Supporting Information

de Groot et al. 10.1073/pnas.1009136107

SI Materials and Methods

Patr-A- and -B-Transfected Cell Lines. Thirty-two wild-caught West African chimpanzees originating from Sierra Leone represent the founder animals of the colony that was the subject of the study. Epstein–Barr virus-transformed B-cell lines were used to obtain total RNA, and *Patr* class I amplification was performed as described (1). Target DNA and vector pcDNA I/Neo were digested with NsiI and HindIII (Invitrogen) and subsequently ligated. Competent MC1061/P3 *Escherichia coli* cells were used for transformation. The constructs containing the relevant Patr-A or -B molecules were individually transfected into K562 cells (which lack HLA expression) by electroporation (2). MHC class I-positive cells were enriched by cell sorting using Dynabeads (Dynal) coated with monoclonal antibody W6/32, and were subsequently tested for MHC class I expression. During the course of the study, the sequence of *Patr-B*0201* was corrected and appeared to be identical to *Patr-B*0501*. The IPD-MHC nonhuman primate database (http://www.ebi.ac.uk/ipd/mhc/nhp) records all chimpanzee class I sequences. Many sequences for the different subspecies have been described. All of them group, however, within the lineages that are described for the West African chimpanzee population that is the basis of this study.

Determination of Peptide-Binding Motifs. K562 Patr-A or -B-positive transfectants were cultured in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with G418 and 5% FCS up to 1×10^{10} cells. The MHC class I expression was routinely checked by FACS analysis and mycoplasma infection using the MycoSensor PCR assay kit (Stratagene). Cells were harvested and washed with PBS, and the cell pellet was stored at −80 °C. Subsequently, the cells were lysed with lysis buffer and centrifuged for 60 min at $10{,}000 \times g$ to remove nuclei and insoluble material. Sepharose beads covalently linked to monoclonal antibody W6/32 were used to preclear the lysate by gentle mixing for 60 min. The beads were then washed, and in a last step the MHC-peptide complexes were eluted with 10% acetic acid in water. All purification steps were performed at 4 °C. High molecular mass material (MHC molecules) was removed through Centriprep filtration (3). The peptide pool was prefractionated on a C18 RP-HPLC system (Dr. Maisch GmbH). Fractions were reduced to near dryness and subsequently analyzed by tandem mass spectrometry. Peptides were characterized by nanoflow liquid chromatography using an Agilent 1100 HPLC system (Agilent Technologies) coupled online to a 7-tesla LTQ-FT mass spectrometer (Thermo Electron). The chromatographic system consisted of the following components: ReproSil-Pur C18-AQ (3 μm was used as a resin for the analytical nanocolumn) and AQUA-C18 (5 μm was used as a resin for the trapping column). Columns were prepared in-house. The end of the nanocolumn was drawn to a tip (i.d. ≈5 μm), from which the eluent was sprayed into the mass spectrometer. Peptides were trapped at 5 μL/min on a 1-cm column (100-μm internal diameter, packed in-house) and eluted to a 15-cm column (50-μm internal diameter, packed in-house) at 150 nL/min in a 60-min gradient from 0 to 50% acetonitrile in 0.1% formic acid. The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS acquisition. Full-scan MS spectra were acquired in the FT-ICR with a resolution of 25,000 at a target value of 5,000,000. The two most intense ions were then isolated for accurate mass measurements by a selected ion-monitoring scan in the Fourier transform ion cyclotron resonance with a resolution of 50,000 at a target accumulation value of 50,000. The selected ions were then fragmented in the linear ion trap using collision-induced dissociation at a target value of 10,000. The raw data were converted to peak lists using Bioworks Browser software version 3.1. For protein identification, MS/MS data were submitted to the human IPI database using Mascot version 2.1 (Matrix Science). All reported hits were assessed manually, and peptides with Mascot scores lower than 35 were discarded.

