A Long N-terminal-extended Nested Set of Abundant and Antigenic Major Histocompatibility Complex Class I Natural Ligands from HIV Envelope Protein*

Yolanda Samino[‡1], **Daniel López**[‡], **Sara Guil**[‡2], **Loredana Saveanu**[§], **Peter M. van Endert**[§], **and Margarita Del Val**[‡3]

From the ‡Centro Nacional de Microbiología, Instituto de Salud Carlos III, 28220 Madrid, Spain and §INSERM Unit 580, Université René Descartes Paris 5, 75015 Paris, France

Viral antigens complexed with major histocompatibility complex (MHC) class I molecules are recognized by cytotoxic T lymphocytes on infected cells. Assays with synthetic peptides identify optimal MHC class I ligands often used for vaccines. However, when natural peptides are analyzed, more complex mixtures including long peptides bulging in the middle of the binding site or with carboxyl extensions are found, reflecting lack of exposure to carboxypeptidases in the antigen processing pathway. In contrast, precursor peptides are exposed to extensive cytosolic aminopeptidase activity, and fewer than 1% survive, only to be further trimmed in the endoplasmic reticulum. We show here a striking example of a nested set of at least three highly antigenic and similarly abundant natural MHC class I ligands, 15, 10, and 9 amino acids in length, derived from a single human immunodeficiency virus gp160 epitope. Antigen processing, thus, gives rise to a rich pool of possible ligands from which MHC class I molecules can choose. The natural peptide set includes a 15-residue-long peptide with unprecedented 6 N-terminal residues that most likely extend out of the MHC class I binding groove. This 15-mer is the longest natural peptide known recognized by cytotoxic T lymphocytes and is surprisingly protected from aminopeptidase trimming in living cells.

Newly synthesized viral proteins are proteolytically processed before MHC[4] class I heavy chain-β_2m-peptide complex formation in the lumen of the endoplasmic reticulum (ER) (1). Peptides of 8–10 residues bind to MHC class I molecules usually by means of two major anchor residues at positions 2 and C-terminal in the antigenic peptide (2, 3). Some peptides may be directly produced by the proteasome in their final form, whereas others are generated as precursor peptides (4). These precursor peptides must display the correct C terminus of the final antigenic peptides, as the evidence suggests the absence of carboxypep-

tidases in the ER (1, 5). The limited number of studies on extraction and identification of natural peptides derived from a known foreign antigen shows that they constitute complex mixtures, mostly including C-terminal extensions of the minimal epitope (6–9). On the contrary, peptides are exposed to extensive cytosolic aminopeptidase activity, and fewer than 1% of them survive and are rescued by the transporters associated with antigen processing (TAP) and translocated into the ER (10, 11).

In most cases it is assumed that the natural MHC class I ligand is the one that has the canonical anchor sites, the minimal length, and the optimal antigenicity when tested as synthetic peptide. N-terminal extensions of such potential epitopes have been used to study trimming by aminopeptidases purified from the ER or microsomal fractions. There are several reports which show that the two enzymes, mouse endoplasmic reticulum aminopeptidase associated with antigen processing (ERAAP)/human ERAP1 and human leukocyte-derived arginine aminopeptidase (L-RAP)/ERAP2, can indeed trim the precursors to yield the putative epitope in solution in the absence of the peptide-binding MHC class I molecule and that they usually stop at around 8–9-aa-long products and in front of an *X*-proline bond (12–17). These properties suffice to consider them *bona fide* trimming enzymes for antigen presentation. There is also an intriguing report showing that MHC class I molecules can direct this trimming (18), implying that trimming might occur on the formed complex.

We have previously studied the natural peptides endogenously processed in living cells from the HIV gp160 glycoprotein and presented by the murine MHC class I molecule D[d] (19). Unexpectedly, the natural situation was more complex than having a single type of peptide-MHC class I complex, as we found evidence of a nested set of two or three equally abundant and equally antigenic peptides that differed in the N terminus. This suggested the existence of an intracellular pool of peptides derived from this natural protein in infected cells that are available for binding to MHC class I molecules. The shortest natural D[d] ligand derived from gp160 is a 9-mer with the canonical anchor motifs for L[d] (20), and there were also some hints of a marginal *in vivo* L[d]-restricted CTL response to gp160 (21). This prompted us to study the natural L[d] ligands derived from this full-length viral protein to gain insight into the nature of this pool of peptides and the mechanisms of antigen processing and presentation.

The analysis of the natural L[d] peptidic ligands resulting from the endogenous processing of the envelope glycoprotein shows the presence of at least three different MHC-peptide complexes in infected cells. Two peptides coincide with those described in D[d] (19), a 9-mer and a 10- or 11-mer. The third peptidic species corresponds to a 15-mer peptide with an N-terminal extension of 6 residues, which probably protrude out of the L[d] binding groove. The 15-mer peptide binds to the presenting molecule L[d], is recognized by 9-mer-selected CTL with an

antigenicity only slightly lower than the optimal 9-mer, and is more abundant in infected cells than this minimal epitope. The presence of N-terminal extensions in all these high affinity natural ligands in the face of an aminopeptidase activity at the site of complex formation suggests that trimming cannot proceed once the complexes are formed. This complex mixture of natural ligands for L^d parallels and expands the natural ligands of D^d from the same epitope, suggesting that antigen processing gives rise to a rich pool of possible ligands from which MHC class I molecules can choose. These results may have general implications for the rational design of vaccines.

EXPERIMENTAL PROCEDURES

Mice, Cell Lines, and Recombinant Vaccinia Viruses (rVV)—BALB/c mice (H-2^d haplotype) were bred in our animal facilities in accordance with national regulations. All cell lines were cultured in Iscove's modified Dulbecco's medium supplemented with 10% fetal bovine serum and 5×10^{-5} M β-mercaptoethanol. The P13.1 cell line is a derivative from mouse mastocytoma P815 cells (H-2^d) by transfection with the lacZ gene encoding β-galactosidase (22). For infection, untransfected murine kidney L cells (Ltk$^-$ cells) and L cells transfected with D^d (23) or L^d (24) were used. Transfectants with single and multiple mutations in the sequence of the L^d molecule were also used (25). They were named with the residue number followed by the single-letter code for the new aa introduced. For stability assays, the TAP-deficient human lymphoblastoid T2 cells transfected with L^d were employed (26). rVV-ENV (vSC25) encodes the envelope glycoprotein gp160 from the strain IIIB of HIV-1 under the control of the vaccinia early-late promoter 7.5k (27). Its DNA sequence was confirmed. The HIV glycoprotein can be detected by Western blot, but its intermediate to low level of expression precludes direct detection among the major proteins synthesized in the infected cell. The parental Western Reserve strain was used as control.

