dileucine motif in mice.* Vaccine, 2002. **20**(27-28): p. 3413-21.
20. Li, F., et al., *HIV-1 CTL-based vaccine immunogen selection: antigen diversity and cellular response features.* Current HIV Research, 2007. **5**: p. 95-105.
21. Li, F., et al., *Peptide selection for human immunodeficiency virus type 1 CTL-based vaccine evaluation.* Vaccine, 2006. **24**(47-48): p. 6893-904.
22. Altfeld, M., 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): p. 7330-40.
23. Malhotra, U., et al., *Enhanced detection of human immunodeficiency virus type 1 (HIV-1) Nef-specific T cells recognizing multiple variants in early HIV-1 infection.* J Virol, 2007. **81**(10): p. 5225-37.
24. Yang, O.O., *CTL ontogeny and viral escape: implications for HIV-1 vaccine design.* Trends Immunol, 2004. **25**(3): p. 138-42.
25. Yang, O.O., et al., *Efficient lysis of human immunodeficiency virus type 1-infected cells by cytotoxic T lymphocytes.* J Virol, 1996. **70**(9): p. 5799-806.
26. Yang, O.O., et al., *Suppression of human immunodeficiency virus type 1 replication by CD8+ cells: evidence for HLA class I-restricted triggering of cytolytic and noncytolytic mechanisms.* J Virol, 1997. **71**(4): p. 3120-8.
27. Serwanga, J., et al., *Host HLA B*allele-associated multi-clade Gag T-cell recognition correlates with slow HIV-1 disease progression in antiretroviral therapy-naive Ugandans.* PLoS ONE, 2009. **4**(1): p. e4188.
28. Gillespie, G.M., et al., *Cross-reactive cytotoxic T lymphocytes against a HIV-1 p24 epitope in slow progressors with B*57.* AIDS, 2002. **16**(7): p. 961-72.
29. Wilson, C.C., et al., *Development of a DNA vaccine designed to induce cytotoxic T lymphocyte responses to multiple conserved epitopes in HIV-1.* J Immunol, 2003. **171**(10): p. 5611-23.
30. Dubey, S., 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): p. 20-7.

FIGURES

Figure 1: Tracking the frequency of defined epitopes to the vaccine proteins and the database of responses to natural infection and across the vaccine proteins. The blue histograms mark the number of epitopes in the database that span each amino acid position in the protein. The yellow bars track the same thing for the 72 vaccines. The bright blue line tracks the level of coverage the vaccine provides of B clade sequences in the database. The sharp dips are either a consequence of the natural strain carrying very unusual amino acids, or amino acids engineered into the protein enhance the safety of the vaccine. The red line, in contrast, is a "neutral" measure of the overall diversity of each 9-mer, not relative to any one strain; it tracks the frequency of the most common 9 mer, for each nine mer.

Figure 2. Comparison of cumulative (upper) and density (lower) distributions of PTE, known and defined epitope in subtype B strains. The relative variable epitopes are enriched and highly conserved epitopes are under represented in this study. Known epitopes in LANL are over represented as expected.

Figure 3. Coverage by ≥1 and ≥3 epitopes of a Subtype B viral population vs response breadth (number of reactive epitopes). The points are subject-specific estimates of % viral population (Subtype B) covered by ≥1 and ≥3 epitopes versus subject-specific breadth (epitope number) on log scale. The data fit well with both quadratic and cubic regression. Breadth required for 80% and 90% expected coverage is indicated by blue lines.

TABLES

Table 1A. Breadth of response: Distribution of numbers of positive mini-pools among 72 recipients of Merck ad5 *gag/pol/nef* vaccine

	Number (%) of Vaccine Recipients			
No. Epitopes	**Any Gene**	**Gag**	**Pol**	**Nef**
0 pool	14 (19%)	33 (46%)	28 (39%)	34 (47%)
1 pool	15 (21%)	25 (35%)	15 (21%)	26 (36%)
2 pools	5 (7%)	8 (11%)	9 (13%)	8 (11%)
3 pools	9 (13%)	3 (4%)	9 (13%)	2 (3%)
≥4 pools	29 (40%)	3 (4%)	11 (15%)	2 (3%)

Table 1B. : Frequency of unique epitope responses among 72 recipients of Merck ad5 *gag/pol/nef* vaccine

	Number (%) Unique Epitopes Observed		
No. Vaccinees	**Gag**	**Pol**	**Nef**
1	22 (71%)	31 (54%)	9 (43%)
2-3	5 (16%)	23 (40%)	8 (38%)
4-5	2 (6%)	3 (5%)	4 (19%)
≥6	2 (6%)	0 (0%)	0 (0%)
Total	31(100%)	57 (100%)	21 (100%)

Table 2. Expected vaccine coverage: Fraction of target viral sequences covered by reactive epitopes among 72 recipients of Merck ad5 *gag/pol/nef* vaccine