Selection and Synthesis of Indicator and Competitor Peptides. For the various Patr molecules, indicator peptides were selected from the list of eluted natural peptides, and were characterized by high Mascot values. The indicator peptides were labeled at the sixth or seventh position, with a long-chain biotin linked to a cysteine residue. The binding affinity to the relevant Patr molecule was verified by a competition assay with an unlabeled similar competitor peptide, the standardized competitor (SC) (Fig. S2). For each Patr molecule, the data of the indicator peptide that showed the strongest binding capacity overall are presented. For most tested combinations, the observed difference between labeling on the sixth or seventh position was not significant (Fig. S2). Competitor peptides for the different Patr molecules were selected using the SYFPEITHI database (http://www.syfpeithi.de).

For HLA-A*02, hepatitis B virus (HBV) cAG 18–27 was taken as indicator peptide (4), and labeled with a long-chain biotin linked to a cysteine residue at the sixth position. Unlabeled HBV cAG 18–27 was taken as a standardized competitor. HIV-1 p24 15–23 was taken as competitor peptide predicted to have no/low binding affinity.

All peptides were synthesized by solid-phase strategies on an automated multiple-peptide synthesizer (Syro II; MultiSynTech) using Fmoc chemistry, and were analyzed by reverse-phase HPLC (5). The expected molecular masses of the peptides were confirmed using MALDI-TOF mass spectrometry.

Cell-Based Peptide-Binding Competition Assays and Cytotoxic T-Cell Assays. A plate-based cell-based peptide-binding competition assay (CPBCA) was developed to measure the binding affinity of the different peptides using time-resolved fluorescence as readout. Before the assay, the cells were maintained at 26 °C for 2–3 d, as this results in empty MHC class I molecules on the cell surface (6). For the test, a 96-well V-bottom plate was used, and each test well contained 25 μL of indicator peptide (500 nM), 25 μL of competitor peptide (100–0.2 μM, 25–0.05 μM, or 6.25–0.01 μM serial dilution, using nine 2-fold dilution steps; For peptide EEALQAFTY, the following concentration series was taken: 100 μM, 12.5 μM, 5 μM, 1 μM, 0.5 μM, 0.2 μM, 0.1 μM, 0.04 μM, 0.008 μM, 0.001 μM), 25 μL of IMDM supplemented with 2% FCS, and 25 μL of cells (100,000 cells/well). Wells containing indicator peptide only or medium controls were incorporated to determine the maximal and background measures, respectively. The plates were incubated for 24 h at 4 °C. On the following day, unbound peptides were removed by washing the cell pellets two times with 100 μL ice-cold PBS supplemented with 2% FCS. Subsequently, the cell pellets were resuspended and incubated for 30 min at 4 °C in 100 μL europium-labeled streptavidin diluted in DELFIA assay buffer (1/1000) (Perkin-Elmer). Unbound europium-labeled streptavidin was removed by washing five times. Finally, the cell pellets were dissolved in 150 μL of DELFIA Enhancement solution and transferred to a MaxiSorp F96-well plate (Nunc-Immuno). The europium signal was measured using a Victor^3 1420 multilabel counter (PerkinElmer).

Whether the CPBCA can distinguish between binding and no/low-binding peptides was verified using the HLA-A*02-positive

B-cell lines JY and DBB, and a K562 HLA-A2-transfected single-molecule-expressing cell line (SAL). The binding affinity of a peptide was quantified as the concentration giving 50% inhibition of binding of the indicator peptide (IC_{50}). The results show that the CPBCA can differentiate between an MHC/peptide combination predicted to have a high binding affinity versus a combination predicted to have a no/low-binding affinity (Figs. S2 and S5). Previously described cytotoxic T-cell (CTL) assays were reevaluated (7, 8).

