# Increased Sequence Diversity Coverage Improves Detection of HIV-Specific T Cell Responses[1]

**Nicole Frahm,[2]* Daniel E. Kaufmann,[2]* Karina Yusim,[2†] Mark Muldoon,[‡] Can Kesmir,[§¶] Caitlyn H. Linde,* Will Fischer,[†] Todd M. Allen,* Bin Li,* Ben H. McMahon,[†] Kellie L. Faircloth,* Hannah S. Hewitt,* Elizabeth W. Mackey,* Toshiyuki Miura,* Ashok Khatri,[‖] Steven Wolinsky,[#**] Andrew McMichael,[††] Robert K. Funkhouser,[†] Bruce D. Walker,*[‡‡] Christian Brander,[2,3]* and Bette T. Korber[2,3†§§]**

**The accurate identification of HIV-specific T cell responses is important for determining the relationship between immune response, viral control, and disease progression. HIV-specific immune responses are usually measured using peptide sets based on consensus sequences, which frequently miss responses to regions where test set and infecting virus differ. In this study, we report the design of a peptide test set with significantly increased coverage of HIV sequence diversity by including alternative amino acids at variable positions during the peptide synthesis step. In an IFN-$\gamma$ ELISpot assay, these "toggled" peptides detected HIV-specific $CD4^+$ and $CD8^+$ T cell responses of significantly higher breadth and magnitude than matched consensus peptides. The observed increases were explained by a closer match of the toggled peptides to the autologous viral sequence. Toggled peptides therefore afford a cost-effective and significantly more complete view of the host immune response to HIV and are directly applicable to other variable pathogens.** *The Journal of Immunology,* **2007, 179: 6638–6650.**

Human immunodeficiency virus (HIV) sequence diversity is considered one of the major hurdles for the design of an effective and broadly applicable HIV vaccine (1). It is also a considerable challenge for the design of in vitro test sets that will reliably detect virus-specific T cell responses, especially to variable regions of the viral genome (2). However, despite the extraordinary global HIV sequence diversity reported to date, the variability in single positions of the viral genome is often highly restricted, likely reflecting functional constraints of the viral proteins (3–8). These limitations define natural bounds of the evolutionary space in which HIV exists and which can be exploited to design HIV sequences that cover the most prevalent viral populations of specific clades or of group M as a whole (9–11).

Historically, either natural strains or consensus sequences have been used as the basis for peptide synthesis when probing the cellular immune response to HIV (12–15). Since consensus protein sequences, by definition, contain the residue most frequently found in circulating viruses, they provide wider coverage of viral sequence diversity than individual isolate-specific sequences, yet fail to properly represent subdominant variants. We overcame this limitation by incorporating amino acid mixes in the peptide synthesis for positions displaying sequence variability, creating a peptide set that captures sequence diversity in variable regions of the virus. Given the recurrent nature of sequence changes between HIV clades, such a "toggled" peptide set can cover a significant degree of diversity in all group M sequences and may significantly improve the accurate assessment of the host immune response to HIV. Indeed, data presented here establish toggled peptides as a powerful approach to detect significantly stronger and broader CD4 and CD8 T cell responses than consensus peptides. These data demonstrate the usefulness of the toggled peptides approach to more accurately assess the host immune response to HIV, and by inference, to other variable pathogens.

## Materials and Methods

### *Study subjects*

Subjects at different stages of HIV infection (see *Results*) were recruited at three hospitals in the Boston area. Additionally, 11 HIV-negative subjects were recruited at Massachusetts General Hospital. All human subject protocols have been approved by the Partners Human Research Committee, and all subjects provided written informed consent before enrollment.

### *Quantification of surface preference of toggling sites*

The solvent-exposed surface area of each atom in the respective protein (reverse transcriptase and capsid) was calculated by the MSMS algorithm,

*Partners AIDS Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02129; †Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545; ‡School of Mathematics, University of Manchester, Manchester, United Kingdom; §Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark; ¶Theoretical Biology/Bioinformatics, Utrecht University, Utrecht, The Netherlands; ‖Endocrine Unit, Massachusetts General Hospital, Charlestown, MA 02114; #Department of Medicine, Feinberg School of Medicine, Northwestern University Chicago, IL 60611; **International Institute for Nanotechnology, Northwestern University, Evanston, IL 60208; ††Medical Research Council (MRC) Human Immunology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom; ‡‡Howard Hughes Medical Institute, Chevy Chase, MD 20815; and §§ Santa Fe Institute, Santa Fe, NM 87501

Received for publication May 29, 2007. Accepted for publication August 17, 2007.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported in part by Federal funds from the National Institute of Allergy and Infectious Disease and National Institute of Health Contracts Nos. N01-AI-15442 and AI-30024 (to B.D.W.), R01-AI-067077 (to C.B.), R56-AI-071726 (to C.B.), R01-AI-054178 (to T.M.A.), and R21-AI-055421 (to K.Y.). It was also supported by internal directed research funds (LDRD) from the Los Alamos National Laboratory and a Harvard University Center for Acquired Immunodeficiency Syndrome Research grant (to D.E.K).

[2] N.F., D.E.K., K.Y., C.B., and B.T.K. contributed equally to this work.

[3] Address correspondence and reprint requests to Dr. Bette T. Korber, MS K710, T-10, Los Alamos National Laboratory, Los Alamos, NM 87545. E-mail address: btk@lanl.gov; or Dr. Christian Brander, Massachusetts General Hospital, AIDS Research Center, 149 13th Street, Room 5234, Charlestown, MA 02129. E-mail address: cbrander@partners.org

described at: http://www.scripps.edu/mb/olson/people/sanner/html/msms_home.html and implemented in the Chimera program (http://www.cgl.ucsf.edu/chimera/). Backbone atoms were excluded from the comparison of atom-wise solvent-exposed surface area. The cumulative distribution functions of the levels of solvent exposure of each side chain atom in the sets of toggled and conserved positions were compared using a two-sample Kolmogorov-Smirnov test as implemented in Splus 10.0 (16).

### *Toggled peptides*

Toggled peptides were based on the design of a previously published set of consensus B 18-mer peptides overlapping by 10 amino acids (12). The toggled peptides were synthesized on an AAPPTEC (Louisville, KY), model Apex 396 multiple peptide synthesizer using 9-fluorenylmethoxycarbonyl/tertiary butyl (Fmoc/tBu) solid-phase chemistry. An optimized 20 $\mu$mol scale synthesis cycle with a 6-fold molar excess of Fmoc amino acid and 1*H*-benzotriazolium, 1-[bis(dimethyamino)methylene] 5-chlorohexafluorophosphate(1-), 3-oxide/*N*,*N*-diisopropylethyamine activation was used. Each amino acid was double coupled each time for 75 min. For the toggled positions in the sequence, a mixture containing an equimolar amount of the representative Fmoc amino acids was used. Peptides were cleaved from the solid support and deprotected using reagent K (trifluoroacetic acid/phenol/thioanisole/water/ethanedithiol; 82.5/5.0/5.0/5.0/2.5 v/v) for 3 h at ambient temperature. Peptides were precipitated using cold methyl tertiary butyl ether. The precipitate was washed 3 times with methyl tertiary butyl ether followed by freeze-drying. All toggled peptides were characterized by MALDI-mass spectrometry (MS).[4] Edman degradation was used for some toggled peptides to confirm the presence of similar ratios of variants in the mix (data not shown). For peptide design, the variability cutoff was 5% (i.e., all variant amino acids present in at least 5% of clade B database sequences were included in the mix) for Gag p24 and Pol, and 10% for Nef. This gave rise to toggles containing between two and 1296 peptide variants with a median of four variants per toggle product and only five toggled peptides with >100 variants in the mixture. The vast majority (87%, 148/170) of toggles had <20 variants in the toggle mixture and were adjusted for peptide concentrations so that each single peptide was present in the toggle at the same concentration as it was in the consensus overlapping peptide (OLP) preparation. The individual variant peptide concentration in the remaining 13% of toggles containing >20 variants (median 42) was lower depending on the number of variants, since the overall peptide concentration in DMSO would have exceeded solubility if adjusted, and higher solvent concentration could potentially cause adverse effects in the ELISpot assays due to DMSO toxicity.

### *ELISpot assays*

ELISpot assays were performed as described previously (12), using 2 $\mu$g/ml coating Ab. PBMCs were incubated with consensus or toggled peptides overnight. Cells in medium without peptide were used as negative control (≥6 wells), PHA was added as positive control. Responses were considered positive when exceeding all of the following criteria: four times mean negative wells, mean negative wells plus three SDs, 5 spots/well and 55 SFC/million input cells. Assays with more than mean three spots/well in the negative controls were disregarded. Testing 11 HIV-negative individuals with all toggled peptides resulted in 3 weak positive responses, representing a false positive rate of 0.1%. All chronic subjects had detectable responses to both consensus and toggled peptides.

For assessment of HIV-specific CD4 T cell responses, PBMCs were CD8 depleted by RosetteSep (StemCell Technologies) or with Dynabeads (Invitrogen life Technologies) according to the manufacturers' instructions.

For simplicity and to allow for a direct assessment of responses to toggled vs consensus-based peptides on freshly isolated samples, all toggled as well as consensus peptides were tested separately without pooling. All assays for total T cell responses were done in duplicate, whereas assays using CD8-depleted samples were done without replicates.

### *Statistical methods*

To assess the efficacy of toggle peptides, the relative frequency of yes/no outcomes using toggled peptide sets or consensus sequence peptides was predicted. For the breadth of responses, we considered whether there is or is not a response to a particular peptide. The natural tool for this sort of Boolean yes/no problem is binomial logistic regression: it amounts to making a generalized linear model (GLM) for the "logit"-transformed probability of a positive response (17, 18)

$$\text{logit}\,(p) = \ln\left(\frac{p}{1-p}\right)$$

and yields a set of predicted probabilities as well as a list of coefficients that permit one to compute odds ratios and assign $p$ values to them. We used the statistical package R for these calculations (19), following Crawley (20), applying a GLM using a quasi-binomial family for the error function. In these analyses, we considered only points with duplicate assays that were either both positive or both negative for total T cell responses, and excluded discordant duplicates with one positive and one negative response. The $CD4^+$ T cell reactions were not available in duplicate and were defined as either positive or negative based on a single reaction. The data was not over-dispersed when considering the Boolean positive or negative responses data: the dispersion parameter for the quasi-binomial family was 0.78 and 1.47 for the total T cell and for CD4 T cell, respectively. This method allows detailed modeling of the complete response profiles; we also performed nonparametric Wilcoxon tests to compare the overall distributions of responses in the patients, again using the R package.

There were 1060 duplicate reactions performed, and comparing duplicates gave us a foundation for defining a conservative cut-off for a significant difference in ELISpot scores. The absolute value of the difference between all duplicate reactions was determined, and 95% of the duplicates had differences of <160 spot forming cells (SFC)/million (median 40, IR 12–80).

### *Sequencing*

Genomic DNA was isolated from PBMC samples using the QIAamp DNA blood mini kit. Nested PCR protocols with limiting dilution were used to amplify near full-length HIV genomes as previously described (21). Purified PCR products were directly sequenced using Clade B consensus primers. Sequence data were manually edited using Sequencher 4.6 (Gene Code Corporation). All available sequences were screened for subtype and potential recombination or contamination. All sequences were submitted to GenBank and are available under accession numbers DQ886031–DQ886038 for chronic samples, EF090287–EF090289 for controllers, and EF680862–EF680881 for acute samples.