	Expected % of viral population covered			
No. Epitopes	**Any Gene**	**Gag**	**Pol**	**Nef**
0 epitope	17%	16%	10%	35%
1 epitope	22%	25%	21%	10%
2 epitopes	15%	5%	11%	1%
3 epitopes	10%	1%	5%	0%
≥4 epitopes	10%	0%	11%	0%

Table 3. Comparison of PTE and defined epitope with different conservation level

Conservation level	Gag		Pol		Nef	
	Epitope	PTE	Epitope	PTE	Epitope	PTE
Highly variable(0-)	1 (2%)	43 (9%)	6 (6%)	132 (16%)	22 (50%)	81 (39%)
Relative variable(0.1-)	18 (34%)	62 (13%)	21 (21%)	64 (8%)	15 (34%)	63 (30%)
Relative conserved(0.4-)	24 (45%)	228 (46%)	45 (46%)	286 (34%)	7 (16%)	61 (29%)
Highly conserved(0.8-)	10 (19%)	159 (32%)	26 (27%)	360 (43%)	0 (0%)	3 (1%)
Total	**53** (100%)	492 (100%)	**98** (100%)	842 (100%)	**44** (100%)	208 (100%)

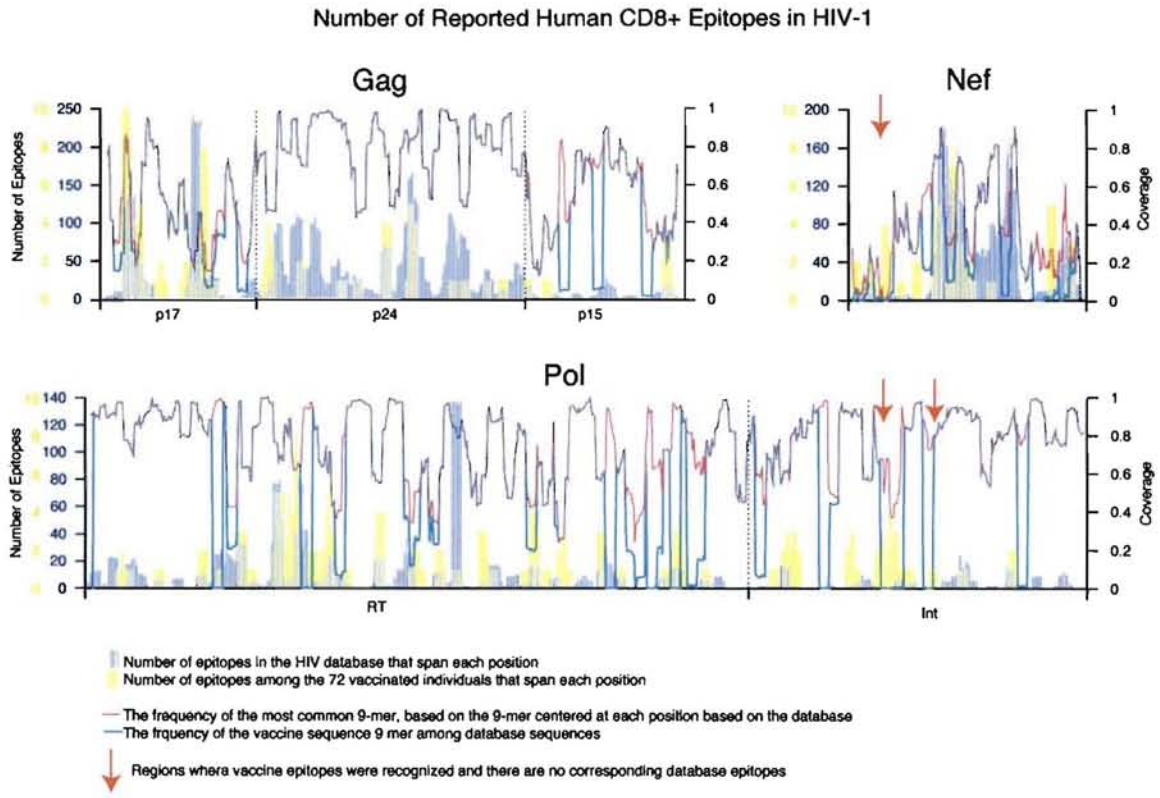

Number of Reported Human CD8+ Epitopes in HIV-1
Gag
Nef
Pol
Number of Epitopes
Coverage
p17
p24
p15
RT
Int
Number of epitopes in the HIV database that span each position
Number of epitopes among the 72 vaccinated individuals that span each position
The frequency of the most common 9-mer, based on the 9-mer centered at each position based on the database
The frquency of the vaccine sequence 9 mer among database sequences
Regions where vaccine epitopes were recognized and there are no corresponding database epitopes

Figure 1.