Calculation of Percentage Inhibition, IC_{50} Values, and Statistical Analysis. The relative inhibition of the indicator peptide was calculated as follows:

$$\text{Inhibition } (\%) = \left[1 - \left(\text{MF}_{\text{indicator + competitor}} - \text{MF}_{\text{background}}\right) / \left(\text{MF}_{\text{indicator}} - \text{MF}_{\text{background}}\right)\right] \times 100\%.$$

Background values were subtracted and values were normalized to 100% maximum activity. All available replicates [three in the case of the performed CPBCA for JY/DBB/HLA-A2(SAL), and five in the case of the performed CPBCA for the four Patr molecules] were entered in a regression analysis. IC_{50} values were estimated using nonlinear least-squares regression with the "R" platform for statistical computing (http://www.r-project.org), using the formula: inhibition = $100/[1 + e^{(\log IC50 - \log \text{competitor conc.} \times \text{slope})}]$, where inhibition is the relative reduction of peptide binding expressed as a fraction, IC_{50} is the estimate of the peptide concentration yielding 50% inhibition, and slope is a parameter indicating the steepness of the curve. IC_{50} and the corresponding 95% confidence intervals (CI) were calculated from the regression results using (*i*) the parameter estimates, (*ii*) corresponding SEs, and (*iii*) degrees of freedom. The IC_{50} values are presented as the antilog values obtained from the regression analyses. Data are considered statistically significantly different when the 95% CIs do not overlap.

1. de Groot NG, et al. (2000) Major histocompatibility complex class I diversity in a West African chimpanzee population: Implications for HIV research. *Immunogenetics* 51: 398–409.
2. Zoet YM, et al. (2005) The single antigen expressing lines (SALs) concept: An excellent tool for screening for HLA-specific antibodies. *Hum Immunol* 66:519–525.
3. Stepniak D, et al. (2008) Large-scale characterization of natural ligands explains the unique gluten-binding properties of HLA-DQ2. *J Immunol* 180:3268–3278.
4. Bertoletti A, et al. (1993) Definition of a minimal optimal cytotoxic T-cell epitope within the hepatitis B virus nucleocapsid protein. *J Virol* 67:2376–2380.
5. Hiemstra HS, et al. (1997) The identification of CD4+ T cell epitopes with dedicated synthetic peptide libraries. *Proc Natl Acad Sci USA* 94:10313–10318.
6. Schumacher TN, et al. (1990) Direct binding of peptide to empty MHC class I molecules on intact cells and in vitro. *Cell* 62:563–567.
7. Balla-Jhagjhoorsingh SS, et al. (1999) Conserved CTL epitopes shared between HIV-infected human long-term survivors and chimpanzees. *J Immunol* 162:2308–2314.
8. Balla-Jhagjhoorsingh SS, et al. (2001) Protection from secondary human immunodeficiency virus type 1 infection in chimpanzees suggests the importance of antigenic boosting and a possible role for cytotoxic T cells. *J Infect Dis* 184:136–143.
9. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14:1188–1190.
10. Macdonald WA, et al. (2009) T cell allorecognition via molecular mimicry. *Immunity* 31:897–908.

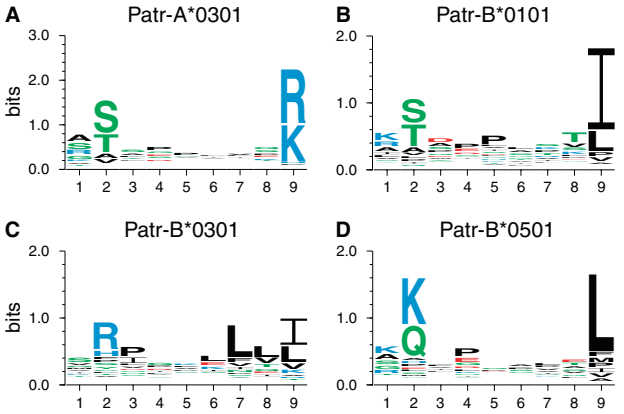

Fig. S1. Peptide-binding motifs determined from the eluted natural peptides. (*A*) Patr-A*0301, (*B*) -B*0101, (*C*) -B*0301, and (*D*) -B*0501. WebLogo version 3.0 was used for visualization of the eluted 9-mer peptides (9). In a sequence logo, the height of a column of letters is equal to the information content at that position. The height of a letter in a column is proportional to the frequency of the corresponding amino acid at that position. The color code of the amino acids is according to chemical property: green, polar; blue, basic; red, acidic; and black, hydrophobic amino acids. As an example, a list of eluted 8-, 9-, and 10-mer natural peptides used to elucidate the peptide-binding motif for Patr-A*0301 is provided (Table S2).