## Results

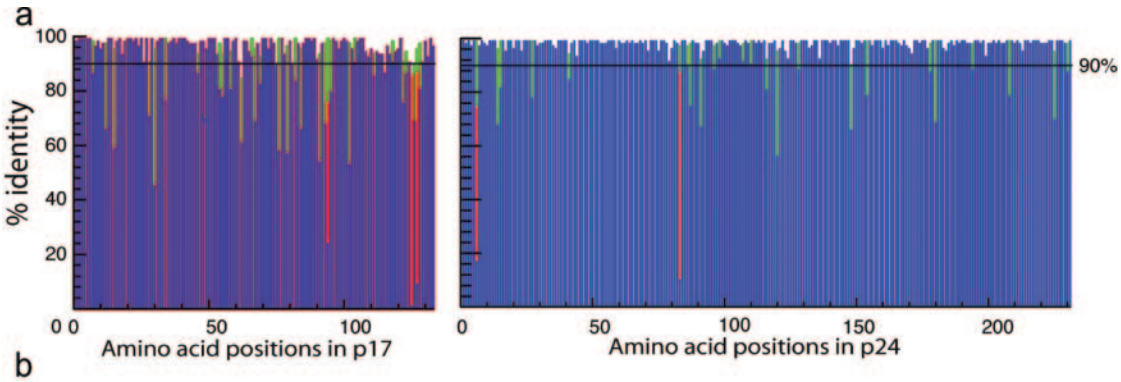
### *The restricted nature of amino acid variation in conserved HIV proteins*

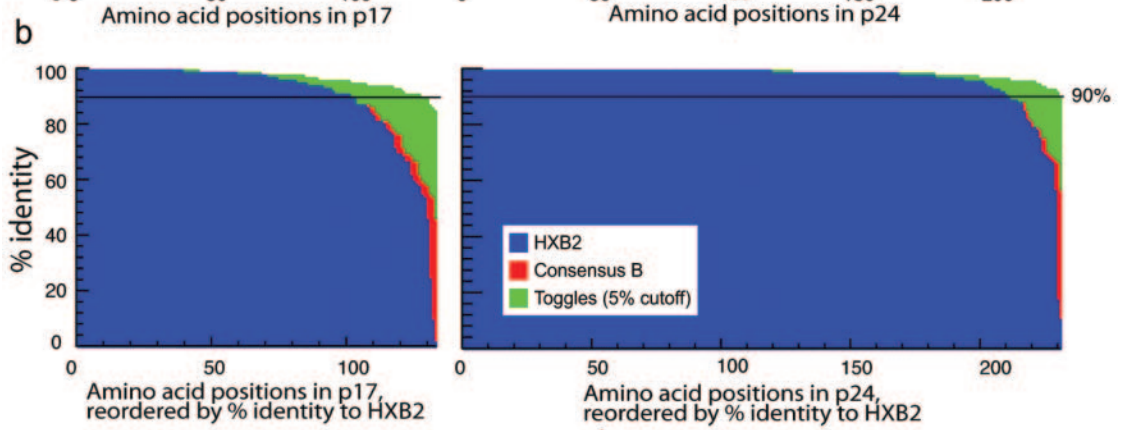
Despite extensive HIV sequence diversity on a full genome level, the variation in individual positions of viral proteins is often restricted to a limited number of amino acid substitutions that recur or "toggle" throughout the different HIV clades (22). Toggling residues often include amino acid pairs such as arginine (R) and lysine (K) or leucine (L) and valine (V), which are biochemically similar and are therefore likely to be tolerated by the virus without serious impacts on its replication efficiency. In addition, these changes are generally readily achieved as they can often be obtained by single nucleotide substitutions (23).

We designed a strategy to incorporate this considerably limited variability into peptide test sets that allow more accurate detection of virus-specific T cell responses. This is illustrated in Fig. 1 (and shown for HIV Pol in Table I), which provides detailed comparisons of amino acid variation in the Gag protein, showing levels of HIV subtype B population coverage at each position achieved by either a natural strain, a consensus sequence, or a toggled sequence incorporating the most common amino acids occurring at a given position. Gag p17 is representative of a more variable protein, while Gag p24 is representative of a highly conserved protein (24). Notably, variable positions are relatively infrequent and generally well dispersed among highly conserved positions, so that within the boundaries of a CD8 T cell epitope, which is typically nine amino acids long, a large fraction of HIV population diversity can be covered by a small number of variants.

This observation can be used for the design of a novel peptide test set, where residues at variable positions are added simultaneously to the peptide synthesis, resulting in the generation of highly and specifically restricted small peptide libraries. To allow for a direct comparison of how toggled peptides perform in in vitro

[4] Abbreviations used in this paper: MS, mass spectrometry; OLP, overlapping peptide; GLM, generalized linear model; SFC, spot forming cell; PTE, potential T cell epitope.

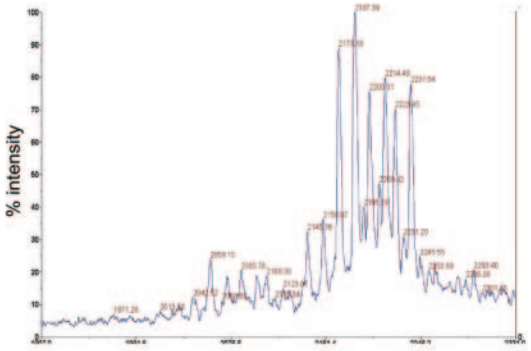



c

| | | |
|---|---|---|
| OLP 35 consensus | WMTNNPPIPVGEIYKRWI | Mass |
| OLP 35 toggled counterpart | WMT(NHS)NPPIPVG(ED)IYK(RK)WI | (kD) |
| OLP 35 toggle variants | WMTNNPPIPVGEIYKRWI | 2214 |
| | ----------D------ | 2220 |
| | ----------K-- | 2186 |
| | ----------D---K-- | 2172 |
| | ---H---------- | 2237 |
| | ---H--------D------ | 2223 |
| | ---H----------K-- | 2209 |
| | ---H--------D---K-- | 2195 |
| | ---S---------- | 2187 |
| | ---S--------D------ | 2173 |
| | ---S----------K-- | 2159 |
| | ---S--------D---K-- | 2145 |

d



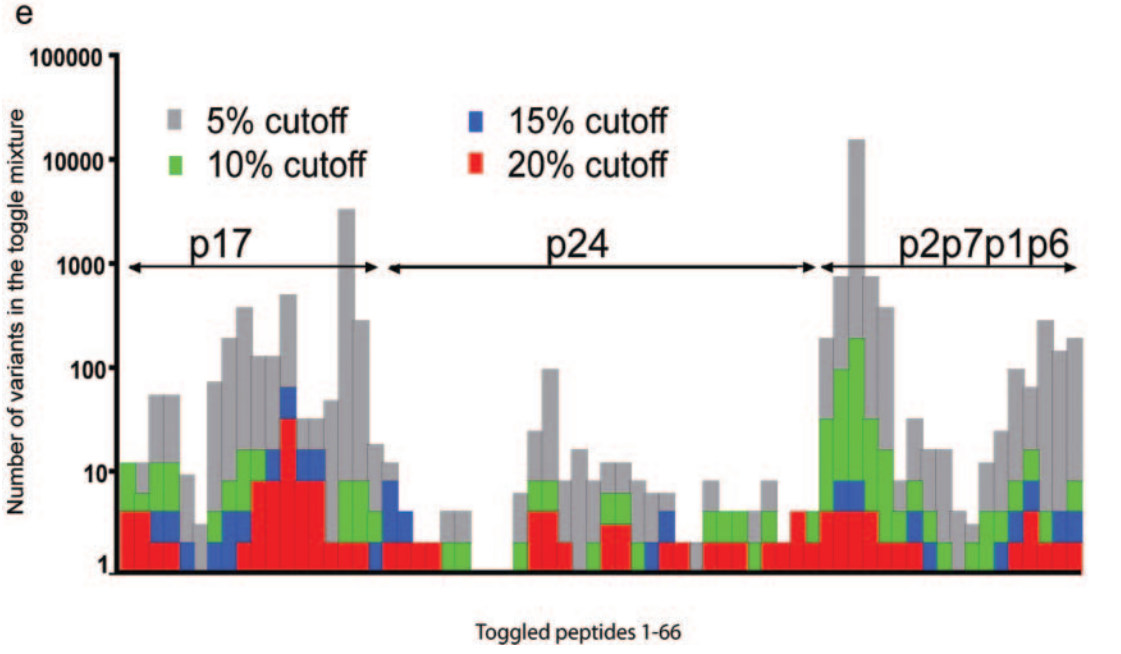



Table I. *Analysis of the sequence diversity in HIV Pol*[a]

| Position[b] | M (427)[c] | A (41) | B (62) | C (87) | D (35) |
|---|---|---|---|---|---|
| In 750 positions, a single amino acid is present in >97% of 427 group M sequences | | | | | |
| 3 | **425R**[d] | **41R** | **62R** | **86R** | **35R** |
| | 1G, 1K | | | 1G | |
| 131 positions have a shared consensus and a recurring alternative amino acid: | | | | | |
| 896 | **331I**, ***92V*** | **22I**, ***19V*** | **59I**, ***3V*** | **81I**, ***4V*** | **33I**, ***2V*** |
| | 2L, 2T | | | 2T | |
| 122 positions change consensus in different clades (A, B, C, D) | | | | | |
| 151 | **335R**, ***90K*** | ***32K***, **9R** | **55R**, ***6K*** | **84R**, ***2K*** | **34R**, ***1K*** |
| | 1G, 1T | | 1T | 1G | |

[a] Of the 1003 amino acid positions in Pol, 750 positions are more than 97% conserved throughout all available group M sequences (e.g, position 3). In the remaining positions, sequence variation occurs but is frequently limited to two or three toggling amino acids, with the same changes occurring in different clades, although at different frequencies (e.g. positions 896 and 151, respectively).

[b] An exemplary amino acid position in Pol alignments is shown of the three different scenarios.

[c] Numbers in parentheses indicate the number of Pol sequences for group M (427 sequences), clade A (41 sequences), clade B (62 sequences), clade C (87 sequences), and clade D (35 sequences) available in the LANL database at the time of analysis.

[d] Of the 427 group M sequences, 425 contain an R at position 3, thereby R represents the consensus M amino acid indicated as underlined. Amino acids in italics represent recurring alternative amino acids occurring in all clades. Very few sequences display unique amino acid changes.

assays relative to other commonly used peptide sets, toggled peptides were based on the design of a previously published set of consensus B 18-mers overlapping by 10 amino acids (12). During peptide synthesis, equimolar mixes of variant amino acids were included at appropriate elongation steps to include common amino acids in variable positions. By default, because the consensus OLP represents the most common amino acid at each position, and the most common amino acid is always included in the toggle set, the consensus OLP sequence was one of the variants in the toggled peptide product (Fig. 1*c*). Of note, over the length of a given 18mer, the consensus peptide is also found among natural strains. To control for synthesis quality and to verify, in a nonquantitative manner, that all variants were contained in the mix, all toggled peptides were characterized by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS, Fig. 1*d*). Remarkably, most p24 toggled peptide mixtures required fewer than 20 variants (median 4) even at the highest level of sequence coverage (i.e., only amino acids present in <5% of database sequences are excluded from toggle positions), while the number of variants is higher in other Gag subunits at this coverage level (Fig. 1*e*).