Synthetic Peptides—Peptides were synthesized in a peptide synthesizer (model 433A; Applied Biosystems, Foster City, CA) and purified when needed by cation exchange (IE) and/or reversed-phase (RP) HPLC. They were quantitated by A_{280} using peptide R10I as a standard, and the identity was confirmed by matrix-assisted laser desorption/ionization time of flight mass spectrometry. All peptide sequences are derived from the sequence ^{313}KIRIQR**GPGRAFVTI**GKIGN-MRQAH^{337}, which forms part of the V3-loop antigenic area from HIV-1 strain IIIB envelope glycoprotein. They are named indicating the first aa, the length, and the last aa. Thus, G9I refers to the nonamer of sequence ^{319}GPGRAFVTI^{327}. The single-letter aa code is used throughout. See Table 1 for an overview.

T Cell Lines and Cytotoxicity Assay—Polyclonal gp160-IIIB-monospecific CTL were generated by immunization of mice with rVV-ENV followed by weekly restimulation of splenocytes with G9I synthetic peptide and interleukin 2, as described (19, 28). They were used as effector cells in standard 6-h cytotoxicity assays. As targets, L/D^d and L/L^d cells were infected overnight with rVV as described (24). For peptide titrations, targets and peptides were incubated for 20 min, and CTLs were then added.

Isolation of Naturally Processed Peptides—L cells (3×10^9 in 30 roller bottles) were infected with rVV at a multiplicity of infection of 3.5 plaque-forming units/cell, and 16 h later naturally processed peptides were extracted from whole cells with trifluoroacetic acid, selected with a Macrosep centrifugal concentrator (Filtron) with a cut-off of 10 kDa, and purified by RP HPLC (19, 29). Control extractions were performed identically, except that pellets of 10^9 P13.1 cells received 10 nmol of synthetic peptides immediately after the addition of trifluoroacetic acid. Pooled RP HPLC fractions that tested positive in cytotoxicity assays

with P13.1 target cells and gp160-specific CTLs were newly chromatographed by IE HPLC chromatography in a Mono S column (19). Fractions from this second column and dilutions thereof were analyzed in new cytotoxicity assays. As the internal standard, a gp160-unrelated peptide was included in all HPLC runs. To ensure comparability of different runs, the actual conductivity gradient was always monitored. Contamination of the HPLC columns was excluded by testing preceding HPLC runs with CTL. Sometimes D^d-restricted CTLs were used in these assays (but not in any other experiment included in this report), because they showed an almost identical sensitivity to variations in peptide sequence as the L^d-restricted CTL (compare Table 1 and Ref. 19). For standardization, serial dilutions of synthetic peptides were tested always in parallel.

Mass Spectrometry—HPLC fractions containing peptides from the cellular extracts were sequenced by quadrupole ion trap electrospray tandem mass spectrometry in a Deca XP LCQ mass spectrometer (Finnigan ThermoQuest, San José, CA) (19). The charge and the mass of the ionic species were determined by high resolution sampling of the mass/charge rank. Collision energy and ion-precursor resolution were improved to optimize the fragmentation spectrum. Some synthetic peptides derived from gp160 were analyzed, and their detection limit was 1 fmol (G9I), 10 fmol (R10I and Q11I), and 50 fmol (K15I).

MHC-Peptide Stability Assay—TAP-deficient T2 cells transfected with L^d were cultured at 26 °C, and 14 h later they were washed and incubated for 2 h at 37 °C with 500 μM concentrations of the different synthetic peptides, as described (19). After washing (time point 0), the cells were further incubated at 37 °C. Aliquots removed at different time points were stained with monoclonal antibody 30-5-7S, which recognizes L^d bound to peptides (30), followed by flow cytometry. Cells incubated without peptide had peak fluorescence intensities close to background staining with second Ab alone. Fluorescence index was calculated at each time point as the ratio of peak channel fluorescence of the sample to that of the control incubated without peptide.

TAP Transport Assay—A TAP translocation assay was performed as described (31) using 5×10^6 murine A20 cells permeabilized with streptolysin O and the reporter peptide RYWANATRSF (R10F) that has an acceptor sequence for glycosylation once transported into the ER (32). R10F was labeled with ^{125}I by using chloramine T (33), and ENV synthetic peptides were used as competitors. Affinities of competitor peptides were expressed as 1/IC$_{50}$.

RESULTS

Differential Recognition Patterns of gp160-specific CTL Lines—In BALB/c mice the CTL response against HIV-1 strain IIIB envelope glycoprotein is mainly restricted by the D^d class I molecule (34). A minor L^d-restricted CTL population was previously described (21). To characterize these responses gp160-specific CTL lines were generated that selectively recognized gp160 peptides presented either by D^d or by L^d (Fig. 1, *A* and *B*, respectively). Additional experiments performed with three other L^d-positive cell lines confirmed the L^d-restricted recognition of the latter CTL line (data not shown). Endogenously synthesized gp160 was also presented to either CTL line in cells that were infected with a rVV that expresses the native glycoprotein, rVV-ENV (Fig. 1, *C* and *D*). Generation of similar long term L^d-restricted gp160-specific CTL lines by selection with the G9I peptide was achieved several times over a period of several years. In accordance with published experience, they were less easy to establish from BALB/c mice than D^d-restricted gp160-specific CTL lines.

Titration curves with different synthetic peptides using L cells transfected with D^d or L^d were performed. The nonamer G9I, which contains

the canonical anchor residues for binding to L^d, was compared with peptides with N-terminal extensions (R10I) and deletions (P8I and G7I). G9I and R10I were equally efficiently recognized, whereas P8I was 1000-fold less antigenic, and G7I was not recognized (Table 1).

Comparable Presentation of gp160 Peptides by Mutants of the Ld Molecule—L cells transfected with different mutants of the L^d allele were used to characterize the L^d-restricted response against gp160. The altered positions in each cellular line are detailed in its respective plot in Fig. 2 and include both single and multiple mutations located in the peptide binding groove. They were chosen to mimic the six changes that exist between L^d and L^q, its closest allele (25). Thus, the cells that have the six positions altered actually express L^q.

Titration curves of G9I and R10I synthetic peptides are shown in Fig. 2. The mutation in residue 116 was the one that singly affected most presentations by MHC class I and recognition by the T-cell receptor. Its negative effect was not compensated by any additional mutation, as all transfectants with additional mutations besides that at 116 were negative. Changes at positions 95 and 97, also located on the floor of the groove as residue 116, caused a milder 10–100-fold decrease in antige-
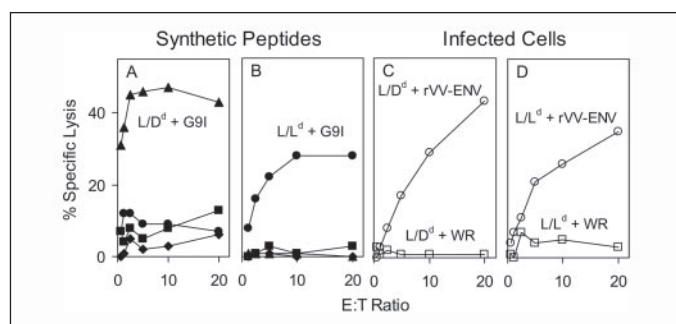
nicity. Residue 116 might be more critical because it contributes to pocket F that hosts the C terminus of the peptide (35) and that has been found also to affect CTL recognition of other epitopes presented by L^d (25). Recognition of rVV-ENV-infected transfectants with mutations at residues 116, 95, and 97 was clearly less efficiently than with the wild-type presenting molecule (data not shown), supporting the synthetic peptide data.