A

Cell line	Indicator peptide	Competitor peptide	IC_{50} (μM)	95% CI
JY	FLPSDCFPSV	FLPSDCFPSV (HBV cAG 18-27) ISPRTLNAW (HIV-1 p24 15-23)	0.3 >100	(0.2-0.4)
DBB	FLPSDCFPSV	FLPSDCFPSV ISPRTLNAW	0.3 >100	(0.2-0.4)
HLA-A2 (SAL)	FLPSDCFPSV	FLPSDCFPSV ISPRTLNAW	0.2 46.5	(0.2-0.2) (17.0-127.4)
MHC molecule				
Patr-A*0301	ATALECVYK/ ATALEYCYK	ATALEYVYK[1] MVHQAISPR (HIV-1 p24 10-18) RTLNAWVKV (HIV-1 p24 18-26) HQAAMQMLK (HIV-1 p24 62-70) RMYSPTSIL (HIV-1 p24 143-151) NPPIPVGEI (HIV-1 p24 121-129)	1.5/0.8 1.6/0.6 0.07/0.1 2.3/1.1 0.2/0.1 5.5/3.2	(1.0-2.2)/(0.5-1.2) (1.2-2.1)/(0.4-0.8) (0.05-0.1)/(0.08-0.2) (1.7-3.2)/(0.8-1.6) (0.2-0.3)/(0.06-0.2) (4.0-7.4)/(1.9-5.4)
Patr-B*0101	LSDMHCRSI/ LSDMHLCSI	LSDMHLRSI[2] ISPRTLNAW RTLNAWVKV TTSTLQEQV (HIV-1 p24 107-115) TTSTLQEQI (HIV-1 p24 107-115) RMYSPTSIL NPPIPVGEI	0.5/1.0 0.6/0.9 0.4/0.1 8.2/5.0 6.9/9.6 0.4/1.4 8.2/27.1	(0.4-0.8)/(0.8-1.2) (0.4-0.8)/(0.7-1.3) (0.3-0.5)/(0.08-0.2) (5.3-12.6)/(3.4-7.3) (4.9-9.6)/(7.3-12.6) (0.3-0.5)/(1.0-2.0) (4.3-15.6)/(14.8-49.8)
Patr-B*0301	GRIDICQLI/ GRIDIKCLI	GRIDIKQLI[3] KRWIILGLN (HIV-1 p24 131-139) RTLNAWVKV RMYSPTSIL NPPIPVGEI	0.1/0.2 0.2/0.5 0.2/0.2 0.7/0.3 51.1/28.7	(0.08-0.2)/(0.2-0.3) (0.1-0.3)/(0.3-0.6) (0.1-0.3)/(0.1-0.3) (0.4-1.1)/(0.2-0.6) (23.0-113.4)/(10.3-79.8)
Patr-B*0501	GQYEQCKQL/ GQYEQVCQL	GQYEQVKQL[4] HQAISPRTL (HIV-1 p24 12-20) RTLNAWVKV EKAFSPEVI (HIV-1 p24 29-37) VKNWMTETL (HIV-1 p24 181-189) YKRWIILGL (HIV-1 p24 130-138) RMYSPTSIL NPPIPVGEI	1.5/1.3 0.5/0.5 0.08/0.2 2.7/4.1 0.1/0.3 0.2/0.8 0.1/0.2 4.0/9.3	(1.0-2.3)/(1.0-1.8) (0.3-0.7)/(0.4-0.7) (0.04-0.1)/(0.1-0.2) (2.0-3.7)/(3.2-5.3) (0.1-0.2)/(0.3-0.4) (0.09-0.3)/(0.4-1.2) (0.07-0.2)/(0.1-0.3) (2.7-6.1)/(5.9-14.7)