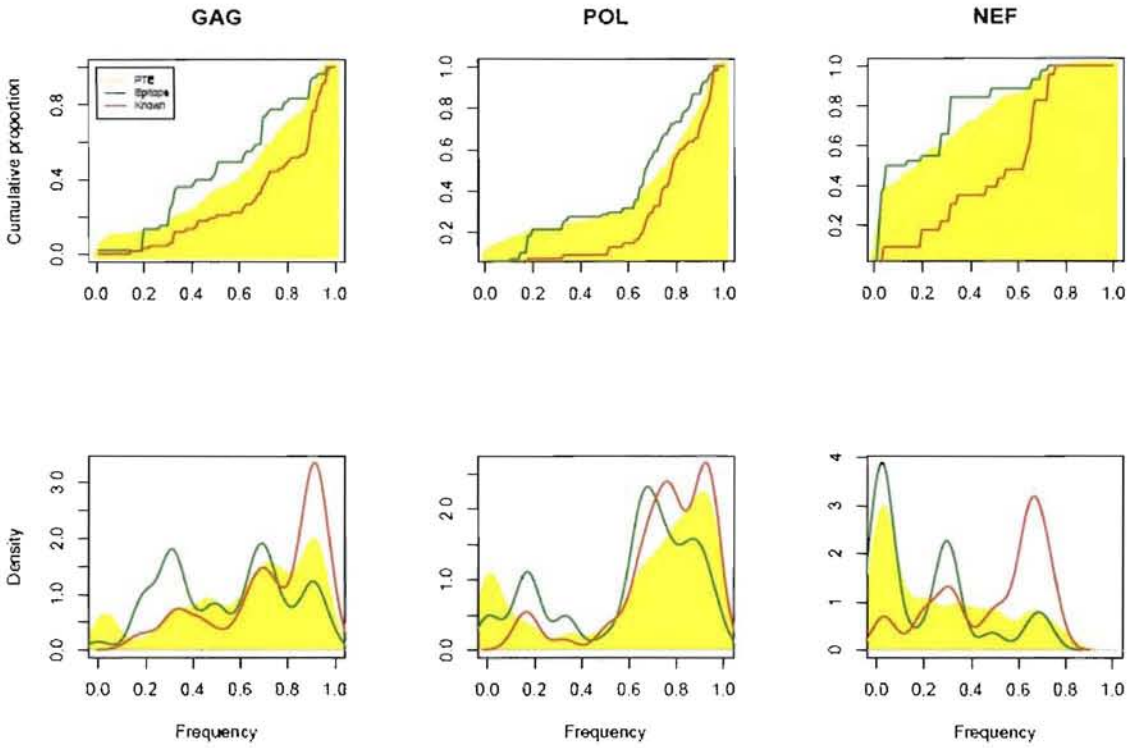
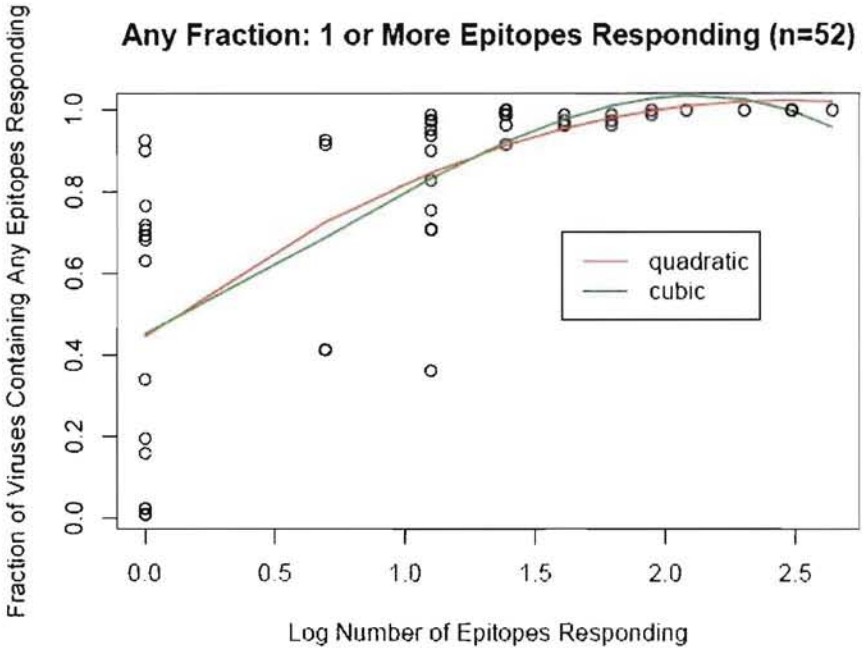

GAG
POL
NEF
Cumulative proportion
Density
Frequency
Frequency
Frequency

Figure. 2


Any Fraction: 1 or More Epitopes Responding (n=52)
Fraction of Viruses Containing Any Epitopes Responding
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
Log Number of Epitopes Responding
quadratic
cubic