Also of importance, the common amino acids found in the toggling positions tend to recur in different regional viral epidemics and thus provide improved in vitro test sets to assess immune responses across clades or in the entire HIV group M sequences (Fig. 2). Furthermore, toggled amino acids often have similar side chain chemistry, probably because such changes are less disruptive to protein structure and function and hence are tolerated by the virus (Table II). Despite their chemical similarities, such conserved changes have been found to still permit immune escape, which is presumably why they recur (25). There are, however, exceptions to this: some combinations of amino acids in a given position actually have very distinctive side chain chemistry (Table II). Nonetheless, these substitutions must still be tolerated in the protein structure and function, reflecting the need of the pathogen to balance effective immune evasion with structural and functional integrity of the viral proteins. For example, a relatively variable position in p24, presumably evolving under immune pressure, is constrained to vary between the amino acids Asn, Ser, and Thr. These amino acids all have hydroxyl groups that can form hydrogen bonds resulting in the termination of an $\alpha$ helix (Fig. 3*a*); retention of an amino acid with the capacity to form a hydrogen bond in this variable position is found throughout the entire primate lentiviral tree. So while the virus is rapidly changing in this position, it is limited in the way it can mutate. Furthermore, T cell epitopes thread in and out of the folded protein with the mutable positions preferentially located at the outside of the protein, likely dictated by functional and structural limitations. This is supported by the observation that the variable positions that toggle are more often found on the surface of the proteins than within (Fig. 3, *b* and *c*, Gag p24, $p = 0.0007$ and reverse transcriptase, $p = 0.0109$, two-sample Kolmogorov-Smirnov test). Thus, positions that can vary and the extent of variation are dictated by the protein, providing natural constraints that ultimately make a toggled peptide approach to coverage feasible.

**FIGURE 1.** Distribution and extent of variable positions in Gag and the consequences in terms of number of variants included in B subtype toggle peptide reagents. In *a*, the bar height represents the proportion of aligned amino acids in clade B sequences that are identical to one of three sequences that can be used as a basis for peptide design: 1) a single B clade protein (HXB2, GenBank *K03455*; blue), 2) a B subtype consensus sequence (22) (red), 3) a clade B toggle design including up to the three most common amino acids in each variable position (green). Additional amino acids for the putative toggles (green bars) were included only if the alternate amino acids had a frequency ≥5% in the population (i.e., 5% cut-off in *e*). Each bar represents a single position in the alignment. Positions are in the same order as they are found in the protein, from the N to C terminus; these figures illustrate the relative dispersion of the variable positions. *b*, Individual amino acid positions were re-sorted along the *x*-axis based on the frequency of amino acids that are identical to HXB2, so that positions that have the highest level of identity with HXB2 (100%) are on the *left*, and those positions unique to HXB2 are on the *right*, illustrating both the relative frequency and levels of variation in variable positions. At each point, the proportional identity of the consensus (red) or the toggles (green) equals or exceeds that of HXB2 (blue). *c*, Sequence composition of exemplary toggled peptide 35: the first amino acid listed in the case shown, residues N4, E12 and R16 reflect the most common (consensus) amino acids in these positions. The residues in parentheses indicate positions in the peptide sequence where more than one amino acid was present in >5% of all the HIV clade B sequences in the Los Alamos HIV Sequence Database. *d*, Mass spectrometric analysis of the toggled peptide for OLP 35. Ten variant sequences produce discrete peaks, whereas two pairs of variants have masses that are too close for separation in the MS analysis (variants 3 and 9, 2186 vs 2187 kDa, and variants 4 and 10, 2172 and 2173 kDa). *e*, The number of single peptide sequences in toggled peptides depends on sequence conservation and the desired degree of viral population coverage: At 5% cut-off (gray bars), the relatively conserved p24 region shows only one toggled peptide that contains close to 100 variants, whereas toggled peptides in the p17 and p15 sequences can exceed 1000 variants. These numbers can be reduced significantly when the population coverage cut-off is set at a lower level (i.e., when only amino acids present in at least 10% (green bars), 15% (blue bars), or 20% (red bars) of sequences are included in the toggled peptide).

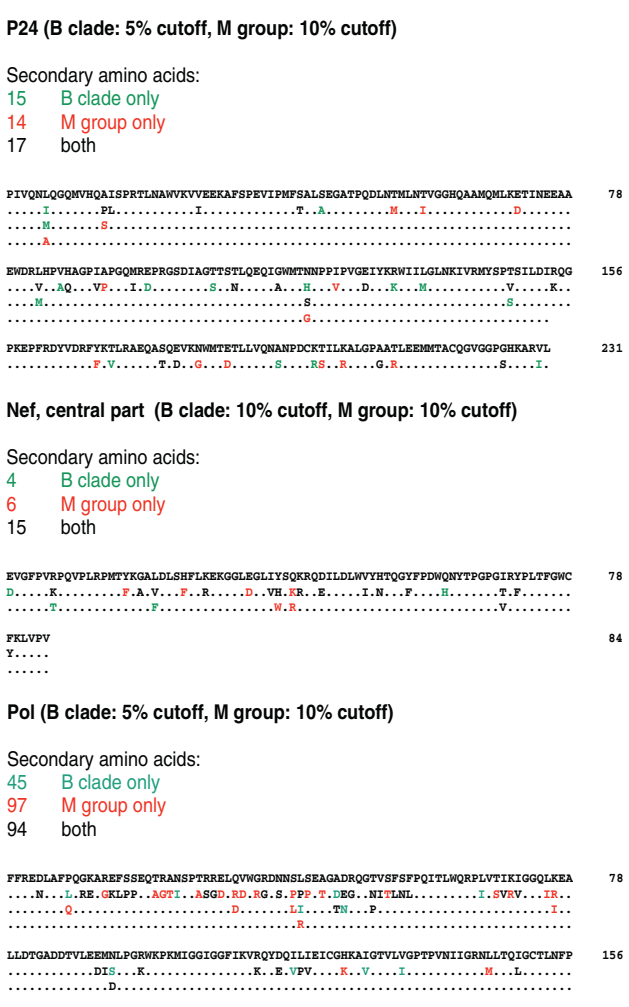



**FIGURE 2.** Los Alamos National Laboratory HIV database sequence alignments were used to calculate common amino acids for clade B and group M at each protein sequence position, for HIV-1 p24, Pol, and the central part of Nef. The p24 alignment contained 125 clade B sequences and 614 group M sequences, the Nef alignment contained 514 clade B sequences and 1291 group M sequences, and the Pol alignment contained 111 clade B sequences and 716 group M sequences. For the purposes of this study, a secondary amino acid was included for p24 and Pol clade B if it was present in at least 5% of the clade B sequences, whereas for the central part of Nef, a secondary amino acid was included if it was present in at least 10% of clade B sequences. Secondary group M amino acids are also shown to demonstrate additional changes needed for a group M toggled peptide set. For group M (for all three genes), an amino acid was included if it was present in at least 10% of the group M sequences. Secondary amino acids are shown in relation to clade B consensus, dots stand for cases when no secondary amino acid was present at the preset percentage of sequences. Secondary amino acids shown in green stand for those amino acids that were present at the required percentage in the clade B sequence alignment only, but not in the group M sequence alignment. Secondary amino acids shown in red stand for those amino acids that were present at the required percentage in the group M sequence alignment only, but not in the clade B sequence alignment. Secondary amino acids shown in black stand for those amino acids that were present at the required percentage in both the clade B sequence alignment and in the group M sequence alignment.

### *Toggled peptide design*

For the initial proof-of-concept studies, sets of toggled peptides covering variation within clade B in the immunologically most highly recognized regions of the virus were used (12): Gag p24, Pol, and the central region of Nef. Non-consensus amino acid (s) that were found in at least 5% of all clade B sequences in the 2004 HIV database were included in the synthesis step for peptides spanning the Gag and Pol sequences. In cases where the HIV database contained multiple sequences per subject, only one sequence per individual was included in the alignment. For the more variable central portion of Nef this cut-off was set at 10%. Using these criteria, all but four positions had >90% population coverage by including one, two, or three amino acids. Most variable positions required only two alternative amino acids (Table II), and of the 178 peptides spanning the three proteins, 25 (14%) were so conserved that the consensus sequence provided adequate coverage. In these cases, no separate toggled counterparts to the consensus sequence were synthesized. All toggled peptides and their respective consensus counterpart, as well as the number of variants included in each toggle preparation, are summarized in Table III.

The B consensus sequence was selected for a comparison to the toggled peptide for several reasons. Despite being a non-natural sequence, recent data show that recombinant proteins based on consensus sequences retain the structural, functional and immunogenic properties of circulating HIV-1 (26–28). In addition, several primary considerations for comparing responses of toggled peptides to a consensus based peptide set include: first, peptides based on consensus sequences are the most widely used ones to screen for HIV-specific responses, and so the current standard. Second, peptides based on a centralized sequence are by definition contained in the toggled peptide—natural sequences will always contain rare amino acids in some positions that would not be included in toggle design at the 5% threshold. Third, as shown in Fig. 1*a*), consensus based peptides provide much better coverage of circulating strains and thus can be expected to detect more responses than peptides based on a natural strain. Fourth, although full length protein HIV population consensus sequences are not found in nature, at least at the level of sampling that is possible, over the length a peptide, the consensus is essentially always found in multiple natural strains in this set. So consensus peptides are in fact as "natural" as peptides based on a real sequence, making the consensus-based peptide set the most appropriate comparison set.

Table II. *Amino acids combinations found in the toggled positions*[a]

| Non-Polar | | Charged | | Polar | | Polar /Non-Polar | | Polar /Charged | | Other | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 21 | IV | 20 | KR +/+ | 6 | *S*N | 6 | *T*A | 5 | GR | 1 | QLE |
| 6 | IL | 15 | DE −/− | 6 | *ST* | 3 | **YF** | 2 | **Y**H | 1 | R*S*P |
| 4 | **F**L | 4 | EK −/+ | 3 | *S*C | 3 | *S*A | 2 | GE | 1 | IVE |
| 2 | ILV | | | 2 | *S*G | 2 | GA | 2 | QH | 1 | AD*T* |
| 2 | IM | | | | | 2 | SP | 2 | QE | | |
| 2 | LP | | | | | 2 | *T*I | 2 | ND | | |
| 2 | AP | | | | | 2 | *T*IV | 2 | *S*R | | |
| 2 | AV | | | | | 2 | *ST*A | 1 | N*S*D | | |
| 1 | MV | | | | | 1 | *S*L | 1 | *T*K | | |
| 1 | **F**LV | | | | | 1 | *S*LV | 1 | NK | | |
| 1 | LM | | | | | 1 | *T*MV | 1 | HN | | |
| | | | | | | 1 | *T*NP | 1 | QLE | | |
| | | | | | | 1 | *ST*P | | | | |
| | | | | | | 1 | NI | | | | |
| | | | | | | 1 | *T*NA | | | | |
| | | | | | | 1 | *T*AI | | | | |

[a] Toggling amino acid combinations occurring at variable positions are shown, along with the number of times each individual combination was observed in the HIV sequence database. Amino acids in **bold** have aromatic side chains (F, Y, and W), those with aliphatic hydroxyl side chains (S and, T) are indicated in *italics.*

### *Toggled peptide sets yield more positive responses than consensus peptides*

Total T cell responses were tested in 17 chronically HIV clade B infected individuals using an IFN-γ ELISpot assay and either toggled or consensus peptides in duplicate wells (12). Applying a GLM (see *Materials and Methods*), we found a significantly higher probability for a positive response using the toggled peptides than the consensus (odds ratio 1.33, $p = 0.0019$). These data are illustrated in Fig. 4*a*, where the number of positive responses for each patient is shown. Although some responses were detected by either the consensus or the toggled peptides only, significantly more reactions were detected using the toggled peptides than the consensus peptides in 14 of 17 patients ($p = 0.0038$, Wilcoxon test). In 2 of 17 cases the number of responses to consensus and toggled peptides were identical, and in only one patient did the consensus OLP peptides elicit more reactions than the toggled peptide set. Thus, although some consensus peptide responses may be lost, there is a significant net gain of responses detected using toggled peptides.