B

MHC molecule	Indicator peptide	Competitor peptide	IC_{50} (μM)	95% CI
Patr-A*0301	ATALECVYK	HQAISPRTL (HIV-1 p24 12-20) ISPRTLNAW (HIV-1 p24 15-23) EKAFSPEVI (HIV-1 p24 29-37) TTSTLQEQV (HIV-1 p24 107-115) TTSTLQEQI (HIV-1 p24 107-115) VKNWMTETL (HIV-1 p24 181-189) YKRWIILGL (HIV-1 p24 130-138) KRWIILGLN (HIV-1 p24 131-139) EEALQAFTY[5]	3.9 0.7 18.1 9.1 6.3 0.1 0.5 0.2 24.8	2.6-5.9 0.4-1.1 11.1-29.4 6.6-12.5 4.1-9.6 0.09-0.2 0.3-0.8 0.2-0.3 5.2-117.4
Patr-B*0101	LSDMHLCSI	MVHQAISPR (HIV-1 p24 10-18) HQAISPRTL EKAFSPEVI VKNWMTETL YKRWIILGL KRWIILGLN HQAAMQMLK (HIV-1 p24 62-70) EEALQAFTY	3.6 1.8 21.1 0.4 1.3 0.6 5.7 22.9	2.3-5.8 0.8-4.4 13.4-33.1 0.2-0.8 0.7-2.5 0.2-1.5 2.4-13.4 11.5-45.7
Patr-B*0301	GRIDIKCLI	MVHQAISPR HQAISPRTL ISPRTLNAW EKAFSPEVI TTSTLQEQV TTSTLQEQI VKNWMTETL YKRWIILGL HQAAMQMLK EEALQAFTY	12.8 7.9 2.7 >100 62.8 24.9 5.8 0.3 >100 >100	7.4-22.4 3.1-20.1 1.6-4.4 29.2-135.1 8.9-69.6 3.2-10.4 0.2-0.5
Patr-B*0501	GQYEQVCQL	MVHQAISPR ISPRTLNAW TTSTLQEQV TTSTLQEQI KRWIILGLN HQAAMQMLK EEALQAFTY	0.3 0.09 2.9 5.2 0.03 1.1 21.2	0.1-0.7 0.06-0.1 1.7-4.7 3.6-7.5 0.02-0.05 0.4-2.7 8.4-53.5

Fig. S2. Overview of the indicator and competitor peptides tested for JY, DBB, HLA-A2(SAL), and the different Patr molecules. The determined IC_{50} values and the corresponding 95% confidence intervals (CI) are provided. (*A*) The IC_{50} values for JY/DBB/HLA-A2(SAL) and the Patr molecules are the result of three and five individual experiments, respectively. For the Patr molecules, the data of both tested indicator peptides are shown. (*B*) The IC_{50} values presented are the result of three individual experiments. C indicates the position where the indicator peptides are labeled. SAL, single-molecule-expressing cell line. [1]ATALEYVYK is a biogenesis of lysosome-related organelles complex-1 subunit 1. [2]LSDMHLRSI is an isoform 1 of fragile X mental retardation syndrome-related protein 1. [3]GRIDIKQLI is a sterol O-acyltransferase 1. [4]GQYEQVKQL is a CCR4-NOT transcription complex, subunit 1 isoform b. [5]EEALQAFTY is an alanine-substituted peptide of the ATP-binding cassette subfamily D member 3-derived peptide EEYLQAFTY (10). The color codes correspond to those in Fig. 2.