The same pattern of recognition was observed in all cases for both R10I and G9I peptides (Fig. 2). None of the mutations in the presenting molecule differentially affected presentation of either one of these envelope synthetic peptides to specific CTL. The simplest model to account for this observation is that both peptides bind with a similar conformation to L^d, as all L^d mutations had the same effect on presentation of either peptide. Because G9I has the canonical length and P2 and P9 anchor residues to L^d, the data suggest that peptide R10I binds as G9I does and that the R residue in R10I extends N-terminally out of the MHC class I binding groove.

Physiological Processing Generates at Least Three Different L^d-associated Peptidic Species—To identify the natural endogenously processed peptides of the gp160 glycoprotein that are associated with L^d, the peptides generated after infection of L/L^d cells with rVV-ENV were acid-extracted. The peptides were separated by RP HPLC, and the fractions collected were analyzed in a cytotoxicity assay, detecting a broad antigenic peak (data not shown). Whole cells rather than antibody-selected L^d cells were used for natural peptide extraction in order not to exclude potential precursor peptides. Because of this, it was interesting to test whether this antigenic activity was detected only when the L^d molecule was present. Consequently, extraction experiments were carried out using the L^d-negative parental Ltk$^-$ cells using similar infected cell equivalents as tested previously with L^d-positive cells. No antigenic peaks were found in the RP HPLC runs from infected Ltk$^-$ cells (data not shown) (19). These results indicate that peptides generated by endogenous processing of the envelope glycoprotein in infected L/L^d cells must be bound to the presenting L^d molecule and are, therefore, neither free from MHC nor bound also to the endogenous H-2^k molecules of L cells.

FIGURE 1. **Different MHC-restricted CTL populations recognize gp160.** L/D^d and L/L^d target cells, preincubated (▲, ●) or not (♦, ■) with 10^{-5} M G9I, were used in cytotoxic assays with gp160-specific CTL lines. Recognition by a D^d-restricted (*A*) or by a L^d-restricted (*B*) CTL line is shown. L/D^d (*C*) or L/L^d cells (*D*) were infected or not with 10 plaque-forming units/cell of rVV-ENV (○) or with the negative control Western Reserve (*WR*) (□) and assayed with D^d-restricted CTL or L^d-restricted CTL, respectively.

TABLE 1
Functional properties of synthetic peptides and overview of results

Synthetic peptide		Termini[a]		Coelution with antigenicity in IE HPLC[b]	Recognition by CTL[c]	L^d/peptide stability[d]
Name	Sequence	N end	C end			
G9I	**GPGRAFVTI**			**1st**	++++	+
P8I	PGRAFVTI	Deleted		1st	+	−
G7I	GRAFVTI	Deleted		1st	−	−
G10G	**GPGRAFVTI**G		Extended	1st	−	ND[e]
R11G	R**GPGRAFVTI**G	Extended	Extended	2nd	−	ND
Q12G	QR**GPGRAFVTI**G	Extended	Extended	2nd	−	ND
G11K	**GPGRAFVTI**GK		Extended	None	++[f]	ND
R12K	R**GPGRAFVTI**GK	Extended	Extended	None	+[f]	ND
I14K	IQR**GPGRAFVTI**GK	Extended	Extended	2nd	−	ND
R15K	RIQR**GPGRAFVTI**GK	Extended	Extended	3rd	ND	−
R10I	R**GPGRAFVTI**	Extended		**2nd**	++++	+
Q11I	QR**GPGRAFVTI**	Extended		**2nd**	++++	+
I12I	IQR**GPGRAFVTI**	Extended		None	++++	+
R13I	RIQR**GPGRAFVTI**	Extended		None	+++	+
K15I	KIRIQR**GPGRAFVTI**	Extended		**3rd**	+++	+
G9I-A	**GA GRAFVTI**			1st	ND	−
R10I-A	RGA**GRAFVTI**			2nd	ND	−
K15I-A	KIRIQRGA**GRAFVTI**			3rd	ND	−

[a] The terms deleted or extended are in reference to the termini of the highlighted G9I peptide (first peptide in the Table). The name and coelution of all natural peptides identified in this study are highlighted.

[b] Three antigenic peaks were detected in IE HPLC runs of infected cell extracts, as shown in Fig. 3, and are referred here to as the 1st, 2nd, and 3rd peaks. When synthetic peptides were individually analyzed by IE HPLC, their elution in any of the fractions of any of these peaks is scored here as coelution. Most peptides with 1, 2, or 4 positive charges coelute with the 1st, 2nd, or 3rd peak, respectively (19).

[c] Peptides were serially diluted in 10-fold steps and sequentially incubated with targets and CTL. Each additional + sign indicates a 10-fold higher antigenicity of the respective peptide.

[d] Relevant MHC/peptide stability assays are displayed in Fig. 5.

[e] ND, not done.

[f] A peptidyl-carboxydipeptidase activity present in fetal calf serum removes the GK C-terminal extension (19, 37) and artificially improves antigenicity.
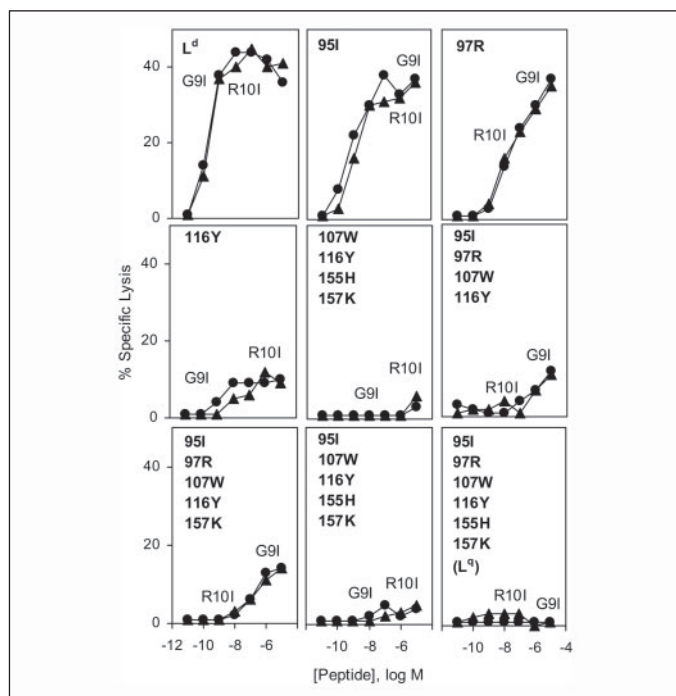
FIGURE 2. **Identical effect of mutations in the L^d molecule on exogenous presentation of gp160 synthetic peptides.** L cells transfected with the L^d molecule or with mutants in different residues as indicated were incubated with serial dilutions of G9I (●) or R10I (▲) synthetic peptides and then with L^d-restricted CTL at an E:T ratio of 10:1.