To investigate whether the detection of CD4 T cell responses would be similarly enhanced by the use of toggled peptides, 17 samples were obtained from HIV infected individuals who either spontaneously controlled HIV replication to levels below 2000 RNA copies/ml in the absence of treatment ("controllers") or who were at an early stage of HIV infection with high CD4 counts. Subjects fulfilling these criteria generally have stronger CD4 T cell responses to HIV than people with cronic disease (29), whereas the responses in the previous total T cell analyses were likely mediated largely by CD8 T cell responses (13, 30). Using CD8 depleted PBMC, these individuals were only tested against the subset of

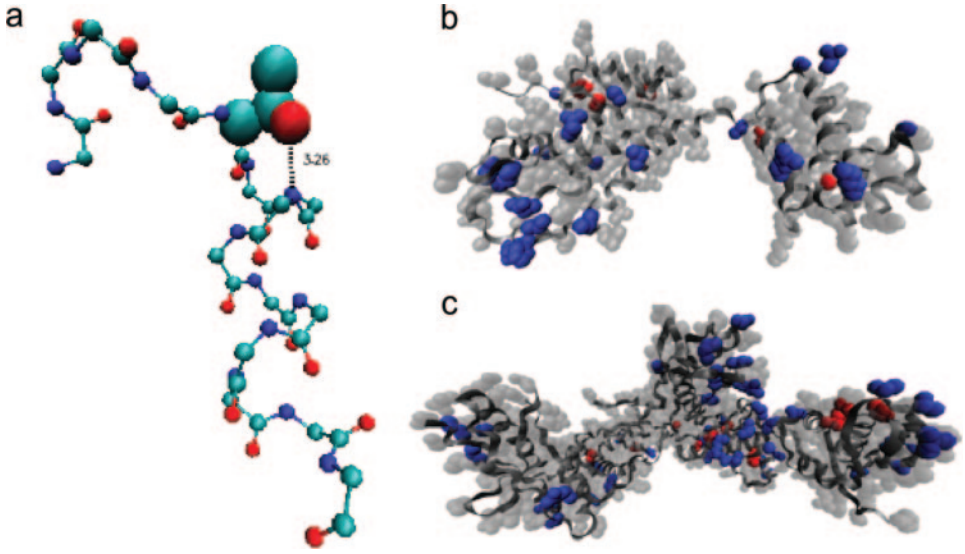



**FIGURE 3.** Location of toggling residues relative to conserved positions shown on the protein structure for Gag-p24 (*a* and *b*) and reverse transcriptase (*c*). Simple reasons for the particular residues observed to toggle are often evident. For example, *a* shows a position in Gag-p24, which caps a short helix, with clade B toggling residues N and T; the hydrogen bonding capability of the side chains of these two amino acids determines the end of the α helix. Toggling in sequences across the entire M group and the primate lentiviruses also allows serine residues, toggling between N, S, and T, but the hydroxyl group is always preserved. For *b* and *c*, the protein backbone is shown in a ribbon representation, and the side chains are illustrated as van der Waals spheres. Nontoggling residues are conserved (each have a consensus that is present in >90% of B clade sequences) and are shown in gray, whereas positions that exhibit toggling behavior in clade B are colored. Residues deemed by inspection to be on the surface are shown in blue, while buried residues are shown in red. A preference for toggling residues to be on the surface is evident.

Table III. *Consensus and toggled peptide sequences*

| | Consensus OLP | | Toggled Peptide[a] | Variants /Toggle |
|---|---|---|---|---|
| 18 | SQVSQNYPIVQNLQGQMV | To-18 | (SN)(QPK)VSQNYPIVQN (LIM) QGQMV | 18 |
| 19 | IVQNLQGQMVHQAISPR | To-19 | IVQN (LIM) QGQMVHQ (AP) (IL) SPR | 12 |
| 20 | QMVHQAISPRTLNAWVKV | To-20 | QMVHQ (AP) (IL) SPRTLNAWVKV | 4 |
| 21 | PRTLNAWVKVVEEKAF | To-21 | PRTLNAWVKV (VI) EEKAF | 2 |
| 22 | WVKVVEEKAFSPEVIPMF | To-22 | WVKV (VI) EEKAFSPEVIPMF | 2 |
| 23 | AFSPEVIPMFSALSEGA | To-23 | AFSPEVIPMF (ST) AL (SA) EGA | 4 |
| 24 | PMFSALSEGATPQDLNTM | To-24 | PMF (ST) AL (SA) EGATPQDLNTM | 4 |
| 25 | GATPQDLNTMLNTVGGH | | | |
| 26 | NTMLNTVGGHQAAMQMLK | | | |
| 27 | GGHQAAMQMLKETINEEA | | | |
| 28 | LKETINEEAAEWDRLHPV | To-28 | LKETINEEAAEWDR (LVM) HP (VA) | 6 |
| 29 | AAEWDRLHPVHAGPIA | To-29 | AAEWDR (LVM) HP (VA) (HQ) AGP (IV) A | 24 |
| 30 | LHPVHAGPIAPGQMREPR | To-30 | (LVM) HP (VA) (HQ) AGP (IV) APGQ (MI) R (ED) PR | 96 |
| 31 | IAPGQMREPRGSDIA | To-31 | (IV) APGQ (MI) R (ED) PRGSDIA | 8 |
| 32 | MREPRGSDIAGTTSTL | To-32 | (MI) R (ED) PRGSDIAG (TS) TS (TN) L | 16 |
| 33 | SDIAGTTSTLQEQIGWM | To-33 | SDIAG (TS) TS (TN) LQEQI (GA)WM | 8 |
| 34 | STLQEQIGWMTNNPPIPV | To-34 | S (TN) LQEQI (GA)WMT (NHS) NPPIPV | 12 |
| 35 | WMTNNPPIPVGEIYKRWI | To-35 | WMT (NHS) NPPIPVG (ED) IYK (RK) WI | 12 |
| 36 | PVGEIYKRWIILGLNKIV | To-36 | PVG (ED) IYK (RK) WII (LM) GLNKIV | 8 |
| 37 | WIILGLNKIVRMYSPTSI | To-37 | WII (LM) GLNKIVRMYSP (TVS) SI | 6 |
| 38 | IVRMYSPTSILDIRQGPK | To-38 | IVRMYSP (TVS) SILDI (RK) QGPK | 6 |
| 39 | SILDIRQGPKEPFRDYV | To-39 | SILDI (RK) QGPKEPFRDYV | 2 |
| 40 | GPKEPFRDYVDRFYKTLR | To-40 | GPKEPFRDYVDRFYK (TV) LR | 2 |
| 41 | YVDRFYKTLRAEQASQEV | To-41 | YVDRFYK (TV) LRAEQA (ST) Q (ED) V | 8 |
| 42 | LRAEQASQEVKNWMTETL | To-42 | LRAEQA (ST) Q (ED) VKNWMTETL | 4 |
| 43 | EVKNWMTETLLVQNA | To-43 | (ED) VKNWMTETLLVQN (AS) | 4 |
| 44 | MTETLLVQNANPDCKTIL | To-44 | MTETLLVQN (AS) NPDC (KR)TIL | 4 |
| 45 | NANPDCKTILKALGPAA | To-45 | N (AS) NPDC (KR)TILKALGP (AG) A | 8 |
| 46 | TILKALGPAATLEEMMTA | To-46 | TILKALGP (AG) ATLEEMMTA | 2 |
| 47 | AATLEEMMTACQGVGGPGH | To-47 | (AG) ATLEEMMTACQGVGGP (GS) H | 4 |
| 48 | ACQGVGGPGHKARVLAEA | To-48 | ACQGVGGP (GS) HKAR (VI) LAEA | 4 |
| 76 | EVGFPVRPQVPLRPMTYK | To-076 | [ED]VGFPV[RKT]PQVPLRPMTYK | 6 |
| 77 | QVPLRPMTYKAAVDLSHF | To-077 | QVPLRPMTYK[AG]A[LVF]DLSHF | 6 |
| 78 | YKAAVDLSHFLKEKGGL | To-078 | YK[AG]A[LVF]DLSHFL[KR]EKGGL | 12 |
| 79 | SHFLKEKGGLEGLIYSQK | To-079 | SHFL[KR]EKGGLEGL[IV][YH]SQ[KR] | 16 |
| 80 | GLEGLIYSQKRQDILDLW | To-080 | GLEGL[IV][YH]SQ[KR]RQ[DE]ILDLW | 16 |
| 81 | QKRQDILDLWVYHTQGYF | To-081 | Q[KR]RQ[DE]ILDLW[VI]Y[HN]TQG[YF]F | 32 |
| 82 | LWVYHTQGYFPDWQNY | To-082 | LW[VI]Y[HN]TQG[YF]FPDW[QH]NY | 16 |
| 83 | QGYFPDWQNYTPGPGIRY | To-083 | QG[YF]FPDW[QH]NYTPGPG[ITV]R[YF] | 24 |
| 84 | NYTPGPGIRYPLTFGWCF | To-084 | NYTPGPG[ITV]R[YF]PLTFGWCF | 6 |
| 85 | RYPLTFGWCFKLVPV | To-085 | R[YF]PLTFGWC[YF]KLVPV | 4 |
| 145 | FFREDLAFPQGKAREF | To-145 | FFRE(DN)LAF(PL)Q(GR)(KE)AR(EK)(FL) | 64 |
| 146 | AFPQGKAREFSSEQTRA | To-146 | AF(PL)Q(GR)(KE)AR(EK)(FL)(SP)(SP)EQTRA | 128 |
| 147 | REFSSEQTRANSPTRREL | To-147 | R(EK)(FL)(SP)(SP)EQTRA(NI)SPT(RS)(RG)EL | 128 |
| 148 | RANSPTRRELQVWGR | To-148 | RA(NI)SPT(RS)(RG)ELQVWG(RG) | 16 |
| 149 | TRRELQVWGRDNNSLSEA | To-149 | T(RS)(RG)ELQVWG(RG)D(NS)NS(LPI)SEA | 48 |
| 150 | GRDNNSLSEAGADRQGTV | To-150 | G(RG)D(NS)NS(LPI)SEAG(ADT)(DEN)(RG)QG(TNP)(VI) | 1296 |
| 151 | EAGADRQGTVSFSFPQI | To-151 | EAG(ADT)(DEN)(RG)QG(TNP)(VI)S(FL)(SN)(FL)PQI | 864 |
| 152 | GTVSFSFPQITLWQRPLV | To-152 | G(TNP)(VI)S(FL)(SN)(FL)PQITLWQRP(LI)V | 96 |
| 153 | QITLWQRPLVTIKIGGQL | To-153 | QITLWQRP(LI)VT(IV)K(IV)GGQL | 8 |
| 154 | LVTIKIGGQLKEALL | To-154 | (LI)VT(IV)K(IV)GGQLKEALL | 8 |
| 155 | IGGQLKEALLDTGADDTV | To-155 | (IV)GGQLKEALLDTGADDTV | 2 |
| 156 | LLDTGADDTVLEEMNL | To-156 | LLDTGADDTVLE(ED)(MI)(NSD)L | 12 |
| 157 | DDTVLEEMNLPGRWKPKM | To-157 | DDTVLE(ED)(MI)(NSD)LPG(RK)WKPKM | 24 |
| 158 | NLPGRWKPKMIGGIGGFI | To-158 | (NSD)LPG(RK)WKPKMIGGIGGFI | 6 |
| 159 | KMIGGIGGFIKVRQYDQI | To-159 | KMIGGIGGFIKV(RK)QY(DE)Q(IV) | 8 |
| 160 | FIKVRQYDQILIEICGHK | To-160 | FIKV(RK)QY(DE)Q(IV)(LP)(IV)EICGHK | 32 |
| 161 | QILIEICGHKAIGTVLV | To-161 | Q(IV)(LP)(IV)EICGHKA(IV)GTVL(VI) | 32 |
| 162 | GHKAIGTVLVGPTPVNII | To-162 | GHKA(IV)GTVL(VI)GPTPVNII | 4 |
| 163 | LVGPTPVNIIGRNLLTQI | To-163 | L(VI)GPTPVNIIGRNLLTQ(IL) | 4 |
| 164 | IIGRNLLTQIGCTLNFPI | To-164 | IIGRNLLTQ(IL)GCTLNFPI | 2 |
| 165 | QIGCTLNFPISPIETVPV | To-165 | Q(IL)GCTLNFPISPI(ED)TVPV | 4 |
| 166 | PISPIETVPVKLKPGM | To-166 | PISPI(ED)TVPVKLKPGM | 2 |
| 167 | TVPVKLKPGMDGPKVKQW | | | |
| 168 | GMDGPKVKQWPLTEEKIK | | | |
| 169 | QWPLTEEKIKALVEI | To-169 | QWPLTEEKIKAL(VI)EI | 2 |
| 170 | EEKIKALVEICTEMEK | To-170 | EEKIKAL(VI)EICTEMEK | 2 |
| 171 | LVEICTEMEKEGKISKI | To-171 | L(VI)EICTEMEKEGKISKI | 2 |
| 172 | MEKEGKISKIGPENPY | | | |
| 173 | ISKIGPENPYNTPVFAIK | | | |
| 174 | PYNTPVFAIKKKDSTKWR | To-174 | PYNTPVFAIKKKDST(KR)WR | 2 |
| 175 | IKKKDSTKWRKLVDFREL | To-175 | IKKKDST(KR)WRKLVDFREL | 2 |

