

Genome-wide Characterization of a Viral Cytotoxic T Lymphocyte Epitope Repertoire*[§]

Received for publication, July 10, 2003, and in revised form, August 1, 2003
Published, JBC Papers in Press, September 5, 2003, DOI 10.1074/jbc.M307417200

Weimin Zhong, Pedro A. Reche, Char-Chang Lai, Bruce Reinhold, and Ellis L. Reinherz[‡]

From the Laboratory of Immunobiology and Department of Medical Oncology, Dana-Farber Cancer Institute and Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115

A genome-wide search using major histocompatibility complex (MHC) class I binding and proteasome cleavage site algorithms identified 101 influenza A PR8 virus-derived peptides as potential epitopes for CD8⁺ T cell recognition in the H-2^b mouse. Cytokine-based flow cytometry, ELISPOT, and cytotoxic T lymphocyte assays reveal that 16 are recognized by CD8⁺ T cells recovered directly *ex vivo* from infected animals, accounting for greater than 70% of CD8⁺ T cells recruited to lung after primary infection. Only six of the 22 highest affinity MHC class I binding peptides comprise cytotoxic T lymphocyte epitopes. The remaining non-immunogenic peptides have equivalent MHC affinity and MHC-peptide complex half-lives, eliciting T cell responses when given in adjuvant and with T cell receptor-ligand avidity comparable with their immunogenic counterparts. As revealed by a novel high sensitivity nanospray tandem mass spectrometry methodology, failure to process those predicted epitopes may contribute significantly to the absent response. These results have important implications for rationale design of CD8⁺ T cell vaccines.

ity is critical in the development of CD8⁺ T cell-based vaccines.

Of the many CTL epitopes potentially available in a viral genome, antiviral CD8⁺ T cell responses usually focus on only one or two immunodominant peptides (6). Apparently, the great majority of potential CTL epitopes are silent under physiological conditions. The mechanisms for this phenomenon, termed immunodominance, are poorly understood. Current rational design of CD8⁺ T cell vaccines is mainly focused on the induction of antiviral immunity by immunodominant CTL epitopes, assuming that protective CD8⁺ T cell immunity is mediated by recognition of immunodominant CTL epitopes. Such a vaccination regimen is indeed efficient in conferring immune protection against a variety of viral infections in murine models (7–9). However, many viruses, especially RNA viruses, evade immune clearance of CD8⁺ T cell-mediated mechanisms by mutating immunodominant CTL epitopes, as observed in a number of persistent, as well as acute, viral infections (10, 11).

To circumvent this escape mechanism, new vaccination strategies that target conserved CTL epitope regions of viral proteins are needed. Preliminary studies suggest that subdominant CTL epitopes may represent attractive potential candidates to induce a broader degree of antiviral CD8⁺ T cell immunity. It has been shown that both subdominant and even so-called “non-immunogenic” epitopes are capable of eliciting strong CD8⁺ T cell responses, if appropriately primed *in vivo* (12, 13). Furthermore, several recent studies (14, 15) have documented a partial immune protection against lethal viral challenge following immunization with subdominant or non-immunogenic CTL epitopes. Thus, these epitopes are capable of participating in the effector function of cell-mediated antiviral immune responses. However, selection of appropriate candidate peptides has been difficult, requiring comprehensive biochemical and functional characterization of the natural CTL epitope population of an infectious agent. Traditional strategies for CTL epitope identification, such as overlapping peptide synthesis and antigen screening using T cell hybridoma-based immune recognition strategies, have limited power.