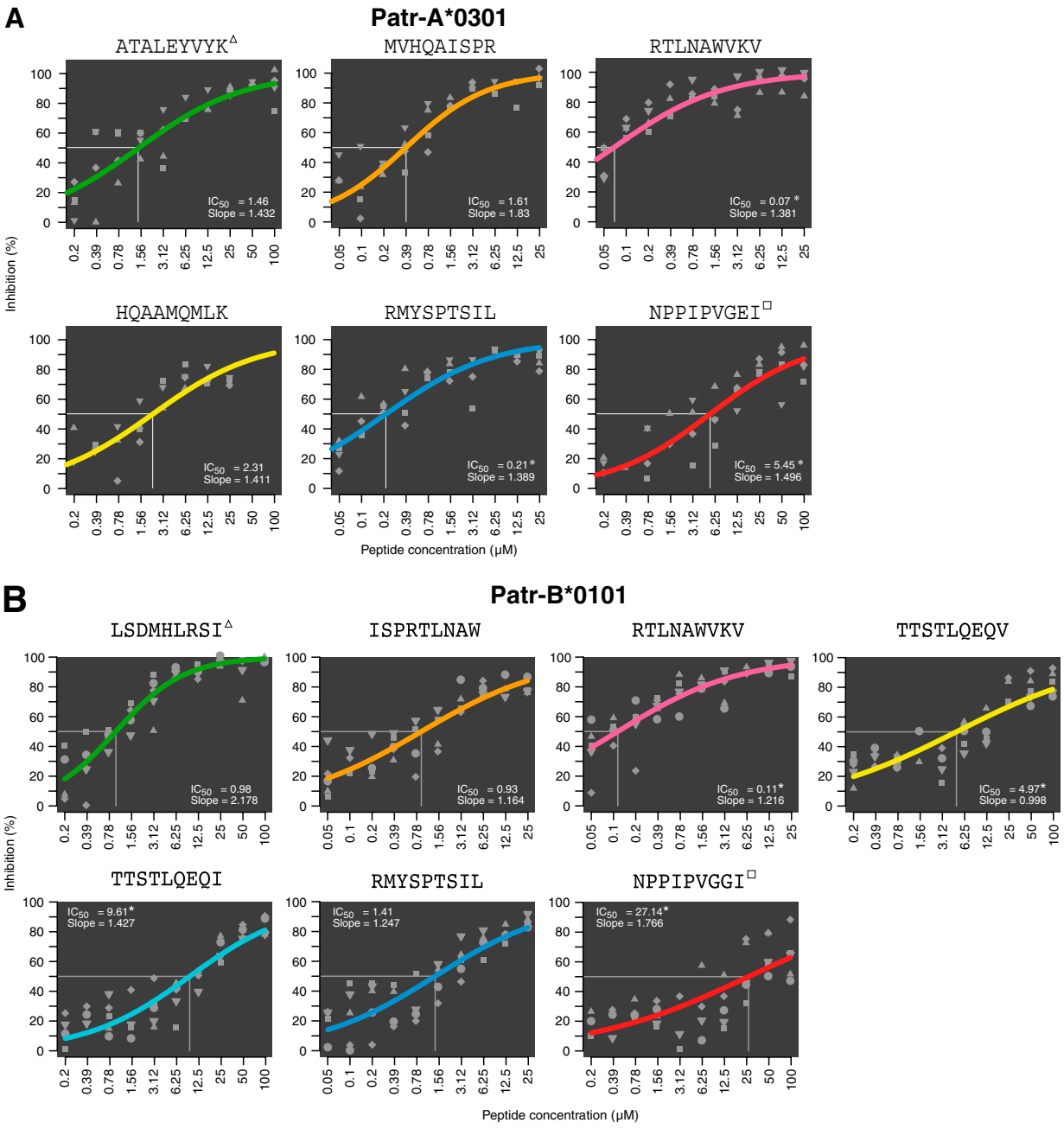

Fig. S3. (Continued)