(*Table continues*)

Table III. (*Continued*)

| | Consensus OLP | | Toggled Peptide[a] | Variants /Toggle |
|---|---|---|---|---|
| 176 | WRKLVDFRELNKRTQDFW | To-176 | WRKLVDFRELNK(RK)TQDFW | 2 |
| 177 | ELNKRTQDFWEVQLGIPH | To-177 | ELNK(RK)TQDFWEVQLGIPH | 2 |
| 178 | FWEVQLGIPHPAGLKKKK | To-178 | FWEVQLGIPHP(AS)GLKKKK | 2 |
| 179 | PHPAGLKKKKSVTVLDV | To-179 | PHP(AS)GLKKKKSVTVLDV | 2 |
| 180 | KKKSVTVLDVGDAYFSV | | | |
| 181 | LDVGDAYFSVPLDKDFRK | To-181 | LDVGDAYFSVPLD(KE)(DE)FRK | 4 |
| 182 | SVPLDKDFRKYTAFTI | To-182 | SVPLD(KE)(DE)FRKYTAFTI | 4 |
| 183 | DFRKYTAFTIPSINNETPGI | To-183 | (DE)FRKYTAFTIPS(IVT)NNETPG(IV) | 12 |
| 184 | PSINNETPGIRYQYNVL | To-184 | PS(IVT)NNETPG(IV)RYQYNVL | 6 |
| 185 | PGIRYQYNVLPQGWK | To-185 | PG(IV)RYQYNVLPQGWK | 2 |
| 186 | QYNVLPQGWKGSPAIF | | | |
| 187 | QGWKGSPAIFQSSMTKIL | To-187 | QGWKGSPAIFQ(SC)SMT(KR)IL | 4 |
| 188 | IFQSSMTKILEPFRK | To-188 | IFQ(SC)SMT(KR)ILEPFRK | 4 |
| 189 | MTKILEPFRKQNPDIVIY | To-189 | MT(KR)ILEPFRKQNP(DE)(IL)VIY | 8 |
| 190 | RKQNPDIVIYQYMDDLYV | To-190 | RKQNP(DE)(IL)VIYQYMDDLYV | 4 |
| 191 | IYQYMDDLYVGSDLEI | | | |
| 192 | DLYVGSDLEIGQHRTKI | To-192 | DLYVGSDLEI(GE)QHR(TAI)KI | 6 |
| 193 | LEIGQHRTKIEELRQHLL | To-193 | LEI(GE)QHR(TAI)KIEELRQHLL | 6 |
| 194 | KIEELRQHLLRWGFTTPDK | To-194 | KIEELRQHLL(RK)WG(FL)TTPDK | 4 |
| 195 | LRWGFTTPDKKHQKEPPF | To-195 | L(RK)WG(FL)TTPDKKHQKEPPF | 4 |
| 196 | DKKHQKEPPFLWMGYELH | | | |
| 197 | PFLWMGYELHPDKWTV | | | |
| 198 | YELHPDKWTVQPIVLPEK | To-198 | YELHPDKWTVQPI(VM)LPEK | 2 |
| 199 | TVQPIVLPEKDSWTVNDI | To-199 | TVQPI(VM)LPEKDSWTVNDI | 2 |
| 200 | EKDSWTVNDIQKLVGKL | | | |
| 201 | NDIQKLVGKLNWASQIYA | To-201 | NDIQKLVGKLNWASQIY(AP) | 2 |
| 202 | KLNWASQIYAGIKVKQL | To-202 | KLNWASQIY(AP)GIKV(KR)(QE)L | 8 |
| 203 | IYAGIKVKQLCKLLRGTK | To-203 | IY(AP)GIKV(KR)(QE)LCKLLRG(TA)K | 16 |
| 204 | QLCKLLRGTKALTEVIPL | To-204 | (QE)LCKLLRG(TA)KALTEV(VI)PL | 8 |
| 205 | TKALTEVIPLTEEAELEL | To-205 | (TA)KALTEV(VI)PLT(EK)EAELEL | 8 |
| 206 | PLTEEAELELAENREILK | To-206 | PLT(EK)EAELELAENREILK | 2 |
| 207 | ELAENREILKEPVHGVYY | | | |
| 208 | LKEPVHGVYYDPSKDLIA | To-208 | LKEPVHGVYYDPSK(DE)L(IV)A | 4 |
| 209 | YYDPSKDLIAEIQKQGQGQW | To-209 | YYDPSK(DE)L(IV)AE(ILV)QKQG(QLE)GQW | 36 |
| 210 | EIQKQGQGQWTYQIY | To-210 | E(ILV)QKQG(QLE)GQWTYQIY | 9 |
| 211 | GQGQWTYQIYQEPFKNLK | To-211 | G(QLE)GQWTYQIYQEP(FY)KNLK | 6 |
| 212 | IYQEPFKNLKTGKYARMR | To-212 | IYQEP(FY)KNLKTGKYA(RK)(MTV)(RK) | 24 |
| 213 | LKTGKYARMRGAHTNDVK | To-213 | LKTGKYA(RK)(MTV)(RK)G(AT)HTNDV(KR) | 48 |
| 214 | MRGAHTNDVKQLTEAVQK | To-214 | (MTV)(RK)G(AT)HTNDV(KR)QLTEAVQK | 24 |
| 215 | VKQLTEAVQKIATESIVI | To-215 | V(KR)QLTEAVQK(IV)(AT)TE(SC)IVI | 16 |
| 216 | QKIATESIVIWGKTPKFK | To-216 | QK(IV)(AT)TE(SC)IVIWGK(TI)PKF(KR) | 32 |
| 217 | VIWGKTPKFKLPIQKETW | To-217 | VIWGK(TI)PKF(KR)LPIQKETW | 4 |
| 218 | FKLPIQKETWEAWWTEYW | To-218 | F(KR)LPIQKETW(ED)(AT)WW(TM)EYW | 16 |
| 219 | TWEAWWTEYWQATWIPEW | To-219 | TW(ED)(AT)WW(TM)EYWQATWIPEW | 8 |
| 220 | YWQATWIPEWEFVNRPPL | | | |
| 221 | EWEFVNRPPLVKLWYQL | | | |
| 222 | PPLVKLWYQLEKEPIVGA | To-222 | PPLVKLWYQLEK(ED)PI(VIE)GA | 6 |
| 223 | QLEKEPIVGAETFYVDGA | To-223 | QLEK(ED)PI(VIE)GAETFYVDGA | 6 |
| 224 | GAETFYVDGAANRETKL | To-224 | GAETFYVDGAANRETK(LS) | 2 |
| 225 | DGAANRETKLGKAGYV | To-225 | DGAANRETK(LS)GKAGYV | 2 |
| 226 | ETKLGKAGYVTDRGRQKV | To-226 | ETK(LS)GKAGYVT(DN)(RK)GRQKV | 8 |
| 227 | YVTDRGRQKVVSLTDTTNQK | To-227 | YVT(DN)(RK)GRQKVV(SPT)(LI)(TA)DTTNQK | 48 |
| 228 | VSLTDTTNQKTELQAIHL | To-228 | V(SPT)(LI)(TA)DTTNQKTELQAI(HY)L | 24 |
| 229 | QKTELQAIHLALQDSGL | To-229 | QKTELQAI(HY)LALQDSG(LSV) | 6 |
| 230 | IHLALQDSGLEVNIV | To-230 | I(HY)LALQDSG(LSV)EVNIV | 6 |
| 231 | QDSGLEVNIVTDSQYAL | To-231 | QDSG(LSV)EVNIVTDSQYAL | 3 |
| 232 | NIVTDSQYALGIIQA | | | |
| 233 | SQYALGIIQAQPDKSESEL | To-233 | SQYALGIIQAQPD(KR)SESE(LI) | 4 |
| 234 | AQPDKSESELVSQIIEQL | To-234 | AQPD(KR)SESE(LI)V(SN)QIIEQL | 8 |
| 235 | ELVSQIIEQLIKKEKVYL | To-235 | E(LI)V(SN)QIIEQLIKKEKVYL | 4 |
| 236 | QLIKKEKVYLAWVPAHK | | | |
| 237 | VYLAWVPAHKGIGGNEQV | To-237 | VYLAWVPAHKGIGGNEQ(VI) | 2 |
| 238 | HKGIGGNEQVDKLVSAGI | To-238 | HKGIGGNEQ(VI)DKLVS(AST)GI | 6 |
| 239 | QVDKLVSAGIRKVLFL | To-239 | Q(VI)DKLVS(AST)GIR(KR)VLFL | 12 |
| 240 | SAGIRKVLFLDGIDKA | To-240 | S(AST)GIR(KR)VLFLDGIDKA | 6 |
| 241 | VLFLDGIDKAQEEHEKYH | To-241 | VLFLDGIDKAQ(ED)(ED)HEKYH | 4 |
| 242 | KAQEEHEKYHSNWRAMA | To-242 | KAQ(ED)(ED)HEKYH(SN)NWRAMA | 8 |
| 243 | KYHSNWRAMASDFNLPPV | To-243 | KYH(SN)NWRAMASDFNLPP(VI) | 4 |
| 244 | MASDFNLPPVVAKEIVA | To-244 | MASDFNLPP(VI)VAKEIVA | 2 |
| 245 | PPVVAKEIVASCDKCQLK | To-245 | PP(VI)VAKEIVA(SC)CDKCQLK | 4 |
| 246 | VASCDKCQLKGEAMHGQV | To-246 | VA(SC)CDKCQLKGEAMHGQV | 2 |
| 247 | LKGEAMHGQVDCSPGIW | | | |