In the present study, we explore combined bioinformatic and functional approaches to rapidly identify, from the genome of an infectious agent, CD8⁺ T cell peptide targets for CTL epitope-based vaccine design using the PR8 strain of influenza A virus as a model system. Initially, SYFPEITHI, a well established computer algorithm for prediction of MHC binding peptides (16), was used to pre-select all possible peptides with potential mouse D^b and K^b binding capacities from the genome of the influenza A virus. This method identified 148 potential binders. A second new computer algorithm, Probabilistic Model of Proteosomal Cleavage (PMPC), was then used to prospect the likelihood of proteasomal cleavage of the SYFPEITHI-predicted D^b and K^b binders. Of these, 101 peptides scored with a

Antigen-specific recognition of peptides bound to major histocompatibility complex class I (MHCI)¹ molecules is mediated by CD8⁺ T lymphocytes. The stretches of linear peptide sequence recognized by these T cells, referred to as T cell epitopes, generally consist of 8–11 contiguous amino acids (1, 2). CD8⁺ cytotoxic T lymphocyte (CTL) responses play a central role in protective immunity against viral and intracellular bacterial infections (3–5). Therefore, identification of MHCI-bound (*i.e.* “restricted”) T cell epitopes that can induce cellular immu-

* This work was supported by National Institutes of Health Grant AI50900 and an award from the Molecular Immunology Foundation. 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.

[§] The on-line version of this article (available at <http://www.jbc.org>) contains Supplemental Material and Supplemental Figs. S1–S11 and Tables S1 and S2.

[‡] To whom correspondence should be addressed: Dana-Farber Cancer Inst., 44 Binney St., Boston, MA 02115. Tel.: 617-632-3412; Fax: 617-632-3351; E-mail: ellis_reinherz@dfci.harvard.edu.

¹ The abbreviations used are: MHCI, major histocompatibility complex class I; CTL, cytotoxic T lymphocyte; PMPC, Probabilistic Model of Proteosomal Cleavage; HPLC, high pressure liquid chromatography; mAb, monoclonal antibody; ER, endoplasmic reticulum; pMHC, peptide/MHC complex; NA, neuraminidase; HA, hemagglutinin; HAU, HA unit; IFN γ , interferon γ ; NP, nucleoprotein; PA, polymerase acidic protein; BAL, bronchoalveolar lavage; SFC, spot-forming cells; PB, polymerase basic protein; rIL-2, recombinant human interleukin-2; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid; MS/MS, tandem mass spectrometry; NS, non-structural protein; M, matrix protein; TCR, T cell receptor; APC, antigen-presenting cell.

high likelihood to be proteolytically processed by eukaryotic proteasomes. Detailed functional analysis of all potential CTL epitopes revealed that 16 peptides, 10 of which have not been identified previously, could be recognized by CD8⁺ T cells recovered directly *ex vivo* from lungs of influenza A virus-infected B6 mice. This number of naturally processed and *in vivo* presented viral CTL epitopes is considerably larger than previously thought. By investigating peptide binding interaction with MHC molecules, T cell repertoire potential, and aspects of viral protein processing, parameters influencing immune epitope display and recognition are revealed.

EXPERIMENTAL PROCEDURES

CTL Epitope Prediction—All potential T cell epitopes encoded within the influenza A virus (strain A/Puerto Rico/8/34; PR8) proteome were predicted using the motif matrix-based algorithm SYFPEITHI (16) (syfpeithi.bmi-heidelberg.com/scripts/MHCServer.dll/home.htm). In the case of the K^b molecule, we chose a K^b-matrix specific for predicted 8-mers, given that most peptides that bind to K^b are of this length (17). As D^b binds both 9- and 10-mers (17), two D^b-specific matrices were used. We selected peptides that had a score of at least 51% of the optimum using the relevant matrices. The number of the predicted K^b and D^b binders was 73 and 75, respectively, corresponding to 1.6% of all possible peptides from the PR8 proteome (see Table I and Supplemental Table S1).

The number of potential CTL epitopes was further restricted by employing a PMPC, to predict whether a peptide C terminus is likely to result from proteasomal cleavage. This new method for the prediction of proteasome cleavage was developed using the SRILM statistical language model toolkit (18) and will be described in detail.² Accuracy of the model is ~90% as determined by using a set of 932 natural MHC-I-restricted peptides.