FIGURE 3. **At least three natural peptides generated by endogenous gp160 antigen processing.** Peptides extracted from rVV-ENV-infected L/L^d cells were separated by RP HPLC. Antigenic fractions 13–22 from this column were pooled and separated by IE HPLC. This was followed by a cytotoxicity assay of the collected fractions. The results obtained with 1.5×10^8 (●), 0.18×10^8 (▲), and 0.09×10^8 (■) cell equivalents of the peptide extracts are plotted. At the right, the secondary scale indicates the monitored conductivity of the NaCl gradient (*dotted line*) employed for elution.

FIGURE 4. **Identification by mass spectrometry of G9I in infected cell extracts.** Tandem mass spectrometry fragmentation spectrum obtained after quadrupole ion trap mass spectrometry of fraction 22 of the IE HPLC run of extracts of L/L^d cells infected with rVV-ENV and shown in Fig. 3. The bicharged parental ion corresponding to the G9I peptide has a mass/charge ratio of 458.4 (*arrow*). Ions generated in the fragmentation are detailed, whereas the sequence deduced from the indicated fragments is shown in the *lateral box*.

Even though the gradient used for elution of the natural peptides from the RP HPLC column was rather flat, several different overlapping synthetic peptides from the antigenic region encompassing gp160 residues 313–337 eluted within the antigenic peak (data not shown). This is probably because all these peptides have related, low hydrophobicity, as all of them are positively charged. To determine the peptidic species present in the antigenic peak obtained in the RP HPLC analysis of infected cells, the pool of positive fractions was further separated by IE HPLC. The fractions were again analyzed by cytotoxicity assays. Fig. 3 shows three different antigenic peaks generated as a result of the endogenous processing of the envelope glycoprotein in L/L^d cells.

The first antigenic peak (fractions 21–23) coelutes with monocharged species. Quadrupole/ion trap mass spectrometry analysis showed that fractions 21–23 recognized by CTL unequivocally contained G9I peptide (Fig. 4). Mass spectrometry fragmentation of the synthetic peptide G9I gave a very similar fragmentation profile, confirming identification (data not shown). Thus, the optimal and minimal G9I is a natural L^d ligand in infected cells.

The second peak in Fig. 3 (fractions 36–40) from infected L/L^d cells mainly groups some bi-charged peptides. These first two peaks surprisingly correspond to the elution times of the peptides derived from this epitope that we have previously described as natural D^d ligands (19). As in that report, five candidate coeluting synthetic peptides that contained the minimal G9I antigenic core were considered. Three, R11G, Q12G, and I14K, were excluded as natural L^d ligands because they were not recognized as synthetic peptides by CTL (Table 1). In contrast, the other two candidate synthetic peptides coeluting with the second peak in Fig. 3, the singly and doubly N-terminal-extended peptides R10I and Q11I, were as antigenic as the minimal G9I nonamer (Fig. 5A). Unfortunately, attempts to identify R10I or Q11I peptides (both with two charges) analyzing the second antigenic peak by mass spectrometry were compromised by the lack of sensitivity. Indeed, fragmentation of synthetic R10I and Q11I peptides is at least 10-fold less efficient than that of
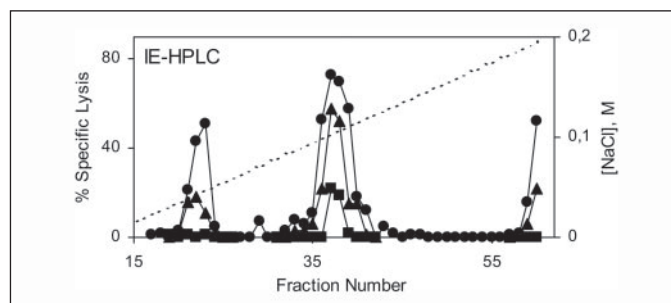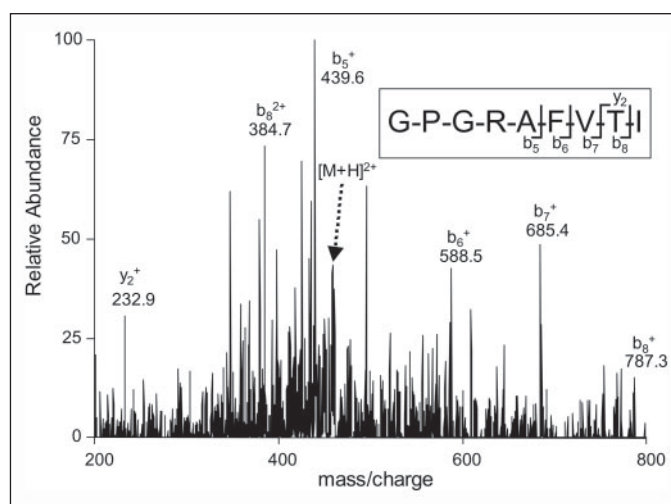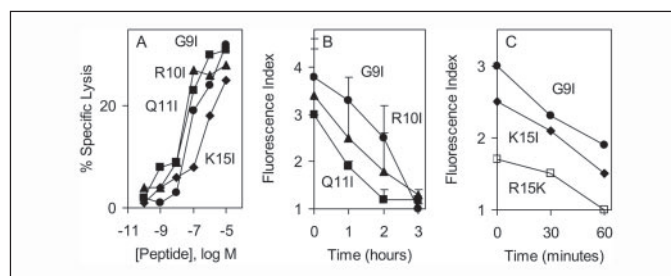

FIGURE 5. **Antigenicity and MHC stability assay of several gp160 synthetic peptides.** *A*, recognition by CTL of titration curves of G9I (●), R10I (▲), Q11I (■), and K15I (♦) synthetic peptides on L/L^d cells. *B* and *C*, stability of cell surface MHC-peptide complexes on TAP-deficient T2/L^d cells was assessed by flow cytometry. □, R15K. *Bars*, S.D.

synthetic G9I (see "Experimental Procedures"). The presence of an additional Arg in these two peptides as compared with G9I may have contributed to further decreased efficiency of detection of fragmented ions, already close to the detection limit for the natural G9I peptide. An even stronger lack of sensitivity also compromised positive identification of K15I by mass spectrometry (see next paragraph). We concluded

that either R10I or Q11I or both are naturally processed and presented by L^d in infected cells.