(*Table continues*)

Table III. (*Continued*)

| | Consensus OLP | | Toggled Peptide[a] | Variants /Toggle |
|---|---|---|---|---|
| 248 | GQVDCSPGIWQLDCTHL | | | |
| 249 | GIWQLDCTHLEGKIILVA | To-249 | GIWQLDCTHLEGK(IV)ILVA | 2 |
| 250 | HLEGKIILVAVHVASGYI | To-250 | HLEGK(IV)ILVAVHVASGYI | 2 |
| 251 | VAVHVASGYIEAEVIPA | To-251 | VAVHVASGYIEAEVIP(AT) | 2 |
| 252 | GYIEAEVIPAETGQETAY | To-252 | GYIEAEVIP(AT)ETGQ(ED)TAY | 4 |
| 253 | PAETGQETAYFLLKLAGR | To-253 | P(AT)ETGQ(ED)TAYF(LI)LKLAGR | 8 |
| 254 | AYFLLKLAGRWPVKTIH | To-254 | AYF(LI)LKLAGRWPV(KT)T(IV)H | 8 |
| 255 | AGRWPVKTIHTDNGSNF | To-255 | AGRWPV(KT)T(IV)HTDNG(SRP)NF | 12 |
| 256 | TIHTDNGSNFTSTTVKAA | To-256 | T(IV)HTDNG(SRP)NF(TI)S(TAN)(TAS)VKAA | 108 |
| 257 | NFTSTTVKAACWWAGIK | To-257 | NF(TI)S(TAN)(TAS)VKAACWWAGIK | 18 |
| 258 | KAACWWAGIKQEFGIPY | | | |
| 259 | GIKQEFGIPYNPQSQGVV | To-259 | GIKQEFGIPYNPQSQGV(VI) | 2 |
| 260 | PYNPQSQGVVESMNKELK | To-260 | PYNPQSQGV(VI)ES(MI)N(KN)ELK | 8 |
| 261 | VVESMNKELKKIIGQVR | To-261 | V(VI)ES(MI)N(KN)ELKKIIGQVR | 8 |
| 262 | ELKKIIGQVRDQAEHLK | | | |
| 263 | QVRDQAEHLKTAVQMAVF | | | |
| 264 | LKTAVQMAVFIHNFKRK | | | |
| 265 | AVFIHNFKRKGGIGGYSA | To-265 | AVFIHNFKRKGGIG(GE)YSA | 2 |
| 266 | RKGGIGGYSAGERIVDII | To-266 | RKGGIG(GE)YSAGERI(VI)DII | 4 |
| 267 | SAGERIVDIIATDIQTK | To-267 | SAGERI(VI)DIIA(TS)DIQT(KR) | 8 |
| 268 | DIIATDIQTKELQKQITK | To-268 | DIIA(TS)DIQT(KR)ELQKQITK | 4 |
| 269 | TKELQKQITKIQNFRVYY | To-269 | T(KR)ELQKQITKIQNFRVYY | 2 |
| 270 | TKIQNFRVYYRDSRDPLW | To-270 | TKIQNFRVYYRD(SN)RDP(LIV)W | 6 |
| 271 | YYRDSRDPLWKGPAKLLW | To-271 | YYRD(SN)RDP(LIV)WKGPAKLLW | 6 |
| 272 | LWKGPAKLLWKGEGAVVI | To-272 | (LIV)WKGPAKLLWKGEGAVVI | 3 |
| 273 | LWKGEGAVVIQDNSDIKV | To-273 | LWKGEGAVVIQDNS(DE)IKV | 2 |
| 274 | VIQDNSDIKVVPRRKAKI | To-274 | VIQDNS(DE)IKVVPRRK(AV)KI | 4 |
| 275 | KVVPRRKAKIIRDYGKQM | To-275 | KVVPRRK(AV)KIIRDYGKQM | 2 |
| 276 | KIIRDYGKQMAGDDCVA | | | |
| 277 | KQMAGDDCVASRQDED | To-277 | KQMAGDDCVA(SG)RQDED | 2 |

[a] Amino acids in parentheses correspond to amino acid mixes added at the synthesis step.

peptides that spanned p24 and the center of Nef due to the reduced yields of CD8 depleted PBMC. The statistical results for these samples using the GLM model paralleled the results obtained for the total T cell responses above, as the chance to detect responses using toggled peptides was twice as high compared with using consensus based OLP (odds ratio 2.14, $p = 0.0025$ and Fig. 4*b*).

*Responses to the toggled peptides are of greater magnitude than responses to consensus OLP*

We next assessed whether toggled peptides not only detected more, but also stronger *in vitro* responses. The magnitude of responses was plotted as magnitude in reaction to either the consensus OLP or the corresponding toggled peptide, and treated as paired events (Fig. 5). Responses where both OLP and toggled peptides reacted were separated from those where only the toggled peptides or the consensus OLP was positive. This approach showed overall significantly stronger responses using toggled peptides relative to the consensus OLP for both the total PBMC ELISpot (Fig. 5*a*) and the CD4 T cell ELISpot (Fig. 5*b*) comparisons ($p = 2.8 \times 10^{-2}$ and $p = 4.1 \times 10^{-5}$, respectively, Wilcoxon test). In addition, the number of the increased responses was significantly greater for the toggled peptides than the consensus when only one of the two peptide test sets scored positive (total T cells: $p = 7.3 \times 10^{-8}$, CD4 T cells: $p = 3.3 \times 10^{-6}$, using a 1-sample proportions test). When both test sets elicited positive responses, only the total T cell ELISpot showed significantly more samples with gain in responses among the toggled peptides (total T cells: $p = 0.0013$, CD4 T cells: $p = 0.33$).

*Increased response rates are due to a better match of the toggle to the autologous sequence*

Toggled peptides could give a stronger response relative to the consensus if T cells are more efficiently triggered by a peptide variant sequence that is found among the toggled peptides but not the consensus. To better resolve how increased coverage of autologous sequences contributes to increased breadth and magnitude of responses, autologous HIV sequences from the study subjects were compared with the consensus sequence and to the toggled peptides (Fig. 6). The magnitude of detected responses was significantly higher when the match with the autologous sequence was improved by the toggles, relative to those cases where the consensus was as close to the autologous sequences as the toggles ($p = 0.0054$, Wilcoxon test, Fig. 6*a*). There were fewer sequences available for this comparison in CD4 data, as only Gag and Nef peptides were tested for CD4 T cell responses (Fig. 6*b*). There was a trend toward greater magnitude of responses elicited by toggles than consensus peptides, which did not, however, reach statistical significance ($p = 0.1288$).

## Discussion

Using the novel concept of building sequence diversity into in vitro Ag test sets, we have shown that toggled peptides represent a cost-effective way to more accurately capture the extent of the virus-specific T cell immune response to a highly variable pathogen such as HIV.

The benefits of such more comprehensive assessments of HIV-specific T cell responses may be crucial in evaluating the magnitude and breadth of T cell responses at the population level. The rank ordering of number of responses per individuals and the magnitude of responses to different peptides both differ between toggle and consensus (Fig. 4), which would impact the analyses of potential associations between breadth or magnitude of the virus-specific response and in vivo viral control (12, 13, 31). Just as assessing the T cell immune response by IFN-$\gamma$ expression alone gives only a limited view of T cell function (32), using only one peptide to probe T cell specificity fails to adequately detect T cell

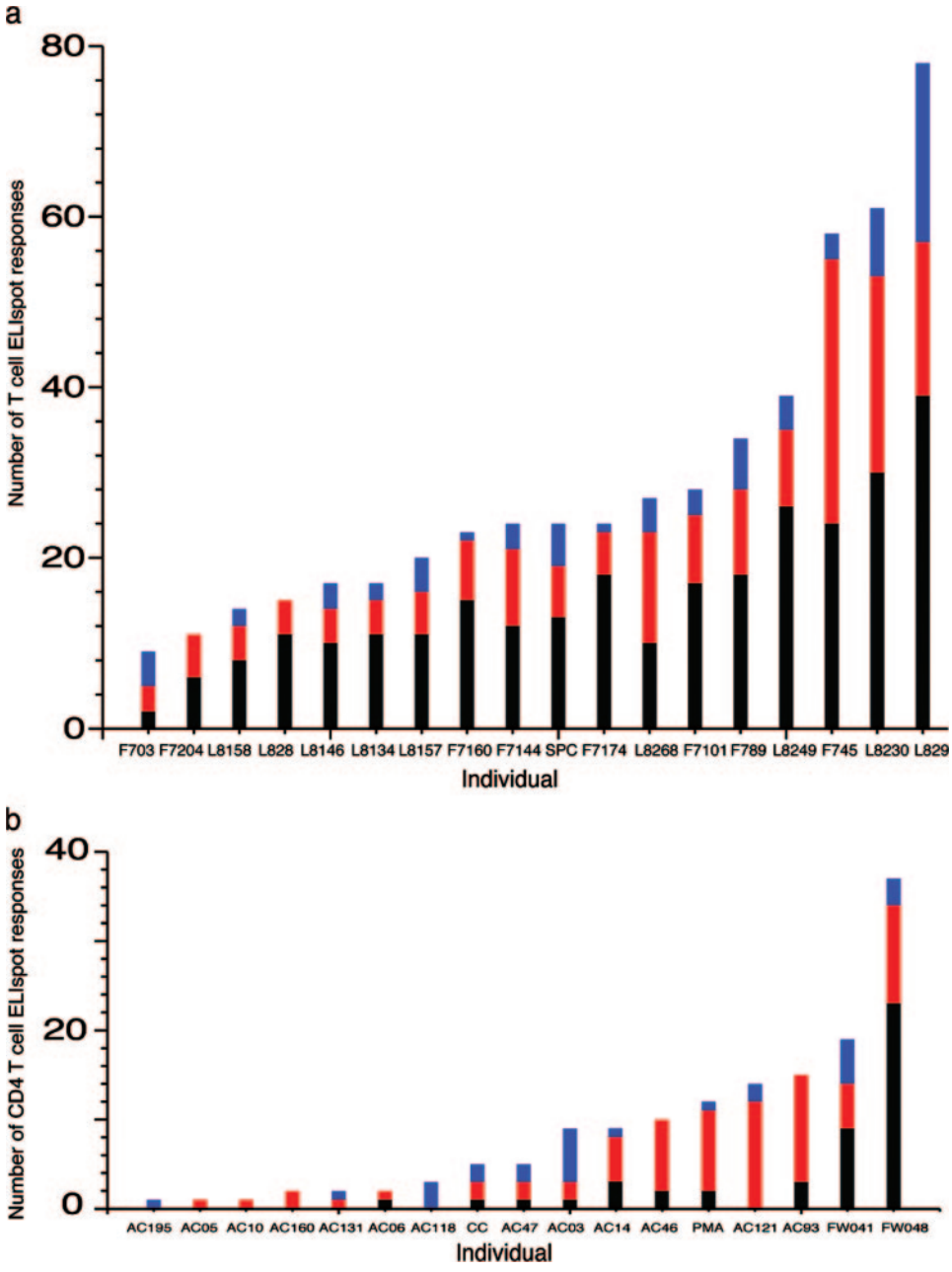



**FIGURE 4.** T cell responses to different peptides, by individual. *a*, The number of peptides that elicited a positive ELISpot using total T cells from 17 different individuals. *b*, The number of peptides that elicited a positive ELISpot response using isolated $CD4^+$ T cells sampled from 17 different individuals. The total number of peptides that elicited a response is indicated by the total height of the bar for each patient. The black portion accounts for responses to peptides for which both the consensus and the toggled counterpart elicited a positive response. Responses that were detected only by the toggled mixture but not when using the counterpart consensus OLP are shown in red; responses that were detected only using the consensus OLP but not the toggle mixture are shown in blue. Significantly more responses were detected using toggled peptides for total ($p = 0.0038$) and CD4 T cells ($p = 0.04194$; both $p$ values are based on Wilcoxon test). Toggled and consensus peptides comprised Gag p24 (amino positions in the HXB2 reference strain 1–231), Pol (HXB2 positions 1–1003), and the central region of Nef (HXB2 Nef positions 65–148) for *a*, and Gag p24 and the central region of Nef only for *b*.