Synthetic Peptides—A panel of 148 peptides derived from influenza PR8 virus was synthesized at New England Peptide, Inc., Fitchburg, MA, by solid phase strategy using Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry on a Gilson AMS 422 synthesizer (Middleton, WI). HPLC analysis showed that the purity of the synthesized crude peptides was 86–92%. All peptides had expected masses as confirmed by mass spectrometry. The subscript numbers of each peptide indicate the amino acid positions starting from the N terminus of the corresponding PR8 protein. Crude peptides were used for initial peptide binding assays. Selected crude peptides were purified to >96% by reverse phase HPLC for further experiments in this study.

MHCI-Peptide Binding and Dissociation Assays—Binding of individual synthetic peptides from PR8 virus to mouse H-2 D^b and K^b molecules was evaluated by measuring stabilization of expression of MHC-I molecules on the surface of the TAP-deficient mutant cell line RMA-S using flow cytometric analysis with anti-D^b (clone HB-27) and anti-K^b (clone B24-8-3) mAb (19). SD₅₀ (peptide concentration that stabilizes 50% of the maximal number of D^b or K^b molecules) was calculated (Abelbeck Software) from a non-linear regression of the data to the hyperbolic equation, $f = f_{\max} (P)/(SD_{50} + (P))$, where f is the mean fluorescence acquired at different concentrations of a peptide, (P) is the corresponding peptide concentrations tested, and f_{\max} is the maximal fluorescence induced by a peptide.

The dissociation rate between peptides and MHC-I molecules was determined using 1 μg/ml of GolgiPlug (containing monensin; BD Biosciences) to block the transportation of newly synthesized MHC-I molecules from endoplasmic reticulum (ER) to the surface of RMA-S cells (20). A half-life measure ($t_{1/2}$, the time required for 50% of the molecules to decay) was used to compare the stability of different peptide/MHC (pMHC) complexes.

Infection of Mice with Influenza Virus—The working stock of influenza A/PR8/8/34 virus (PR8) used in this study was propagated and produced from a PR8 seed virus kindly provided by Dr. David Woodland, The Trudeau Institute, Saranac Lake, NY. The virus stock was titered either directly by hemagglutinin assay to obtain its hemagglutinin unit (HAU) or in 10-day old embryonated chicken eggs to determine its egg infectious dose (EID₅₀).

Female C57BL/6 mice were purchased from Taconic (Albany, NY) and housed under specific pathogen-free conditions at the animal core facility of Dana-Farber Cancer Institute prior to infection with influ-

enza virus at 6–10 weeks of age. Mice were infected by inoculating 3000 EID₅₀ of PR8 viral particles intranasally under anesthesia.

Intracellular IFN γ Staining—Intracellular IFN γ staining was performed using the Cytotfix/Cytoperm kit (BD Biosciences). Briefly, cells from the bronchoalveolar lavage (BAL) were pooled from 12–15 mice on day 10 after intranasal inoculation with PR8 virus. The cells were then cultured for 6 h in the presence of 10 μg/ml PR8-derived synthetic peptides and 1 μg/ml GolgiPlug. After culture, the responder cells were washed and stained with rat anti-CD8 phycoerythrin conjugate, followed by intracellular staining with rat anti-mouse IFN γ fluorescein isothiocyanate conjugate (BD Biosciences). Stained cells were acquired on a BD Biosciences FACScan flow cytometer, and the data were analyzed using CellQuest software (BD Biosciences). The results are expressed as the percentage of CD8⁺IFN γ ⁺ cells among total CD8⁺ T cells. BAL cells recovered from B6 mice infected with the virus for 10 days were stained with phycoerythrin-conjugated NP₃₆₆₋₃₇₄/D^b and PA₂₂₄₋₂₃₃/D^b tetrameric reagents (The Trudeau Institute, Saranac Lake, NY) and Cy-Chrome-conjugated anti-mouse CD8 α mAb (BD Biosciences).