An Unexpectedly Long, N-terminal-extended Peptide Is Physiologically Processed from HIV-1 Envelope Protein—Finally, the third antigenic peak extracted from infected cells and shown in Fig. 3 (fractions 59–60) coelutes with several V3-loop synthetic peptides with as many as 4 positive charges. This constitutes a set of 18 different candidate coeluting synthetic peptides 15–20 aa in length that contain the minimal G9I antigenic core. All have double N- and C-terminal extensions of the optimal G9I natural ligand, with the single exception of peptide K15I, with a 6-aa N-terminal extension. Analysis started with the latter. Fig. 5*A* shows that synthetic K15I was only 10-fold less antigenic for G9I-selected CTL that the core G9I peptide. This suggested that K15I may be the naturally processed peptide in this HPLC peak. Because it is an unexpectedly long candidate MHC class I ligand, we made sure that no traces of contamination with other antigenic peptides were responsible for this high antigenicity. Even after overloading an IE HPLC column, all activity coeluted with the single A_{280} peak of the purified K15I peptide (Fig. 6*A*).

We next used TAP-deficient T2 cells transfected with L^d to assay the relative stability of the MHC class I complexes with the different gp160 peptides. It was found that G9I, R10I, Q11I, and K15I peptides induced similar numbers of similarly stable L^d-peptide surface complexes (Fig. 5, *B* and *C*).

Exchange for alanine of the proline residue that serves as the canonical anchor abolished interaction with L^d of the N-extended K15I and R10I peptides as well as of G9I (Table 1). This indicates that they bind to L^d using the canonical anchors and, thus, with an N terminus extending out of the L^d peptide groove. Binding to several MHC class I allotypes of non-natural peptides with N-terminal extensions out of the groove has been shown before to be tolerated without significant rearrangement in the MHC structure (36). Other peptides with N-terminal extensions shorter than the 6 aa, present in K15I but that did not coelute with any natural antigenic activity from infected cells, were also remarkably antigenic for the G9I-selected CTL lines (Table 1), strongly supporting that K15I and the other extended variants bound to L^d and were recognized by CTL in the same conformation as R10I and G9I. Collectively, these results strongly suggest that K15I is naturally processed and presented by L^d in infected cells expressing HIV ENV.

As opposed to K15I, N-extended peptides I14I and R13I as well as I12I, which lack one or two positive charges with respect to K15I, respectively, could be excluded as natural peptides. Indeed, no antigenic activity was found around fraction 45 or 30 of the IE HPLC run of the infected cell extracts (Fig. 3), where they elute, respectively.

The remaining four-charged synthetic peptides coeluting with the third antigenic peak in Fig. 3 had both N- and C-terminal extensions. The shortest was R15K and was analyzed next. As a synthetic peptide, R15K did not coelute with the antigenic peak in the RP HPLC column that was run as the first purification step of the infected cell extracts. In addition, in contrast to K15I, R15K formed complexes with L^d with poor efficiency (Fig. 5*C*). Moreover, part of its activity is actually due to conversion to the more antigenic R13I peptide by an angiotensin-converting enzyme-like present in serum. This carboxy-dipeptidase activity removes pairs of C-terminal residues, as described previously (37) and confirmed by us by IE HPLC (data not shown). This was also the case for other peptides with the same two-residue GK carboxyl extension (Table 1, footnote d) but not for K15I (data not shown). Peptides with this ENV sequence and with C-terminal extensions longer than 2 aa do not profit from angiotensin-converting enzyme activity to enhance their increasingly poorer antigenicity (38). Already, a single C-terminal extension of
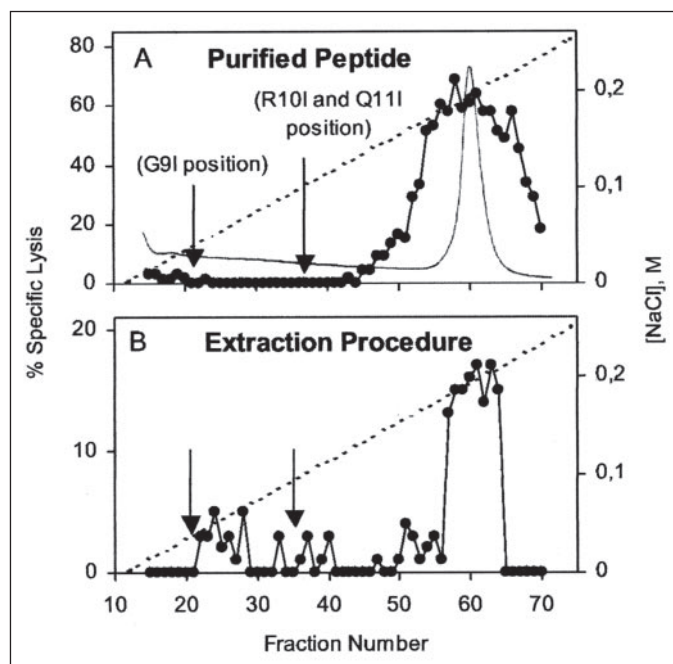

FIGURE 6. **The long synthetic peptide K15I is not contaminated and is not modified by the extraction procedure.** *A*, purified peptide K15I (0.52 nmol) was separated by IE HPLC and monitored by A_{280} (*solid line*), and fractions were collected and assayed with CTL (●) at a resulting concentration per assay well of 10^{-5} M. The elution positions of G9I and bicharged R10I or Q11I peptides are also indicated. At the *right*, the secondary scale indicates the monitored conductivity of the NaCl gradient (*dotted line*) employed for elution. *B*, peptides were extracted from uninfected L^d-positive P13.1 cells to which 0.59 nmol of synthetic K15I had been added before extraction and separated by RP HPLC. Antigenic fractions were selected, further separated by IE HPLC, and tested with CTL.

the minimal G9I peptide, as in peptide G10G, showed the lack of tolerance of L^d and CTL to such extensions (Table 1). As expected, several peptides with double N and C extensions were also very poorly presented by L^d (Table 1). Because of this, we excluded R15K and all other longer, doubly extended, 4-charged peptides as potential natural peptides. In summary these results support the conclusion that, in addition to G9I and R10I/Q11I peptides, the N-terminal-extended K15I peptide is a natural L^d ligand derived from HIV envelope glycoprotein.

Serial dilutions of IE HPLC antigenic fractions indicated comparable antigenic activity for all of them (Fig. 3). Because G9I, R10I, and Q11I were similarly recognized by L^d-restricted CTL, whereas K15I was 10-fold less efficiently recognized, the results suggest that gp160-derived L^d ligands comprise equivalent amounts of the former three peptides and around 10-fold more of the much longer K15I peptide. It is uncertain which fraction of each peptide-MHC complex is located intracellularly or at the cell surface as infection proceeds.