responses. Although autologous peptides are frequently considered the most appropriate test set, peptide test sets reflecting the individuals' autologous viral sequences are prohibitively expensive (33), and given the changing nature of the immune response and viral sequence over time in a given individual, may only allow a limited view of the true extent of the immune response. Thus, toggled peptides may even outperform autologous peptides because they predominantly probe immune responses generated by the incoming virus, but also cover possible escape variants that may have induced de novo responses later in infection (34).

An alternative method has recently been described that also presents HIV sequence diversity in improved peptide test sets, the potential T cell epitope (PTE) approach. The PTE approach is based on a biometric algorithm that includes frequently found k-mer sequences from circulating strains in a sequence database to build a standardized panel of HIV peptides for CTL based vaccine evaluation (35). Similarly to the toggled peptides, PTE peptides detect significantly more and stronger responses than consensus based peptides (36). Toggled peptides and PTE peptides each have different virtues and issues. Each peptide has to be separately synthesized for the PTE approach, so it is more expensive, and the coverage attained with PTE peptides is somewhat reduced relative to toggled peptides (B. Korber and K. Yusim, manuscript in preparation). Toggled peptides provide a clear linear map along a protein, facilitating ease of interpretation and direct comparison to traditional single sequence overlapping peptide methods, but they are mixtures of peptides that may be less reproducible in synthesis, and they will provide less precise information regarding specific peptide reactivity than the PTE approach. Based on similar reasoning regarding the advantages of covering T cell epitope diversity, two recent vaccine Ag designs have been proposed that incorporate variants in an attempt to elicit more broadly cross-reactive T cell responses through vaccination. These are mosaic proteins, which are sets of in silico recombinant full-length HIV proteins that in combination provide

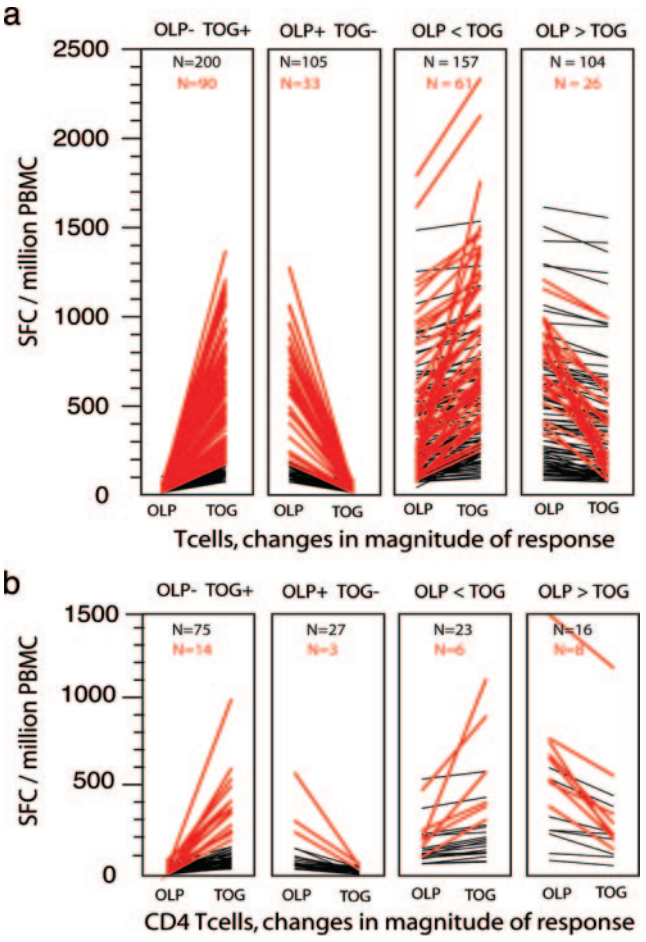



**FIGURE 5.** Toggled peptides elicit stronger responses than consensus overlapping peptides. The magnitude of gains and losses in responses measured by SFC/million between matched pairs of consensus OLP and toggled peptides are shown for total T cell responses (*a*) and $CD4^+$ T cell responses (*b*). Relative gains in toggled peptides are separated from relative gains in OLP consensus peptides, and increases that bring a response from positive to negative are considered separately from changes in the magnitude of two positive responses. Ties are not included in this figure, only responses with different magnitudes; nor are responses to peptides that were so conserved that no toggled version of the peptide needed to be synthesized. Black numbers indicate the total number of changes in a category, red numbers indicate the number of changes in the subset of responses that differed by >160 SFC/million. Changes greater than 160 SFC/million (shown in red lines) were considered significant as this cut-off reflects changes in magnitudes that were greater than 95% of experimental error in duplicate wells.

optimal coverage of potential T cell epitopes and exclude rare or unique epitopes (9), and COT+ (11), which complements a center of tree (COT) full protein sequence with a set of protein fragments designed to optimize potential epitope coverage.

Despite the overall net advantage in using toggled peptides to detect T cell reactivity, some positive responses to the standard test peptide sets were lost or diminished when using the toggled peptides. Although the consensus sequence was always present in the toggled peptide, some responses, especially weak ones, may potentially be lost due to the dilution of the reactive peptide sequence in the toggle synthesis mix, or due to possible antagonistic or stochastic effects caused by the mixture of multiple sequence variants in the toggle preparation. Specific cases where a consensus OLP elicited response is not detected using toggled peptides could be identified and resolved by testing single variant sequences and all possible combinations thereof. This, in turn, could provide potentially crucial information for polyvalent vaccine Ag design, identifying nonimmunogenic or antagonistic sequence variants with good binding affinity that would preferentially be excluded in vaccine immunogens. In addition, testing toggled peptides in assay systems that detect multiple effector functions may also reveal specific sequence variants that could act as superagonists or mediate wide cross-reactivity, and which could thus represent premier candidates for vaccine immunogen design.

Finally, because relatively conserved regions of the viral genome were chosen for this study, the observed benefits of using toggled peptides when compared with consensus sequence based test sets reflect conservative increases, which may be even more significant when applying this approach to more variable proteins in HV and other variable pathogens (33). However, while designing toggled peptides for highly variable targets is feasible, it will, as illustrated in Fig. 1*e*, likely require balancing the desired threshold for sequence diversity coverage with the complexity of the toggled peptide mixtures. Indeed, while there was no overall association between toggle complexity and the magnitude of responses ($p = 0.1$, Spearman one-sided test), focusing on responses to toggled peptides containing >20 variants revealed an inverse correlation between the number of variants and the detected magnitude of the response ($p = 0.009$), suggesting that the inclusion of additional variants above a certain threshold may reduce the strength of the response in vitro. In contrast, these more diverse toggled peptides correspond to more variable regions of HIV genome, where the consensus peptides have a lower response detection rate (12). Thus, more population-representative, complex toggled peptides may detect responses where their consensus counterpart failed to do so, suggesting that a lower magnitude may be a reasonable price to pay for a more complete view of the HIV-specific immune response. One approach to minimize toggle complexity would allow for an experimentally determined maximum number of variants in each toggle preparation to limit dilution effects. In addition, rational limitations to specific polymorphisms, for instance amino acids that revisit the same sequence space in different clades within the M group, could help developing toggled peptides suitable as global reagents while not being affected by extensive dilution effects. Such limitations to diversity (3) make biological sense given the context of what we know regarding the predictability of drug resistance mutations and some immune escape mutations (37, 38). Of note, M group toggles perfectly match some sequences from multiple clades and recombinant forms (data not shown), emphasizing the global usefulness of the toggle approach especially considering the increasing frequency with which recombinant viruses are being described.

Taken together, toggled peptides are a cost effective means to enable better probing and descriptions of the immune response across populations, and have clear applications for other variable pathogens such as hepatitis C virus, for which truly comprehensive assessments of an individual's immune response are limited by sequence diversity in the infecting viral isolate.

## Disclosures

The authors have no financial conflict of interest.

## References

1. Brander, C., N. Frahm, and B. D. Walker. 2006. The challenges of host and viral diversity in HIV vaccine design. *Curr. Opin. Immunol.* 18: 430–437.
2. Frahm, N., and C. Brander. 2007. HIV viral diversity and escape from cellular immunity. *Curr. Infect. Dis. Rep.* 9: 161–166.

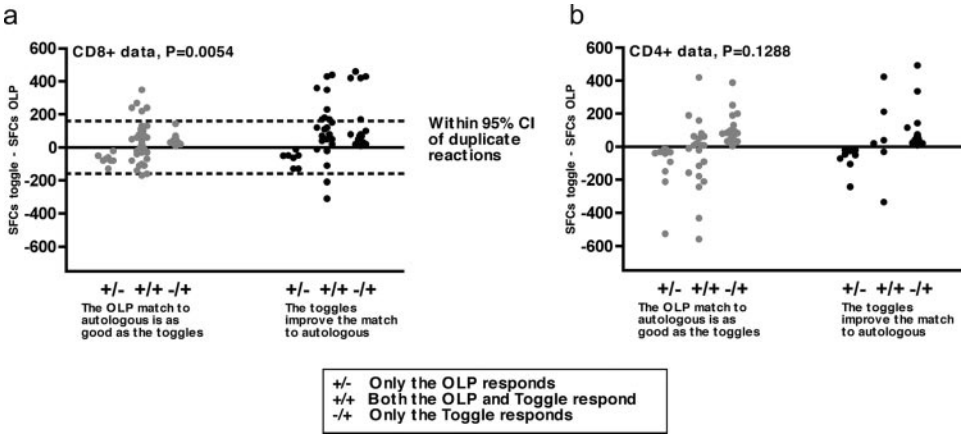



**FIGURE 6.** Correlation between the magnitude of response and the match to the autologous sequence. The magnitude of gains and losses in responses measured by spot forming cells between matched pairs of consensus OLP and toggled peptides are shown. As in Fig. 5, ties are not included in this figure, only responses with different numbers of SFC/million, nor are responses to peptides that were so conserved that no toggled version of the peptide was generated. Two groups of responses were considered: when the OLP sequence matched the autologous peptide as well as the toggles, and when the toggles improved the match to autologous peptide. The data in the two groups is compared by a Wilcoxon test. Within each category, the following situations were separated: $+/-$ only response to the OLP is detected, $-/+$ only response to the toggle is detected, $+/+$ response to both is detected. *a*, PBMC data, autologous sequences were available for 6 patients. The dashed lines show 95% confidence interval (160 SFC/million, see Fig. 5) of duplicate reactions. When the toggles improved the match to consensus, the responses were significantly higher than when the OLP match to autologous was as good as the toggles ($p = 0.0054$). *b*, $CD4^+$ T cell data, autologous sequences were available for 14 patients, of which peptides matching the autologous viral sequence in p24 and Nef were available for 13 and 7 subjects, respectively.