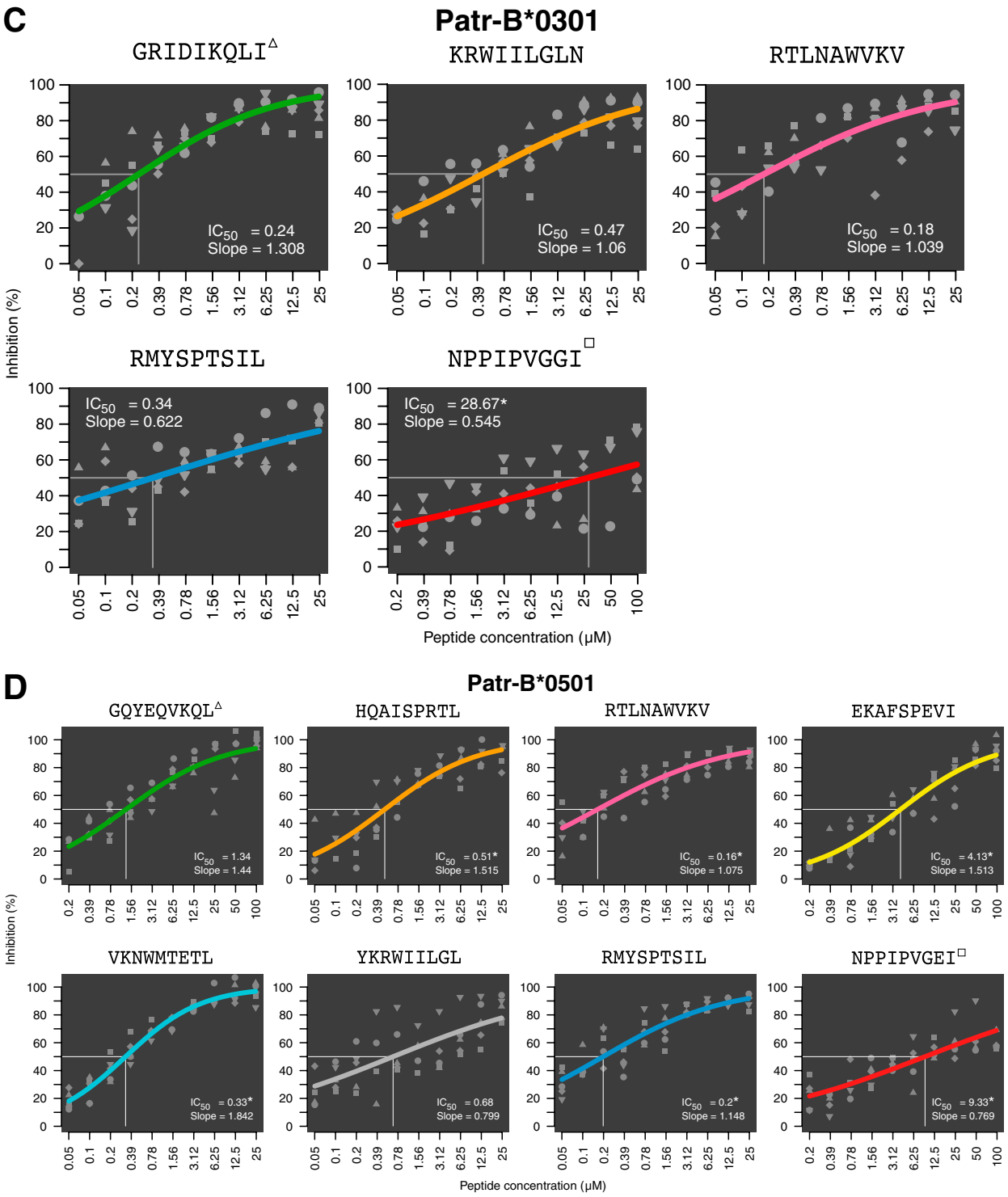

Fig. S3. Individual dose-inhibition curves for the competitor peptides predicted to bind tested for (*A*) Patr-A*0301, (*B*) -B*0101, (*C*) -B*0301, and (*D*) -B*0501. The indicated IC_{50} values (μM) are derived from the regression curves of five individual experiments. As indicator peptides, ATALECVYK biotin-labeled at the p6 position (indicated with a C) and LSDMHLCSI, GRIDIKCLI, and GQYEQVCQL biotin-labeled at the p7 position were used, respectively. A triangle marks the standardized competitor (SC; peptide similar to the biotin-labeled indicator peptide). A square indicates the Gag-derived control peptide. An asterisk indicates a binding affinity significantly higher or lower than the SC.

```
Pep 1: PivQNAQGQmvHQslSPRTL
Pep 2: vHQslSPRTLNAwVKViEek
Pep 3: NAwVKViEekafsPEVIPmF
Pep 4: afsPEVIPmFSALSEGATPq
Pep 5: SALSEGATPqDLNmMLNiVG
Pep 6: DLNmMLNiVGGHQAAMqMLK
Pep 7: GHQAAMqMLKDTINEEAAew
Pep 8: DTINEEAAewDRlHPvhAGP
Pep 9: DRlHPvhAGPipPGQMREPR
Pep10: ipPGQMREPRGSDIAGTTST
Pep11: GSDIAGTTSTpQEQigwMTN
Pep12: pQEQigwMTNPPiPvGDIYk
Pep13: PPiPvGDIYkRwiILGLNKI
Pep14: RwiILGLNKIVRMYSPVSIL