The Shorter Peptidic Species Are Not Generated from the Longer Ones by the Extraction Procedure—Because the K15I natural peptide contains and is longer than the mono-charged G9I and the two bi-charged peptides, the latter three could potentially be artificially generated from K15I by an aminopeptidase activity or by a chemical reaction during the biochemical isolation of naturally processed peptides. To exclude this possibility, similar extraction experiments as those performed with rVV-ENV-infected cells were carried out with K15I-pulsed L^d-positive target cells. After RP and IE HPLC, collected fractions from this control experiment were tested with CTL. Activity was only detected at the expected elution position for K15I. No production of shorter peptides was found (Fig. 6*B*). In addition, in a similar experiment with R10I and Q11I synthetic peptides, no evidence of generation of G9I from these two peptides during extract preparation nor alteration of the IE HPLC
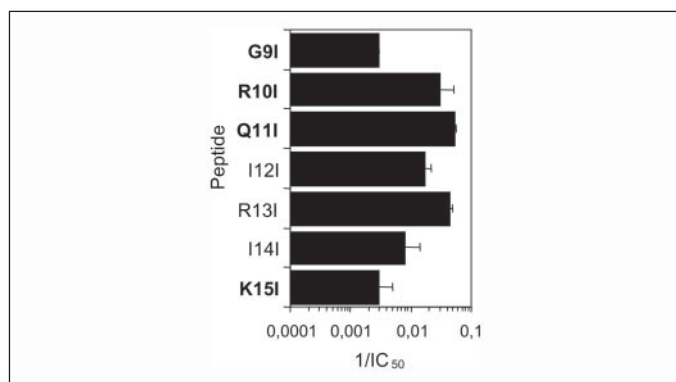
FIGURE 7. **Transport by TAP.** The indicated synthetic peptides were assayed for transport by murine TAP in an *in vitro* competition assay. Affinities of competitor peptides are expressed as $1/IC_{50}$.

the long ER residence time of L^d as well as the lack in murine cells of the counterpart gene for human ERAP2, which by synergizing with ERAP1 increases the efficiency of trimming this HIV epitope *in vitro* (17). The presence of N-terminal extensions in all these high affinity natural ligands in the face of an aminopeptidase activity at the site of complex formation indicates that trimming, at least for the HIV ENV N-extended peptides described here, cannot proceed significantly once the complexes are formed.

Although this is the first report of an N-terminal- elongated nested set of natural MHC class I ligands, there are six previous reports of C-terminal-elongated peptides that also appear to bind to MHC class I molecules with protruding extensions (7, 8, 40–43). This appears as a natural consequence of the reported absence of C-terminal trimming activity in the ER (1, 5). As opposed to our results, the antigenicity of the elongated peptides was either not tested or at least 250-fold lower than that of the standard minimal peptide, which leaves open their quantitative contribution to the CTL response. Notably, five of the six reports on C-terminal extensions involve human MHC ligands, mostly of A2. In contrast, this report reveals N-terminal-extended natural ligands in murine cells. It is intriguing to recall that human cells have an additional gene for an aminopeptidase, ERAP2 (16), that operates in a concerted fashion with ERAP1 (17). Thus, it is uncertain whether high affinity N-extended long peptides similar to those described by us in this report and suggested previously (19) will also quickly and efficiently be protected from trimming by MHC class I molecules in human cells. Therefore, it remains to be established whether this abundant presence of highly antigenic N-terminal-extended ligands is an exceptional behavior of L^d, whether it is more general to mouse cells and MHC class I molecules, or whether it extends to other species.

We identified peptides with 1–2- and 6-aa-long N-terminal extensions of the minimal and optimal HIV epitope as natural L^d ligands. It was surprising that those peptides with 3- or 4-aa extensions, namely I12I, and R13I, which were as antigenic as the natural ligands, were not present in infected cells. Transport by TAP did not seem to be very discriminating in this regard. The N-extended natural peptides persisted despite ER trimming. In addition, at least the spectra of products generated *in vitro* either by human ERAP1 alone or by ERAP1 and ERAP2 (17) do not coincide with the spectrum and relative abundance of natural peptides that we found in infected cells. Thus, we conclude that it is more likely that ER aminopeptidase trimming does not play a major role in determining the composition of the pool of natural L^d ligands derived from this epitope. We suggest that cytosolic processing and trimming coupled with resistance to cytosolic degradation (11, 44) are probably the major determinants of the final composition of gp160-derived peptides presented by L^d.

The most notable exception was peptide R13I, which is transported best, most resistant to trimming (17), and as antigenic as the natural ligand K15I. Yet, it is not a natural L^d ligand because it does not coelute with any natural antigenic peak. Thus, it is probably neither produced in the cytosol nor in the ER. Otherwise, it would have probably bound to L^d. Simple experiments with synthetic peptides would have qualified R13I as a very probable natural peptide. Thus, it represents a paradigm where analysis of natural endogenous peptides is critically required to unequivocally establish the physiological relevance of a group of very antigenic synthetic peptides.

The complex mixture of natural nested ligands for L^d parallels and expands the natural ligands of D^d from the same epitope (19), suggesting that antigen processing gives rise to a rich pool of possible ligands from which MHC class I molecules can choose. In these two examples we show at the same time the extended diversity of peptides generated by

elution pattern was observed either (data not shown) (19). These data, thus, confirm the endogenous generation of G9I and R10I or Q11I in rVV-ENV-infected cells, reinforcing their identification as candidate natural peptides in infected cells.

Transport to the ER by TAP—Next, efficiency of TAP-mediated transport into the ER of all peptides with 1–6-aa N-terminal extensions was assayed *in vitro*. As shown in Fig. 7, all peptides were transported by TAP with a low to intermediate affinity. Of note, two of the natural ligands, G9I and K15I, were among the peptides least efficiently transported by murine TAP but still with an affinity compatible with their physiological transport in infected cells (39).

DISCUSSION

The analysis of the natural L^d peptidic ligands resulting from the endogenous processing of the HIV envelope glycoprotein shows the presence of at least three different MHC-peptide complexes in infected cells. The first peptide is identified by mass spectrometry as the nonamer G9I ([319]GPGRAFVTI[327]) with the canonical anchor motif for binding to L^d at positions 2 and 9. The second species corresponds to the N-terminal-extended R10I 10-mer or Q11I 11-mer, which are as antigenic and abundant as the canonical G9I. The third peptidic species corresponds to the 15-mer K15I peptide, the longest MHC class I natural ligand recognized by CTL described so far. The peptide has an N-terminal extension of 6 residues, binds to the presenting molecule L^d with an antigenicity for G9I-selected CTL only slightly lower than the optimal 9-mer, and is more abundant in infected cells than the minimal epitope.