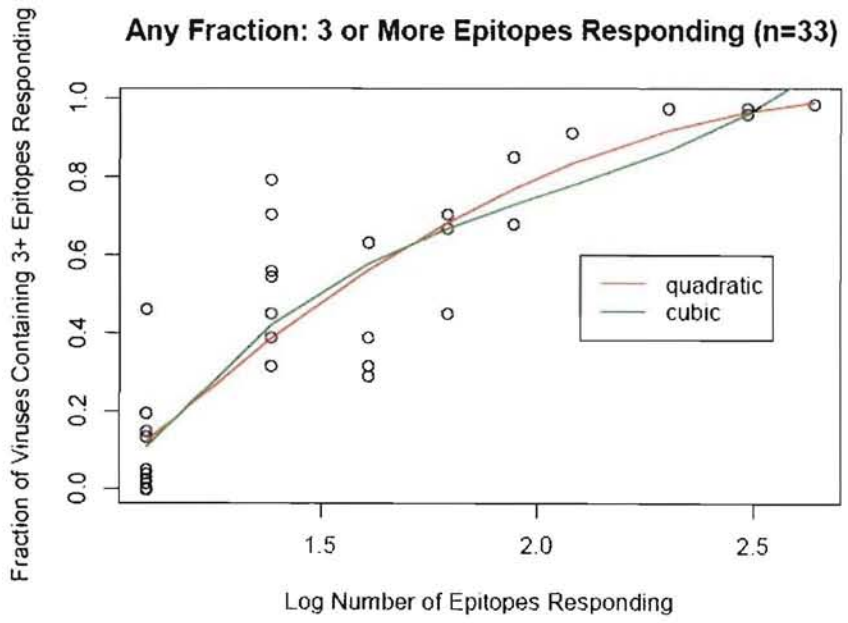

Any Fraction: 3 or More Epitopes Responding (n=33)
Fraction of Viruses Containing 3+ Epitopes Responding
0.0
0.2
0.4
0.6
0.8
1.0
1.5
2.0
2.5
Log Number of Epitopes Responding
quadratic
cubic

Figure 3

Table 1S. Positive mini-pools and 9-mer epitopes indentified from 72 recipients of Merck ad5 *gag/pol/nef* vaccine

PID	Gene	Positive mini-pool peptide	9-mer epitope
6224	Nef	NTAATNADCAWLEAQE	AATNADCAW
6238	Pol	QGWKGSPAIFQSSMTK	SPAIFQSSM
			LRPGGKKK
6239	Gag	EKIRLRPGGKKKYKLK	RLRPGGKKK
6239	Nef	MTYKGAVDLSHFLKEK	Y
6262	Nef	LSHFLKEKGGLEGLIH	GAVDLSHFL
6267	Gag	EKIRLRPGGKKKYKLK	HFLKEKGGL
6293	Gag	GKKKYKLKHIVWASRE	NO
6293	Gag	NLQGQMVHQAISPRTL	KYKLKHIVW
6293	Nef	PAADRVRRTEPAAVGV	HQAISPRTL
6293	Pol	YTAFTIPSINNETPGI	DRVRRTEPA
6293	Pol	SDLEIGQHRTKIEELR	YTAFTIPSI
6293	Pol	YPGIKVRQLCKLLRGT	
6293	Pol	GQWTYQIYQEPFKNLK	
6293	Pol	ITTESIVIWGKTPKFK	ITTESIVIW
6293	Pol	GQVRDQAEHLKTAVQM	EHLKTAVQM
6293	Pol	GIWQLDCTHLEGKVIL	
6293	Pol	HLEGKVILVAVHVASG	
6293	Pol	YRDSRNPLWKGPAKLL	
6295	Gag	EKIRLRPGGKKKYKLK	RLRPGGKKK
6295	Gag	PVGEIYKRWIILGLNK	RWIILGLNK
			QRGNFRNQ
6295	Gag	MMQRGNFRNQRKTVKC	MMQRGNFRN
6295	Nef	WRFDSKLAFHHVAREL	R
6295	Pol	YTAFTIPSINNETPGI	WRFDSKLAF
6295	Pol	AVFIHNFKRKGGIGGY	YTAFTIPS
6295	Pol	YFLLKLAGRWPVKTIH	KRKGGIGGY
6295	Pol	RWPVKTIHTDNGSNFT	NO
6296	Pol	EHEKYHSNWRAMASDF	IHTDNGSNF
6297	Gag	TVATLYCVHQKIDVKD	NO
6297	Gag	QAISPRTLNAWVKVVE	ATLYCVHQK
6297	Gag	VGGPGHKARVLAEAMS	NO
6297	Nef	CAWLEAQEDEEVGFPV	NO
6297	Nef	DEEVGFPVRPQVPLRP	NO
6297	Nef	MTYKGAVDLSHFLKEK	
6297	Pol	PEKDSWTVNDIQKLVG	AVDLSHFLK
6298	Gag	NLQGQMVHQAISPRTL	NO
6298	Nef	RERMRRAEPAADRVRR	GQMVHQAIS
			RMRRAEPAA