Pep15: VRMYSPVSILDIkQGPKEPF
Pep16: DIkQGPKEPFRDYVDRFFKt
Pep17: RDYVDRFFKtlrAEQAtQVK
Pep18: lrAEQAtQVKWMTeTLLvQN
Pep19: WMTeTLLvQNANPDCKsILr
Pep20: ANPDCKsILraLGgATLEEM
Pep21: aLGgATLEEMMtACQGVGGP
Pep22: MtACQGVGGPgHKArVlAEA
```

Fig. S4. Individual peptides present in the Gag pools of the cellular immune-response assays. Twenty-two 20-mer peptides overlapping by 10 amino acids spanning the HIV-1_{SF2} (ARP-788; MRC) Gag amino acid residues 135–364 (8).

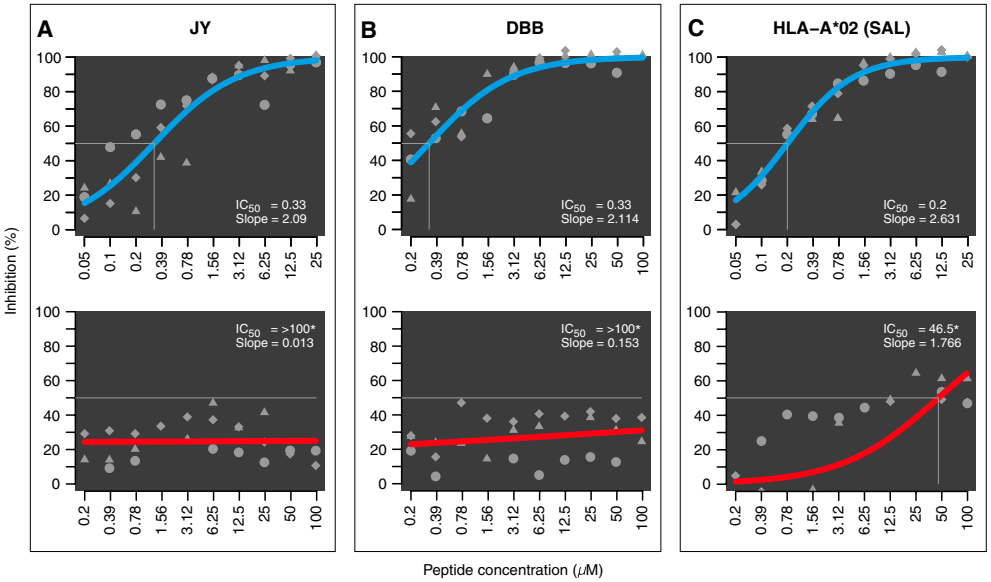

Fig. S5. Dose-inhibition curves for (*A*) JY, (*B*) DBB, and (*C*) HLA-A*02(SAL). As an indicator peptide, HBV cAG 18–27 biotin-labeled at the p6 position was used. The blue lines represent competition with unlabeled HBV cAG 18–27 (standardized competitor peptide). The red lines represent competition with HIV-1 p24 15–23, a peptide predicted to have no/low binding affinity. The indicated IC_{50} values (μM) are derived from the regression curves of three individual experiments. An asterisk indicates a binding affinity significantly lower than the standardized competitor peptide.

Other Supporting Information Files

Table S1 (DOC)
Table S2 (DOC)