Single Cell ELISPOT Assay—Single cell IFN γ -ELISPOT assay was performed as described previously (21) except enriched CD8⁺ T cells from the lung of B6 mice infected with influenza PR8 virus for 10 days were used as effectors. The results were expressed as the number of antigen-specific spot-forming cells (SFC) per 1 × 10⁴ CD8⁺ T cells. Assay resolution was determined by plating serially diluted PB1₇₀₃₋₇₁₁-specific CD8⁺ CTL (see below) into the anti-mouse IFN γ mAb-coated ELISPOT plates. Upon activation with the corresponding peptide, a linear relationship between the numbers of CD8⁺ T cells plated and the numbers of IFN γ -producing spots was observed, with a 1/50,000 cells responsive frequency (data not shown).

Generation of Short Term CD8⁺ T Cell Lines—Mice were injected subcutaneously with 50 μg of selected PR8-derived synthetic peptides emulsified in 100 μl of complete Freund's Adjuvant at the base of the tail. 10 to 15 days following peptide immunization, 1.5 × 10⁷ splenocytes were cultured in T25 flasks in 10 ml of complete RPMI 1640 medium containing 0.5–1 μg/ml of the corresponding peptide. Cell cultures were fed with 10 units/ml of recombinant human IL-2 (rIL-2; BD Biosciences) on day 4. After an initial 7-day culture, viable cells were selected by centrifugation through Lympholyte-M (Cedarlane, Hornby, Ontario, Canada). 1.0 to 1.5 × 10⁶ responder cells were then restimulated weekly with 1.5 × 10⁷ irradiated (3000 rads) splenocytes from naïve B6 mice in T25 flasks in the presence of 1 μg/ml of peptide and 10 units/ml of rIL-2. Lympholyte-enriched viable CD8⁺ T cells were assayed for cytotoxic activity in a ⁵¹Cr release assay on day 5 after three to five rounds of restimulation *in vitro*.

⁵¹Cr Release Assay—To generate PR8-infected target cells, 2 × 10⁶ mouse EL-4 lymphoma cells were washed and resuspended in 400 μl of serum-free RPMI medium. The cells were then incubated with 10 HAU (standard dose) of PR8 viral particles for 1 h at 37 °C. Alternatively, the same number of EL-4 cells was infected with 200 HAU (high dose) of the viral particles as indicated. Virally infected cells were transferred to 6-well plates containing 6 ml of complete RPMI medium per well and incubated overnight. To generate peptide-pulsed target cells, 1–2 × 10⁶ EL-4 cells were incubated with 20 μg/ml of individual influenza peptides in 500 μl of complete RPMI medium for 1 h at 37 °C. Both PR8-infected and peptide-pulsed EL-4 target cells were washed twice and then labeled with 150 μl of ⁵¹Cr (Na⁵¹CrO₄; PerkinElmer Life Sciences) for 90 min at 37 °C. Unpulsed but ⁵¹Cr-labeled EL-4 cells were used as control target cells. After washing three times, 1 × 10⁴ target cells were incubated with titrated concentrations of CD8⁺ T effectors in a final volume of 200 μl. 100 μl of supernatants were removed after 5 h incubation for γ radiation counting.

Immunoprecipitation—RMA-S cells and EL-4 cells were either pulsed with PR8 virus-derived synthetic peptides or infected with a high dose of live PR8 viral particles as described above. Untreated RMA-S cells and EL-4 cells were prepared as controls. The cells were then lysed on ice for 10 min with immunoprecipitation assay buffer (20 mM Tris, pH 8.0, 1 mM EDTA, 100 mM NaCl) containing 1.5% CHAPS, 5 μg/ml leupeptin, and 25 μg/ml trypsin inhibitor (all from Sigma). Peptide-bound D^b molecules were immunoprecipitated by anti-D^b mAb (clone HB-27)-coated GammaBindPlus-Sepharose (Amersham Biosciences), washed, covered with 50 μl of Tris buffer, and stored at –80 °C before analysis by mass spectrometry.