6298	Nef	PAADRVRRTEPAAVGV	RVRRTEPAA	
6298	Pol	QLTEAVQKITTESIVI	QLTEAVQKI	VQKITTES
6298	Pol	GIWQLDCTHLEGKVIL	THLEGKVIL	
6303	Nef	MGGKWSKRSVPGWSTV	WSKRSVPGW	
6303	Pol	WTVQPIVLPEKDSWTV	IVLPEKDSW	
6303	Pol	ITTESIVIWGKTPKFK	ITTESIVIW	
6314	Gag	PVGEIYKRWIILGLNK	RWIILGLNK	
6314	Nef	WRFDSKLAFHHVAREL	WRFDSKLAF	
6314	Pol	QGWKGSPAIFQSSMTK	AIFQSSMTK	
6314	Pol	AVFIHNFKRKGGIGGY	KRKGGIGGY	
6321	Gag	TVATLYCVHQKIDVKD	ATLYCVHQK	
6323	Gag	QPSLQTGSEELRSLYN	NO	
				LYNTVATL
6323	Gag	EELRSLYNTVATLYCV	SLYNTVATL	Y
6323	Gag	SQNYPIVQNLQGQMVH	NO	
6323	Gag	NLQGQMVHQAISPRTL	NO	
6323	Pol	LQAIYLALQDSGLEVN	LQAIYLALQ	
6323	Pol	ASDFNLPPVVAKEIVA		
6326	Gag	MMQRGNFRNQRKTVKC	RNQRKTVKC	
6326	Nef	GAVSRDLEKHGAITSS	RDLEKHGAI	
6326	Pol	KIQNFRVYYRDSRNPL	KIQNFRVYY	
6327	Nef	PMSQHGIEDPEKEVLE	IEDPEKEVL	
				IETVPVKL
6327	Pol	MAPISPIETVPVKLKP	PIETVPVKL	K
6327	Pol	RTKIEELRQHLLRWGL	IEELRQHLL	
6331	Gag	EKIRLRPGGKKKYKLK	RLRPGGKKK	
6331	Gag	GKKKYKLKHIVWASRE	KYKLKHIVW	
6331	Nef	RERMRRAEPAADRVRR	RMRRAEPAA	
6331	Nef	PAADRVRRTEPAAVGV	RVRRTEPAA	
6331	Pol	ILEPFRKQNPDIVIYQ	NO	
6331	Pol	AEIQKQGQGQWTYQIY	KQGQGQWTY	
6332	Nef	MGGKWSKRSVPGWSTV	WSKRSVPGW	
6333	Nef	PAADRVRRTEPAAVGV	RVRRTEPAA	
6333	Pol	WTVQPIVLPEKDSWTV	IVLPEKDSW	
6333	Pol	LTEEAELELAENREIL	NO	
6333	Pol	QLTEAVQKITTESIVI	VQKITTESI	
6337	Nef	PMSQHGIEDPEKEVLE	NO	
6337	Pol	MAPISPIETVPVKLKP	NO	
6337	Pol	QGWKGSPAIFQSSMTK	NO	
6337	Pol	WGKTPKFKLPIQKETW	NO	
6340	Gag	GKKKYKLKHIVWASRE	KYKLKHIVW	
6340	Gag	HQKIDVKDTKEALEKI		
6340	Gag	PVGEIYKRWIILGLNK		
6340	Nef	LSHFLKEKGGLEGLIH	HFLKEKGGL	
6340	Pol	IVTDSQYALGIIQAQP		
6340	Pol	GEAMHGQVDCSPGIWQ		