To our knowledge this is the first report of such a complex N-extended nested set of natural peptides with related antigenicity, including such a long naturally processed antigenic peptide, K15I. The underlying reason for our findings may be the exceptionally high affinity for L^d of the longer N-terminally extended gp160 peptides. Although L^d seems to need trimming by ERAAP for providing roughly one-half of its regular epitopes (13), the K15I to R10I peptides would rather contribute to the fraction of ERAAP-independent complexes, as they are spared from trimming *in vivo* and presented by L^d. Thus, ERAAP-independent MHC class I ligands appear to include not only peptides already produced in their final size in the cytosol but also long high affinity peptides as the two/three identified in this report. Therefore, the working hypothesis is that these N-extended peptides will escape trimming because they bind faster and more quantitatively to L^d than do other sets of lower affinity extended epitopes to other MHC class I allotypes, as shown for an undecapeptide (31). Further possible contributing factors might be

processing (45) as well as the limitations imposed by the selectivity of the combined specificities of the generative and destructive proteases in the antigen processing pathway (43). The existence of a collection of potential MHC ligands may underlie the promiscuous presentation by 5 human, 1 chimpanzee, 1 macaque, and 6 murine MHC class I allotypes of peptides from this very same HIV epitope (46). Our results emphasize the complexity of antigen presentation, which is sometimes minimized in epitope detection assays with synthetic peptides. It remains to be proven whether nested sets of natural peptides are the rule or the exception. It is important to note that, as opposed to the generalized use of low resolution, single column HPLC in our recent reports and in some of the best characterized C-terminal-extended nested sets of peptides (8, 43), high resolution HPLC techniques have been applied. However, it is also true that extensions by as little as one residue on either end of the peptide usually result in marked drops in affinity to MHC. Here we show that a 10-fold drop in antigenicity as in the K15I natural ligand is compatible with abundant MHC-peptide complexes *in vivo*.

Analysis of the cellular peptides eluted from L^d molecules was restricted to those 8–11 aa in length (20). Yet, only 9- or 10-aa-long peptides with both P2 and carboxyl anchors were identified that differed only in the degree of central bulging. This is the most frequent structural adaptation for accommodating peptides up to 14 residues in length and with good affinity by MHC class I molecules (9, 42, 47–51) and always has a drastic effect on CTL recognition, giving rise to fully new CTL specificities. Thus, the occurrence of K15I and R10I or Q11I as natural peptides would not have been predicted from these reports. Several lines of results indicated that R10I binds with the same anchors as the canonical G9I, with the N-terminal Arg residue extending out of the groove. The lack of alternative anchor motifs to L^d as well as the high antigenicity of K15I with G9I-selected CTL also strongly suggests this mode of binding. Because the K15I N-extended sequence of six residues is very polar, it may be stable in solution without additional interactions with the external regions of the L^d α_1 and α_2 domains or, later, with the T-cell receptor although these are not ruled out and might contribute to complex stability and antigenicity.

Perhaps the most relevant aspects of our work are the abundant presence of several N-terminal-extended MHC class I ligands in the face of ER aminopeptidase activity as well as the high diversity of physiological MHC-peptide complexes. Because we obtain CTL restricted by L^d or by D^d from the same animal, we assume that several or all 5–7 natural peptide-MHC complexes may be present simultaneously in infected or cross-presenting cells *in vivo*. This would clearly be an advantage for the immune system, particularly to avoid focusing the CTL response to few complexes, which favors the appearance of CTL escape virus mutants. A detailed knowledge of MHC class I peptide ligands and their intracellular generation should be relevant in the development of an HIV vaccine including CTL epitopes.