Nanospray Tandem Mass Spectrometry on a Quadrupole Time of Flight Spectrometer—The nanospray MS/MS used for identification of PR8 virus-derived peptides from the virally infected EL-4 cells in the present study is described in detail in the Supplemental Material. In brief, the MS spectrum of a typical peptide extract is characterized by a set of

² P. A. Reche, J.-P. Glutting, and E. L. Reinherz, manuscript in preparation.

TABLE I
D^b- and K^b-restricted CTL epitopes predicted from influenza PR8 virus

RNA segment	Protein product			No. of potential MHC I binders ^a		No. of potentially cleaved binders ^b	
	Name	Size	ACN ^c	D ^b	K ^b	D ^b	K ^b
		<i>aa</i>					
1	PB2	759	P03428	8	7	6	6
2	PB1	757	P03431	13	13	7	11
3	PA	716	P03433	9	12	8	6
4	HA	566	P03452	10	12	9	11
5	NP	498	P03466	5	10	3	5
6	NA	454	P03468	10	7	8	4
7	M1	252	P03485	5	4	2	3
	M2	97	P06821	1	2	1	1
8	NS1	230	P03496	4	2	1	1
	NS2	121	P03508	3	2	1	3
	Hypothetical protein (HP)	167	QQIVE1	5	4	2	2
Totals				73	75	48	53

^a SYFPEITHI was used to predict the binding possibility of influenza PR8 virus-derived 8-, 9-, and 10-mer peptides to D^b and K^b molecules.

^b Potential D^b and K^b binders predicted by SYFPEITHI were further screened by PMPC for their possible proteolytic cleavage by eukaryotic proteasomes.

^c SWISSPROT database accession number (ACN) of the viral proteins.

peaks spaced by m/z corresponding to doubly charged ions with mass deficits characteristic of peptides. An arbitrary m/z window (of unit resolution) contains multiple parent ions, and an MS/MS spectrum of the selected m/z window contains a complex mixture of fragments. Prior to the analysis of the extract, the MS/MS spectrum of the isolated target peptide (a synthetic standard) is taken under defined (and optimized) collision conditions, and this (reference) spectrum identifies the m/z windows that contain the target peptide's fragments and their relative intensity.

This statistical measure considers the m/z -dependent arrival of the fragment ions as a Poisson process and formally calculates the probability that the MS/MS spectrum of the mixture would support N parent ions that fragment with the m/z distribution contained in the reference MS/MS spectrum. In brief, the reference MS/MS spectrum defines relative arrival rates for the fragmentation of the target peptide whereas the experimental MS/MS spectrum records the net ion arrivals (counts) for the measurement period (*e.g.* 20 min). The m/z values that have positive intensity in the reference spectrum are interpreted in the experimental spectrum as a sum of fragments from the target peptide and chemical background. The probability $P(N)$ that the experimental spectrum supports at least N total events, distributed by the arrival rates of the reference spectrum, is calculated by determining the expected number of each fragment event, $n_{m/z}$, given N total events. If the measured counts in the fragment m/z window are greater than the expected $n_{m/z}$, the associated probability is set to 1 (the unexpectedly large counts are because of chemical background) whereas if the counts are less than $n_{m/z}$, the probability that this is a statistical fluctuation of the Poisson process is calculated. The product of these probabilities over all the fragment m/z values generates the decreasing function $P(N)$.