6340	Pol	AVQMAVFIHNFKRKGG			
6340	Pol	KQITKIQNFRVYYRDS	KIQNFRVYY		
6357	Pol	VESMNKELKKIIGQVR	NO		
6361	Gag	PVGEIYKRWIILGLNK	RWIILGLNK		
6361	Nef	WRFDSKLAFHHVAREL	WRFDSKLAF		
6361	Pol	AVFIHNFKRKGGIGGY	KRKGGIGGY		
6367	Pol	NLPPVVAKEIVASCDK	LPPVVAKEI		
6370	Gag	STLQEQIGWMTNNPPI	NO		
6370	Nef	MTYKGAVDLSHFLKEK	NO		
6371	Pol	WTVQPIVLPEKDSWTV	IVLPEKDSW		
6373	Pol	PLDEDFRKYTAFTIPS	PLDEDFRKY		
6375	Gag	TVATLYCVHQKIDVKD	ATLYCVHQK		
6375	Nef	MTYKGAVDLSHFLKEK	AVDLSHFLK		
6375	Pol	QGWKGSPAIFQSSMTK	NO		
6379	Nef	PAADRVRRTEPAAVGV	DRVRRTEPA		
6380	Gag	QPSLQTGSEELRSLYN	GSEELRSLY		
6380	Gag	TVATLYCVHQKIDVKD	ATLYCVHQK		
6380	Nef	DEEVGFPVRPQVPLRP	VGFPVRPQV		
6380	Nef	MTYKGAVDLSHFLKEK	AVDLSHFLK		
6380	Pol	QGWKGSPAIFQSSMTK	GSPAIFQSS	QGWKGSPA I	AIFQSSMT K
6380	Pol	ILEPFRKQNPDIVIYQ	KQNPDIVIY		
6380	Pol	RTKIEELRQHLLRWGL	ELRQHLLRW	RQHLLRWG L	
6380	Pol	TTPDKKHQKEPPFLWM	NO		
6380	Pol	AEIQKQGQGQWTYQIY	KQGQGQWTY		
6380	Pol	LPIQKETWETWWTEYW	WETWWTEYW		
6388	Gag	LLVQNANPDCKTILKA	ANPDCKTIL		
6388	Nef	FHHVARELHPEYYKDC	NO		
6395	Pol	YTAFTIPSINNETPGI	NO		
6395	Pol	INNETPGIRYQYNVLP	NETPGIRYQ		
6395	Pol	AETGQETAYFLLKLAG	GQETAYFLL		
6396	Gag	QVTNSATIMMQRGNFR	QVTNSATIM	VTNSATIM M	
6396	Pol	LVSAGIRKVLFLDGID			
6396	Pol	SNFTGATVRAACWWAG	SNFTGATVR		
6396	Pol	IKQEFGIPYNPQSQGV	GIPYNPQSQ		
6400	Gag	STLQEQIGWMTNNPPI	NO		
6403	Gag	EKIRLRPGGKKKYKLK	RLRPGGKKK		
6403	Pol	QGWKGSPAIFQSSMTK	SPAIFQSSM	AIFQSSMT K	
6403	Pol	ITTESIVIWGKTPKFK	ITTESIVIW		
6403	Pol	YFLLKLAGRWPVKTIH	NO		
6404	Gag	STLQEQIGWMTNNPPI	STLQEQIGW		
6404	Nef	MGGKWSKRSVPGWSTV	WSKRSVPGW	SKRSVPGW S	

6404	Nef	NTAATNADCAWLEAQE	AATNADCAW	
6404	Nef	DEEVGFPVRPQVPLRP		
6404	Nef	DLWVYHTQGYFPDWQN	HTQGYFPDW	
6404	Pol	NPDIVIYQYMDDLYVG		
6404	Pol	YMDDLYVGSDLEIGQH	DLYVGSDLE	
				IVLPEKDS
6404	Pol	WTVQPIVLPEKDSWTV	TVQPIVLPE	W
6404	Pol	LCKLLRGTKALTEVIP		
				IPLTEEAE
6404	Pol	KALTEVIPLTEEAELE	VIPLTEEAE	L
6404	Pol	ITTESIVIWGKTPKFK		
6404	Pol	WGKTPKFKLPIQKETW		
6404	Pol	LPIQKETWETWWTEYW		
6404	Pol	QDSGLEVNIVTDSQYA	GLEVNIVTD	
				IKKEKVYL
6404	Pol	NQIIEQLIKKEKVYLA	NQIIEQLIK	A
6404	Pol	VASGYIEAEVIPAETG		
6404	Pol	KLAGRWPVKTIHTDNG		
6404	Pol	SNFTGATVRAACWWAG		
6404	Pol	LQAIYLALQDSGLEVN		
6404	Pol	WVPAHKGIGGNEQVDK		
6431	Gag	TVATLYCVHQKIDVKD	NO	
6431	Nef	MTYKGAVDLSHFLKEK	NO	
6431	Pol	QLTEAVQKITTESIVI	NO	
6431	Pol	RWPVKTIHTDNGSNFT	NO	
6431	Pol	WKGPAKLLWKGEGAVV	NO	
6435	Gag	LKETINEEAAEWDRLH	KETINEEAA	
6435	Pol	ILEPFRKQNPDIVIYQ	KQNPDIVIY	
6435	Pol	VVPRRKAKIIRDYGKQ	RKAKIIRDY	
6435	Pol	SAGERIVDIIATDIQT	GERIVDIIA	
				ATLYCVHQ
6436	Gag	TVATLYCVHQKIDVKD	VATLYCVHQ	K
6436	Nef	DEEVGFPVRPQVPLRP	VGFPVRPQV	
6436	Nef	MTYKGAVDLSHFLKEK	AVDLSHFLK	
6436	Pol	NLPPVVAKEIVASCDK	LPPVVAKEI	
6540	Gag	TVATLYCVHQKIDVKD	ATLYCVHQK	
6540	Gag	VDRFYKTLRAEQASQE	DRFYKTLRA	
6540	Nef	DEEVGFPVRPQVPLRP	NO	
6540	Nef	MTYKGAVDLSHFLKEK	AVDLSHFLK	
6540	Pol	IVTDSQYALGIIQAQP		
6540	Pol	GEAMHGQVDCSPGIWQ		
6540	Pol	KQITKIQNFRVYYRDS	KIQNFRVYY	
6569	Gag	EKAFSPEVIPMFSALS	EVIPMFSAL	
6569	Nef	GAVSRDLEKHGAITSS	RDLEKHGAI	
6569	Nef	GGLEGLIHSQKRQDIL	HSQKRQDIL	
6574	Gag	STLQEQIGWMTNNPPI	STLQEQIGW	