3. Martinez-Picado, J., J. G. Prado, E. E. Fry, K. Pfafferott, A. Leslie, S. Chetty, C. Thobakgale, I. Honeyborne, H. Crawford, P. Matthews, et al. 2006. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. *J. Virol.* 80: 3617–3623.
4. Li, B., A. D. Gladden, M. Altfeld, J. M. Kaldor, D. A. Cooper, A. D. Kelleher, and T. M. Allen. 2006. Rapid reversion of sequence polymorphisms dominates early HIV-1 evolution. *J. Virol.* 81: 193–201.
5. Friedrich, T. C., E. J. Dodds, L. J. Yant, L. Vojnov, R. Rudersdorf, C. Cullen, D. T. Evans, R. C. Desrosiers, B. R. Mothe, J. Sidney, et al. 2004. Reversion of CTL escape-variant immunodeficiency viruses in vivo. *Nat. Med.* 10: 275–281.
6. Allen, T. M., M. Altfeld, X. G. Yu, K. M. O'Sullivan, M. Lichterfeld, S. Le Gall, M. John, B. R. Mothe, P. K. Lee, E. T. Kalife, et al. 2004. Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in human immunodeficiency virus type 1 infection. *J. Virol.* 78: 7069–7078.
7. Herbeck, J. T., D. C. Nickle, G. H. Learn, G. S. Gottlieb, M. E. Curlin, L. Heath, and J. I. Mullins. 2006. Human immunodeficiency virus type 1 env evolves toward ancestral states upon transmission to a new host. *J. Virol.* 80: 1637–1644.
8. Liu, Y., J. McNevin, J. Cao, H. Zhao, I. Genowati, K. Wong, S. McLaughlin, M. D. McSweyn, K. Diem, C. E. Stevens, et al. 2006. Selection on the human immunodeficiency virus type 1 proteome following primary infection. *J. Virol.* 80: 9519–9529.
9. Fischer, W., S. Perkins, J. Theiler, T. Bhattacharya, K. Yusim, R. Funkhouser, C. Kuiken, B. Haynes, N. L. Letvin, B. D. Walker, et al. 2007. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nat. Med.* 13: 100–106.
10. Li, F., H. Horton, P. B. Gilbert, J. M. McElrath, L. Corey, and S. G. Self. 2007. HIV-1 CTL-based vaccine immunogen selection: antigen diversity and cellular response features. *Curr. HIV Res.* 5: 97–107.
11. Nickle, D. C., M. Rolland, M. A. Jensen, S. L. Pond, W. Deng, M. Seligman, D. Heckerman, J. I. Mullins, and N. Jojic. 2007. Coping with viral diversity in HIV vaccine design. *PLoS Comput. Biol.* 3: e75.
12. Frahm, N., B. T. Korber, C. M. Adams, J. J. Szinger, R. Draenert, M. M. Addo, M. E. Feeney, K. Yusim, K. Sango, N. V. Brown, et al. 2004. Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. *J. Virol.* 78: 2187–2200.
13. Addo, M. M., X. G. Yu, A. Rathod, D. Cohen, R. L. Eldridge, D. Strick, M. N. Johnston, C. Corcoran, A. G. Wurcel, C. A. Fitzpatrick, et al. 2003. Comprehensive epitope analysis of 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.* 77: 2081–2092.
14. Betts, M. R., D. R. Ambrozak, D. C. Douek, S. Bonhoeffer, J. M. Brenchley, J. P. Casazza, R. A. Koup, and L. J. Picker. 2001. Analysis of total human immunodeficiency virus (HIV)-specific $CD4^+$ and $CD8^+$ T-cell responses: relationship to viral load in untreated HIV infection. *J. Virol.* 75: 11983–11991.
15. Novitsky, V., H. Cao, N. Rybak, P. Gilbert, M. F. McLane, S. Gaolekwe, T. Peter, I. Thior, T. Ndung'u, R. Marlink, et al. 2002. Magnitude and frequency of cytotoxic T-lymphocyte responses: identification of immunodominant regions of human immunodeficiency virus type 1 subtype C. *J. Virol.* 76: 10155–10168.
16. Crawley, M. J. 2002. *Statistical Computing: An Introduction to Data Analysis using S-Plus*. J. Wiley and Sons Limited, Chichester, West Sussex, England.
17. Agresti, A. 2002. *Categorical Data Analysis*. Wiley, Hoboken, NJ.
18. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*. Chapman and Hall, London, U.K.
19. R. Development Core Team. 2006. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
20. Crawley, M. J. 2005. *Statistics: An Introduction Using R*. Wiley, Hoboken, NJ.
21. Allen, T. M., M. Altfeld, S. C. Geer, E. T. Kalife, C. Moore, M. O'Sullivan, K. I. Desouza, M. E. Feeney, R. L. Eldridge, E. L. Maier, et al. 2005. 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.* 79: 13239–13249.
22. Gaschen, B., J. Taylor, K. Yusim, B. Foley, F. Gao, D. Lang, V. Novitsky, B. Haynes, B. H. Hahn, T. Bhattacharya, and B. Korber. 2002. Diversity considerations in HIV-1 vaccine selection. *Science* 296: 2354–2360.
23. Kijak, G. H., J. R. Currier, S. Tovanabutra, J. H. Cox, N. L. Michael, S. A. Wegner, D. L. Birx, and F. E. McCutchan. 2004. Lost in translation: implications of HIV-1 codon usage for immune escape and drug resistance. *AIDS Rev.* 6: 54–60.
24. Yusim, K., C. Kesmir, B. Gaschen, M. M. Addo, M. Altfeld, S. Brunak, A. Chigaev, V. Detours, and B. T. Korber. 2002. Clustering patterns of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1 (HIV-1) proteins reveal imprints of immune evasion on HIV-1 global variation. *J. Virol.* 76: 8757–8768.
25. Goulder, P. J., C. Brander, Y. Tang, C. Tremblay, R. A. Colbert, M. M. Addo, E. S. Rosenberg, T. Nguyen, R. Allen, A. Trocha, et al. 2001. Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* 412: 334–338.
26. Gao, F., E. A. Weaver, Z. Lu, Y. Li, H. X. Liao, B. Ma, S. M. Alam, R. M. Scearce, L. L. Sutherland, J. S. Yu, et al. 2005. Antigenicity and immunogenicity of a synthetic human immunodeficiency virus type 1 group m consensus envelope glycoprotein. *J. Virol.* 79: 1154–1163.
27. Kothe, D. L., J. M. Decker, Y. Li, Z. Weng, F. Bibollet-Ruche, K. P. Zammit, M. G. Salazar, Y. Chen, J. F. Salazar-Gonzalez, Z. Moldoveanu, et al. 2006. Antigenicity and immunogenicity of HIV-1 consensus subtype B envelope glycoproteins. *Virology* 360: 218–234.
28. Kothe, D. L., Y. Li, J. M. Decker, F. Bibollet-Ruche, K. P. Zammit, M. G. Salazar, Y. Chen, Z. Weng, E. A. Weaver, F. Gao, et al. 2006. Ancestral and consensus envelope immunogens for HIV-1 subtype C. *Virology* 352: 438–449.
29. Kaufmann, D. E., P. M. Bailey, J. Sidney, B. Wagner, P. J. Norris, M. N. Johnston, L. A. Cosimi, M. M. Addo, M. Lichterfeld, M. Altfeld, et al. 2004. Comprehensive analysis of human immunodeficiency virus type 1-specific CD4 responses reveals marked immunodominance of gag and nef and the presence of broadly recognized peptides. *J. Virol.* 78: 4463–4477.
30. Feeney, M. E., K. A. Roosevelt, Y. Tang, K. J. Pfafferott, K. McIntosh, S. K. Burchett, C. Mao, B. D. Walker, and P. J. Goulder. 2003. Comprehensive screening reveals strong and broadly directed human immunodeficiency virus

type 1-specific CD8 responses in perinatally infected children. *J. Virol.* 77: 7492–7501.

31. Cao, J., J. McNevin, S. Holte, L. Fink, L. Corey, and M. J. McElrath. 2003. Comprehensive analysis of human immunodeficiency virus type 1 (HIV-1)-specific γ interferon-secreting $CD8^+$ T cells in primary HIV-1 infection. *J. Virol.* 77: 6867–6878.
32. Betts, M. R., M. C. Nason, S. M. West, S. C. De Rosa, S. A. Migueles, J. Abraham, M. M. Lederman, J. M. Benito, P. A. Goepfert, M. Connors, et al. 2006. HIV nonprogressors preferentially maintain highly functional HIV-specific $CD8^+$ T-cells. *Blood* 107: 4781–4789.
33. Altfeld, M., M. M. Addo, R. Shankarappa, P. K. Lee, T. M. Allen, X. G. Yu, A. Rathod, J. Harlow, K. O'Sullivan, M. N. Johnston, et al. 2003. 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.* 77: 7330–7340.
34. Allen, T. M., X. G. Yu, E. T. Kalife, L. L. Reyor, M. Lichterfeld, M. John, M. Cheng, R. L. Allgaier, S. Mui, N. Frahm, et al. 2005. De novo generation of escape variant-specific $CD8^+$ T-cell responses following cytotoxic T-lymphocyte escape in chronic human immunodeficiency virus type 1 infection. *J. Virol.* 79: 12952–12960.
35. Li, F., U. Malhotra, P. B. Gilbert, N. R. Hawkins, A. C. Duerr, J. M. McElrath, L. Corey, and S. G. Self. 2006. Peptide selection for human immunodeficiency virus type 1 CTL-based vaccine evaluation. *Vaccine* 24: 6893–6904.
36. Malhotra, U., F. Li, J. Nolin, M. Allison, H. Zhao, J. I. Mullins, S. Self, and M. J. McElrath. 2007. Enhanced detection of human immunodeficiency virus type 1 (HIV-1) nef-specific T cells recognizing multiple variants in early HIV-1 infection. *J. Virol.* 81: 5225–5237.
37. Goulder, P. J., and D. I. Watkins. 2004. HIV and SIV CTL escape: implications for vaccine design. *Nat. Rev. Immunol.* 4: 630–640.
38. Draenert, R., T. M. Allen, Y. Liu, T. Wrin, C. Chappey, C. L. Verrill, G. Sirera, R. L. Eldridge, M. P. Lahaie, L. Ruiz, et al. 2006. Constraints on HIV-1 evolution and immunodominance revealed in monozygotic adult twins infected with the same virus. *J. Exp. Med.* 203: 529–539.