The absolute magnitude of $P(N)$ depends on experimental details such as collection period, signal intensity, and chemical background. To offset these variations the reference pattern is translated in m/z space (denoted by x) before calculating $P(N)$. The function $P(N, x)$ is then a surface that is displayed as a contour plot, and the signature of detecting the target peptide becomes a prominent peak at zero translation, $P(N, 0)$. A dominant peak at 0 translation means the MS/MS data of the experiment support the intensity ratios of the reference spectrum primarily at the reference m/z values. The Poisson measure is robust in the presence of background overlaps with some of the target peptide's MS/MS peaks; the failure to measure enough counts in any of the expected fragment windows rapidly decreases the probability of detection.

The MS/MS spectra of the synthetic peptides were measured under collision conditions optimized for each peptide, creating a library of reference MS/MS spectra. The pMHC complexes were isolated by immunoprecipitation from EL-4 total cell lysates as described above, and peptides were extracted by mild acid treatment of the antibody-coated beads. The peptide mixture was desalted by reverse phase and loaded into a nanospray tip, and the MS/MS spectra of the m/z windows corresponding to the reference spectra were measured under the pre-established collision conditions. The Poisson measure of the reference spectrum in the mixture MS/MS spectrum is then calculated as outlined above.

RESULTS

Computational Prediction of All Potential H-2 D^b- and K^b-restricted CTL Epitopes Encoded by the Influenza PR8 Virus Genome—Two important biological processes determine whether a peptide segment within a native protein can be naturally presented by a given MHC I molecule: 1) proteolytic processing within the cell required to liberate the peptide (22); and 2) binding of the excised peptide to the MHC I molecule with sufficient affinity to create a stable epitope as a cell surface pMHC for CTL recognition (23). We used a computer-based algorithm, SYFPEITHI (16), to predict all potential D^b and K^b binding epitopes from the PR8 genome of influenza A virus. Influenza A virus is a negatively stranded RNA virus (24) whose genome consists of eight RNA segments that encode 11 protein products (Table I). To identify potential D^b and K^b binders encoded by the viral genome, the amino acid sequences of each of the eleven viral proteins were searched by SYFPEITHI for all possible 8- and 9- or 10-mer peptide fragments with binding potential for K^b and D^b molecules, respectively. Only peptides with an arbitrary binding score of >50% of the maximal binding score were selected as "hits," resulting in a total of 148 peptides, including 73 potential D^b and 75 potential K^b binders (see Table I and Supplemental Table S1). These peptides comprise 1.6% of the total potential numbers of influenza A virus-encoded peptides. The predictions were further verified by RANKPEP, our recently developed computer algorithm (25).

The proteasome is the most critical protease system for generating the C-terminal end of a MHC I-restricted CTL epitope (for review see Refs. 26 and 27). To assess whether the predicted D^b and K^b binders would be processed by proteasomes, we developed a computer-based algorithm, named PMPC, to predict cleavage. As shown in Table I, 101 of 148 PR8-derived peptides showed high likelihood to be cleaved by eukaryotic proteasomes. Thus, by introducing a proteasomal cleavage motif-based algorithm, we reduced the potential numbers of CTL epitopes by 32%, from 148 to 101.

Assuming these sequence algorithms are good predictors of CTL epitopes, the known immunogenic D^b- and K^b-restricted CTL epitopes of the PR8 virus should be identified among the predicted potential CTL epitopes. As shown in Supplemental Table S1, with a single exception (K^b-restricted, subdominant 9-mer peptide PB1₇₀₃₋₇₁₁), all four other known strong H-2^b MHC I binders representing immunogenic CTL epitopes of the PR8 virus in H-2^b mice were anticipated. This list includes the

TABLE II
Experimentally identified strong D^b- and K^b binders and the stability of those MHCI-peptide complexes