6574	Nef	NTAATNADCAWLEAQE	AATNADCAW	
6574	Pol	ITTESIVIWGKTPKFK	ITTESIVIW	
6580	Gag	KELYPLASLRSLFGND	NO	
6580	Nef	PMSQHGIEDPEKEVLE	IEDPEKEVL	
				VRRTEPAA
6589	Nef	PAADRVRRTEPAAVGV	RVRRTEPAA	V
6589	Pol	FWEVQLGIPHPAGLKK	NO	
6589	Pol	KKSVTVLDVGDAYFSV	SVTVLDVGD	
				QNPDIVIY
6589	Pol	ILEPFRKQNPDIVIYQ	FRKQNPDIV	Q
6589	Pol	YMDDLYVGSDLEIGQH	NO	
6598	Gag	KELYPLASLRSLFGND	KELYPLASL	
6598	Nef	PMSQHGIEDPEKEVLE	IEDPEKEVL	
6598	Pol	LPIQKETWETWWTEYW	WETWWTEYW	
6598	Pol	VASGYIEAEVIPAETG	NO	
6610	Gag	EKIRLRPGGKKKYKLK	RLRPGGKKK	
6610	Nef	RERMRRAEPAADRVRR	RMRRAEPAA	
6610	Nef	PAADRVRRTEPAAVGV	RVRRTEPAA	
6610	Pol	QLTEAVQKITTESIVI	QLTEAVQKI	VQKITTES
7760	Gag	AAEWDRLHPVHAGPIA	HPVHAGPIA	
7760	Gag	RAEQASQEVKNWMTET	QASQEVKNW	
7760	Nef	NTAATNADCAWLEAQE	AATNADCAW	
7760	Nef	DPEKEVLEWRFDSKLA	DPEKEVLEW	
7760	Nef	WRFDSKLAFHHVAREL	KLAFHHVAR	
7760	Pol	SDLEIGQHRTKIEELR	NO	
7760	Pol	YQLEKEPIVGAETFYV	EPIVGAETF	
7760	Pol	VAVHVASGYIEAEVIP	NO	
7760	Pol	KELQKQITKIQNFRVY	NO	
7760	Pol	AVFIHNFKRKGGIGGY	NO	
7760	Pol	LEVNIVTDSQYALGII		
7760	Pol	VYLAWVPAHKGIGGNE	YLAWVPAHK	
7760	Pol	KIQNFRVYYRDSRNPL	NO	
7764	Gag	EKAFSPEVIPMFSALS	EVIPMFSAL	
7764	Nef	GGLEGLIHSQKRQDIL	NO	
7764	Nef	GENNCLLHPMSQHGIE	NO	
7764	Pol	ITTESIVIWGKTPKFK	ITTESIVIW	
7764	Pol	EHEKYHSNWRAMASDF	EHEKYHSNW	
7764	Pol	DGAANRETKLGKAGYV	ETKLGKAGY	
7765	Gag	EKIRLRPGGKKKYKLK	RLRPGGKKK	
7765	Pol	LAENREILKEPVHGVY	NO	
7765	Pol	RKGGIGGYSAGERIVD	NO	
				IQNFRVYY
7770	Pol	KIQNFRVYYRDSRNPL	KIQNFRVYY	R
				YKLKHIVW
7773	Gag	GKKKYKLKHIVWASRE	KYKLKHIVW	A
7773	Gag	QVTNSATIMMQRGNFR	NSATIMMQR	