REFERENCES

1. Shastri, N., Schwab, S., and Serwold, T. (2002) *Annu. Rev. Immunol.* **20,** 463–493
2. Parker, K. C., Bednarek, M. A., and Coligan, J. E. (1994) *J. Immunol.* **152,** 163–175
3. Rammensee, H. G., Bachmann, J., Emmerich, N. N., Bachor, O. A., and Stevanovic, S. (1999) *Immunogenetics* **50,** 213–219
4. Cascio, P., Hilton, C., Kisselev, A. F., Rock, K. L., and Goldberg, A. L. (2001) *EMBO J.* **20,** 2357–2366
5. Powis, S. J., Young, L. L., Joly, E., Barker, P. J., Richardson, L., Brandt, R. P., Melief, C. J., Howard, J. C., and Butcher, G. W. (1996) *Immunity* **4,** 159–165
6. Henderson, R. A., Michel, H., Sakaguchi, K., Shabanowitz, J., Appella, E., Hunt, D. F., and Engelhard, V. H. (1992) *Science* **255,** 1264–1266
7. Herr, W., Ranieri, E., Gambotto, A., Kierstead, L. S., Amoscato, A. A., Gesualdo, L., and Storkus, W. J. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 12033–12038
8. Lucchiari-Hartz, M., van Endert, P. M., Lauvau, G., Maier, R., Meyerhans, A., Mann, D., Eichmann, K., and Niedermann, G. (2000) *J. Exp. Med.* **191,** 239–252
9. Flyer, D. C., Ramakrishna, V., Miller, C., Myers, H., McDaniel, M., Root, K., Flournoy, C., Engelhard, V. H., Canaday, D. H., Marto, J. A., Ross, M. M., Hunt, D. F., Shabanowitz, J., and White, F. M. (2002) *Infect. Immun.* **70,** 2926–2932
10. Princiotta, M. F., Finzi, D., Qian, S. B., Gibbs, J., Schuchmann, S., Buttgereit, F., Bennink, J. R., and Yewdell, J. W. (2003) *Immunity* **18,** 343–354
11. Reits, E., Griekspoor, A., Neijssen, J., Groothuis, T., Jalink, K., van Veelen, P., Janssen, H., Calafat, J., Drijfhout, J. W., and Neefjes, J. (2003) *Immunity* **18,** 97–108
12. Fruci, D., Niedermann, G., Butler, R. H., and van Endert, P. M. (2001) *Immunity* **15,** 467–476
13. Serwold, T., González, F., Kim, J., Jacob, R., and Shastri, N. (2002) *Nature* **419,** 480–483
14. York, I. A., Chang, S. C., Saric, T., Keys, J. A., Favreau, J. M., Goldberg, A. L., and Rock, K. L. (2002) *Nat. Immunol.* **3,** 1177–1184
15. Saric, T., Chang, S. C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M., and Goldberg, A. L. (2002) *Nat. Immunol.* **3,** 1169–1176
16. Tanioka, T., Hattori, A., Masuda, S., Nomura, Y., Nakayama, H., Mizutani, S., and Tsujimoto, M. (2003) *J. Biol. Chem.* **278,** 32275–32283
17. Saveanu, L., Carroll, O., Lindo, V., Del Val, M., López, D., Lepelletier, Y., Greer, F., Schomburg, L., Fruci, D., Niedermann, G., and van Endert, P. M. (2005) *Nat. Immunol.* **6,** 689–697
18. Brouwenstijn, N., Serwold, T., and Shastri, N. (2001) *Immunity* **15,** 95–104
19. Samino, Y., López, D., Guil, S., de León, P., and Del Val, M. (2004) *J. Biol. Chem.* **279,** 1151–1160
20. Corr, M., Boyd, L. F., Frankel, S. R., Kozlowski, S., Padlan, E. A., and Margulies, D. H. (1992) *J. Exp. Med.* **176,** 1681–1692
21. Bergmann, C., Tong, L., Cua, R. V., Sensintaffar, J. L., and Stohlman, S. A. (1994) *J. Immunol.* **152,** 5603–5612
22. Carbone, F. R., and Bevan, M. J. (1990) *J. Exp. Med.* **171,** 377–387
23. Abastado, J. P., Jaulin, C., Schutze, M. P., Langlade-Demoyen, P., Plata, F., Ozato, K., and Kourilsky, P. (1987) *J. Exp. Med.* **166,** 327–340
24. Del Val, M., Volkmer, H., Rothbard, J. B., Jonjic, S., Messerle, M., Schickedanz, J., Reddehase, M. J., and Koszinowski, U. H. (1988) *J. Virol.* **62,** 3965–3972
25. Hengel, H., Burke, K., Kyburz, D., Zinkernagel, R. M., and Koszinowski, U. H. (1995) *J. Immunol.* **154,** 4557–4564
26. Crumpacker, D. B., Alexander, J., Cresswell, P., and Engelhard, V. H. (1992) *J. Immunol.* **148,** 3004–3011
27. Chakrabarti, S., Robert-Guroff, M., Wong-Staal, F., Gallo, R. C., and Moss, B. (1986) *Nature* **320,** 535–537
28. López, D., Gil-Torregrosa, B. C., Bergmann, C., and Del Val, M. (2000) *J. Immunol.* **164,** 5070–5077
29. Del Val, M., Schlicht, H. J., Ruppert, T., Reddehase, M. J., and Koszinowski, U. H. (1991) *Cell* **66,** 1145–1153
30. Ozato, K., Hansen, T. H., and Sachs, D. H. (1980) *J. Immunol.* **125,** 2473–2477
31. Lauvau, G., Kakimi, K., Niedermann, G., Ostankovitch, M., Yotnda, P., Firat, H., Chisari, F. V., and van Endert, P. M. (1999) *J. Exp. Med.* **190,** 1227–1240
32. Momburg, F., Roelse, J., Howard, J. C., Butcher, G. W., Hammerling, G. J., and Neefjes, J. J. (1994) *Nature* **367,** 648–651
33. van Endert, P. M., Tampe, R., Meyer, T. H., Tisch, R., Bach, J. F., and McDevitt, H. O. (1994) *Immunity* **1,** 491–500
34. Takahashi, H., Cohen, J., Hosmalin, A., Cease, K. B., Houghten, R., Cornette, J. L., DeLisi, C., Moss, B., Germain, R. N., and Berzofsky, J. A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 3105–3109
35. Balendiran, G. K., Solheim, J. C., Young, A. C. M., Hansen, T. H., Nathenson, S. G., and Sacchettini, J. C. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6880–6885
36. Achour, A., Michaelsson, J., Harris, R. A., Odeberg, J., Grufman, P., Sandberg, J. K., Levitsky, V., Karre, K., Sandalova, T., and Schneider, G. (2002) *Immunity* **17,** 757–768
37. Kozlowski, S., Corr, M., Takeshita, T., Boyd, L. F., Pendleton, C. D., Germain, R. N., Berzofsky, J. A., and Margulies, D. H. (1992) *J. Exp. Med.* **175,** 1417–1422
38. Kozlowski, S., Corr, M., Shirai, M., Boyd, L. F., Pendleton, C. D., Berzofsky, J. A., and Margulies, D. H. (1993) *J. Immunol.* **151,** 4033–4044
39. Daniel, S., Caillat-Zucman, S., Hammer, J., Bach, J. F., and van Endert, P. M. (1997) *J. Immunol.* **159,** 2350–2357
40. Frumento, G., Harris, P. E., Gawinowicz, M. A., Suciu-Foca, N., and Pernis, B. (1993) *Cell. Immunol.* **152,** 623–626
41. Collins, E. J., Garboczi, D. N., and Wiley, D. C. (1994) *Nature* **371,** 626–629
42. Chen, Y., Sidney, J., Southwood, S., Cox, A. L., Sakaguchi, K., Henderson, R. A., Appella, E., Hunt, D. F., Sette, A., and Engelhard, V. H. (1994) *J. Immunol.* **152,** 2874–2881
43. Serwold, T., and Shastri, N. (1999) *J. Immunol.* **162,** 4712–4719
44. York, I. A., Mo, A. X. Y., Lemerise, K., Zeng, W. Y., Shen, Y. L., Abraham, C. R., Saric,

T., Goldberg, A. L., and Rock, K. L. (2003) *Immunity* **18,** 429–440

45. Paz, P., Brouwenstijn, N., Perry, R., and Shastri, N. (1999) *Immunity* **11,** 241–251

46. Korber, B. T. M., Brander, C., Haynes, B. F., Koup, R., Moore, J. P., Walker, B. D., and Watkins, D. (2005) *Theoretical Biology and Biophysics Group*, pp. LA-UR 06-0036, Los Alamos National Laboratory, Los Alamos, NM

47. Guo, H. C., Jardetzky, T. S., Garrett, T. P., Lane, W. S., Strominger, J. L., and Wiley, D. C. (1992) *Nature* **360,** 364–366

48. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A., and Wilson, I. A. (1992)

Science **257,** 919–927

49. Speir, J. A., Stevens, J., Joly, E., Butcher, G. W., and Wilson, I. A. (2001) *Immunity* **14,** 81–92

50. Probst-Kepper, M., Stroobant, V., Kridel, R., Gaugler, B., Landry, C., Brasseur, F., Cosyns, J. P., Weynand, B., Boon, T., and Van den Eynde, B. J. (2001) *J. Exp. Med.* **193,** 1189–1198

51. Probst-Kepper, M., Hecht, H. J., Herrmann, H., Janke, V., Ocklenburg, F., Klempnauer, J., Van den Eynde, B. J., and Weiss, S. (2004) *J. Immunol.* **173,** 5610–5616

A Long N-terminal-extended Nested Set of Abundant and Antigenic Major Histocompatibility Complex Class I Natural Ligands from HIV Envelope Protein

Yolanda Samino, Daniel López, Sara Guil, Loredana Saveanu, Peter M. van Endert and Margarita Del Val

Access the most updated version of this article at doi: 10.1074/jbc.M512263200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 50 references, 28 of which can be accessed free at
http://www.jbc.org/content/281/10/6358.full.html#ref-list-1