Peptide ^a	Sequence/MHC restriction	Proteasomal cleavage ^b	SD ₅₀ ^c	t _{1/2} ^d	T cell epitope ^e
			nM	h	
NP ₅₄₋₆₂	RLIQNSLTI/D ^b	+	12.1	5.59	-
NP₃₆₆₋₃₇₄	ASNENMETM/D ^b	+	15.2	5.17	+
PB1 ₁₄₁₋₁₄₉	TALANTIEVF/D ^b	+	22.4	3.34	-
NA ₂₃₋₃₁	LQIGNIIS/D ^b	-	36.9	5.24	-
NS1 ₁₂₈₋₁₃₆	ILKANFSVI/D ^b	-	43.8	3.87	-
HA ₃₃₂₋₃₄₀	TGLRNIPSI/D ^b	-	54.9	2.56	-
PA₂₂₄₋₂₃₃	SSLENFRAYV/D ^b	+	72.2	4.49	+
PB1 ₁₄₀₋₁₄₈	TALANTIEV/D ^b	-	98.5	6.73	+
NA ₄₅₋₅₃	TGICNQNI/D ^b	+	198.1	8.14	-
HA ₄₆₈₋₄₇₆	SQLKNNAKEI/D ^b	+	306.1	3.14	+
NA ₄₂₅₋₄₃₂	SSISFCGV/K ^b	+	13.4	10.75	-
HA ₃₀₄₋₃₁₁	SSLPFQNI/K ^b	+	39.7	4.85	-
PB1₇₀₃₋₇₁₁	SSYRRPVGI/K ^b	Not considered	47.4	5.91	+
PB1 ₆₅₂₋₆₅₉	KNMEYDAV/K ^b	+	113.3	3.57	-
PA ₃₄₇₋₆₅₄	SLYASPQL/K ^b	+	144.9	3.60	-
M1₁₂₈₋₁₃₅	MGLIYNRM/K ^b	-	188.3	3.53	+
PB2 ₂₂₇₋₂₃₄	VYIEVLHL/K ^b	+	210.7	5.05	-
HA ₄₀₂₋₄₀₉	MNIQFTAV/K ^b	+	167.7	5.95	-
PB1 ₄₄₂₋₄₄₉	SSDDFALI/K ^b	+	170.5	3.75	-
PB1 ₄₆₄₋₄₇₁	RFYRTCKL/K ^b	+	212.4	4.72	+
NS1₁₃₃₋₁₄₀	FSVIFDRL/K ^b	+	312.8	2.98	+
NP ₃₅₋₄₂	IGRFYIQM/K ^b	-	358.3	3.64	-

^a Peptides in bold indicate the known immunogenic CD8⁺ T cell epitopes.

^b Proteasomal cleavage possibility predicted by PMPC algorithm. +, cleavage predicted; -, cleavage not predicted; Not considered, not among the top 1.6% of the potential MHC binders when examined by either SYFPEITHI or RANKPEP algorithm, because no K^b 9-mers were assessed.

^c Peptide concentration that stabilize a half-maximal number of MHCI molecules, as determined by RMA-S assay. A peptide was arbitrarily considered as a strong binder if its SD₅₀ was lower than 500 nM in RMA-S assay. One representative result is shown from two independent experiments with similar results.

^d The time required for 50% of the MHCI molecules to be lost from the surface of RMA-S cells. t_{1/2} for D^b and K^b molecules in the absence of peptides was 1.50 and 2.10 h, respectively. One representative experiment out of two is shown.

^e Identified by flow cytometry intracellular IFN γ assay and/or ELISPOT.

two identified immunodominant CTL epitopes (NP₃₆₆₋₃₇₄ and PA₂₂₄₋₂₃₃) (28, 29), among the top 1.6% of the scoring peptides, and correctly predicts that they and NS1₁₃₃₋₁₄₀ would be proteolytically cleaved. However, M1₁₂₈₋₁₃₅ is not identified by this proteasome prediction algorithm.