7773	Nef	SVPGWSTVRERMRRAE	SVPGWSTVR		
				IPLTEEAE	
7773	Pol	KALTEVIPLTEEAELE	VIPLTEEAE	L	
7773	Pol	GYELHPDKWTVQPIVL			
7773	Pol	WGKTPKFKLPIQKETW			
7773	Pol	TGKYARMRGAHTNDVK			
7773	Pol	TDNGSNFTGATVRAAC	SNFTGATVR		
				IIEQLIKK	
7773	Pol	SELVNQIIEQLIKKEK	NQIIEQLIK	E	
7773	Pol	YQLEKEPIVGAETFYV	PIVGAETFY		
7773	Pol	FGIPYNPQSQGVVESM	GIPYNPQSQ		
				SGLEVNIV	GLEVNIVT
7773	Pol	YLALQDSGLEVNIVTD	LALQDSGLE	T	D
7773	Pol	RWPVKTIHTDNGSNFT			
7773	Pol	GIRKVLFLDGIDKAQD			
7773	Pol	VAVHVASGYIEAEVIP			
7773	Pol	VYLAWVPAHKGIGGNE			
7773	Pol	NLPPVVAKEIVASCDK			
7773	Pol	EHEKYHSNWRAMASDF			
7773	Pol	KIQNFRVYYRDSRNPL			
7773	Pol	LEVNIVTDSQYALGII			
7773	Pol	QWPLTEEKIKALVEIC			
7773	Pol	NPDIVIYQYMDDLYVG			
7773	Pol	YMDDLYVGSDLEIGQH	DLYVGSDLE		
7776	Gag	EKIRLRPGGKKKYKLK	RLRPGGKKK		
				LYNTVATL	
7776	Gag	EELRSLYNTVATLYCV	SLYNTVATL	Y	
7776	Gag	TVATLYCVHQKIDVKD	NO		
7776	Gag	NLQGQMVHQAISPRTL			
7776	Gag	QAISPRTLNAWVKVVE			
7776	Gag	AAEWDRLHPVHAGPIA			
7776	Nef	WRFDSKLAFHHVAREL	AFHHVAREL		
7776	Pol	QWPLTEEKIKALVEIC			
7776	Pol	FWEVQLGIPHPAGLKK			
7776	Pol	PLDEDFRKYTAFTIPS	PLDEDFRKY		
				TAFTIPSI	
7776	Pol	YTAFTIPSINNETPGI	YTAFTIPSI	N	
7776	Pol	IFQSSMTKILEPFRKQ	SMTKILEPF		
7776	Pol	NPDIVIYQYMDDLYVG	NPDIVIYQY		
7776	Pol	TTPDKKHQKEPPFLWM			
7776	Pol	YPGIKVRQLCKLLRGT			
7776	Pol	LAENREILKEPVHGVY			
7776	Pol	QEPFKNLKTGKYARMR			
7776	Pol	ETWWTEYWQATWIPEW			
7776	Pol	YQLEKEPIVGAETFYV	NO		
7776	Pol	NLPPVVAKEIVASCDK	LPPVVAKEI		

7776	Pol	LEVNIVTDSQYALGII	NO	
7776	Pol	QKTELQAIYLALQDSG		
7776	Pol	RKGGIGGYSAGERIVD		
7776	Pol	GATVRAACWWAGIKQE		
7781	Gag	PVGEIYKRWIILGLNK	RWIILGLNK	
7781	Gag	LGPAATLEEMMTACQG	NO	
7782	Pol	KIQNFRVYYRDSRNPL	KIQNFRVYY	
7783	Gag	EKIRLRPGGKKKYKLK	IRLRPGGKK	
7783	Pol	PLDEDFRKYTAFTIPS	PLDEDFRKY	
7783	Pol	IFQSSMTKILEPFRKQ	SMTKILEPF	
7783	Pol	NPDIVIYQYMDDLYVG	NPDIVIYQY	
7783	Pol	YPGIKVRQLCKLLRGT	PGIKVRQLC	
7783	Pol	LAENREILKEPVHGVY	AENREILKE	
7783	Pol	GATVRAACWWAGIKQE	GATVRAACW	
7791	Gag	RVLAEAMSQVTNSATI	VLAEAMSQV	
7791	Gag	KELYPLASLRSLFGND	KELYPLASL	
7791	Nef	PMSQHGIEDPEKEVLE	IEDPEKEV	
				IETVPVKL
7791	Pol	MAPISPIETVPVKLKP	PIETVPVKL	K
7791	Pol	RTKIEELRQHLLRWGL	IEELRQHLL	
7794	Gag	EKIRLRPGGKKKYKLK	NO	
7794	Gag	GKKKYKLKHIVWASRE	NO	
7794	Gag	HIVWASRELERFAVNP	NO	
				NPGLLETS
7794	Gag	LERFAVNPGLLETSEG	FAVNPGLLE	E
7794	Gag	NLQGQMVHQAISPRTL	HQAISPRTL	
7794	Gag	PVHAGPIAPGQMREPR	NO	
7794	Nef	PAADRVRRTEPAAVGV	NO	
7794	Nef	GGLEGLIHSQKRQDIL	NO	
7794	Nef	DLWVYHTQGYFPDWQN	NO	
7794	Nef	GENNCLLHPMSQHGIE	NO	
7798	Pol	SELVNQIIEQLIKKEK	NQIIEQLIK	
7798	Pol	VAVHVASGYIEAEVIP		
7798	Pol	YQLEKEPIVGAETFYV	NO	
				GLEVNIVT
7798	Pol	YLALQDSGLEVNIVTD	LALQDSGLE	D
7798	Pol	SAGERIVDIIATDIQT		
7798	Pol	WRAMASDFNLPPVVAK	NO	
7798	Pol	KELQKQITKIQNFRVY		
7798	Pol	LEVNIVTDSQYALGII	NO	
7798	Pol	NLPPVVAKEIVASCDK	NO	
7798	Pol	RWPVKTIHTDNGSNFT	NO	

[1]Yellow lines indicate mini-pools without deconvolution study due to either limited sample or relative low T-cell response.

[2]Blank lines are the mini-pools that no epitopes can be identified by deconvolution study.