MHCI Binding Affinities of the Predicted D^b- and K^b-associated Epitopes—The 148 candidate epitopes were synthesized, and their binding affinity to D^b and K^b molecules was determined by RMA-S assay. The SD₅₀ (peptide concentration that stabilizes 50% of the maximal number of D^b or K^b molecules) was calculated based upon the binding curve of individual peptides. In this way, it was possible to quantitatively assess D^b and K^b binding affinity. Overall, 22 (10 D^b- and 12 K^b-restricted) peptides bound with high affinity (SD₅₀ < 500 nM) whereas 92 peptides (42 D^b- and 50 K^b-restricted) bound with intermediate affinity (500 nM < SD₅₀ < 15 μ M). The remaining 34 peptides showed minimal (SD₅₀ > 15 μ M) or no detectable MHCI binding, even at the maximal peptide concentration tested (100 μ M). In summary, 114 of 148 (77%) SYFPEITHI-predicted potential CTL epitopes from PR8 virus bound to murine D^b and K^b MHCI molecules with high or intermediate affinity.

As expected from previous studies (30–32), five known immunogenic CTL epitopes of PR8 virus in the H-2^b mouse could be categorized as strong binders when tested by RMA-S assay (Table II). Most notably, however, 17 additional strong MHCI binders were identified in the present study, 10 of which have not been reported previously. Together, the data demonstrate that matrix-based algorithms can serve as powerful computational tools to predict with reasonable accuracy potential MHCI binding epitopes from viral proteins.

Identification of Immunogenic PR8 Virus-derived CD8 T Cell Epitopes by Intracellular IFN γ Staining Assay—Previous studies (28, 30–32) have identified six CTL epitopes in the influenza virus-infected H-2^b mouse including the five strong binders noted in Table II and an intermediate strength K^b binder,

NS2₁₁₄₋₁₂₁ (32). Thus, minimally six peptide epitopes are processed and presented by professional APCs during influenza A virus infection. To gain a genome-wide perspective on the natural CTL epitope repertoire in H-2^b mice directed against the PR8 virus, a high throughput intracellular IFN γ staining flow cytometry assay was used. The recognition of each of the predicted 148 potential CTL epitopes by CD8⁺ T cells was assessed from B6 mice 10 days after intranasal inoculation with PR8 virus. BAL cells harvested directly *ex vivo* from the lung of infected animals were used as the CD8⁺ T cell source, because lung represents the prevalent anatomic site for CD8⁺ T cell-mediated immunity during the acute phase of a primary influenza A virus infection (33). Furthermore, this time interval following a primary viral challenge represents the peak of the virus-specific CD8⁺ T cell response in the murine model of influenza A virus infection (34). As shown in Fig. 1A, comparative experiments revealed that the numbers of IFN γ -secreting NP₃₆₆₋₃₇₄- and PA₂₂₄₋₂₃₃-specific CD8⁺ T cells correlated well with those obtained by immunostaining with the corresponding MHCI tetramers. Hence, intracellular IFN γ staining is a reliable way to measure the frequency of viral-specific CD8⁺ T cells from freshly isolated materials. Representative results obtained by intracellular IFN γ assay are also shown in Fig. 1B for NP₃₆₆₋₃₇₄, PA₂₂₄₋₂₃₃, PB1₇₀₃₋₇₁₁, NS2₁₁₄₋₁₂₁, and M1₁₂₈₋₁₃₅. Note that the control background value of IFN γ staining CD8⁺ T cells in the absence of peptides (no peptide) was typically <1%.

As expected, CD8⁺ T cells recovered by BAL from PR8-infected animals responded dominantly to stimulation with NP₃₆₆₋₃₇₄ and PA₂₂₄₋₂₃₃ peptides (see Fig. 1A and Table III). The frequency of CD8⁺ effector T cells specific for each of these two CTL epitopes consistently accounted for about 13% of the total CD8⁺ T cells recruited into the lung. The other most frequent CD8⁺ population specificity was directed toward a K^b-restricted epitope, PB1₇₀₃₋₇₁₁. Originally reported as a subdominant CTL epitope of PR8 virus in B6 mice (31), the fre-

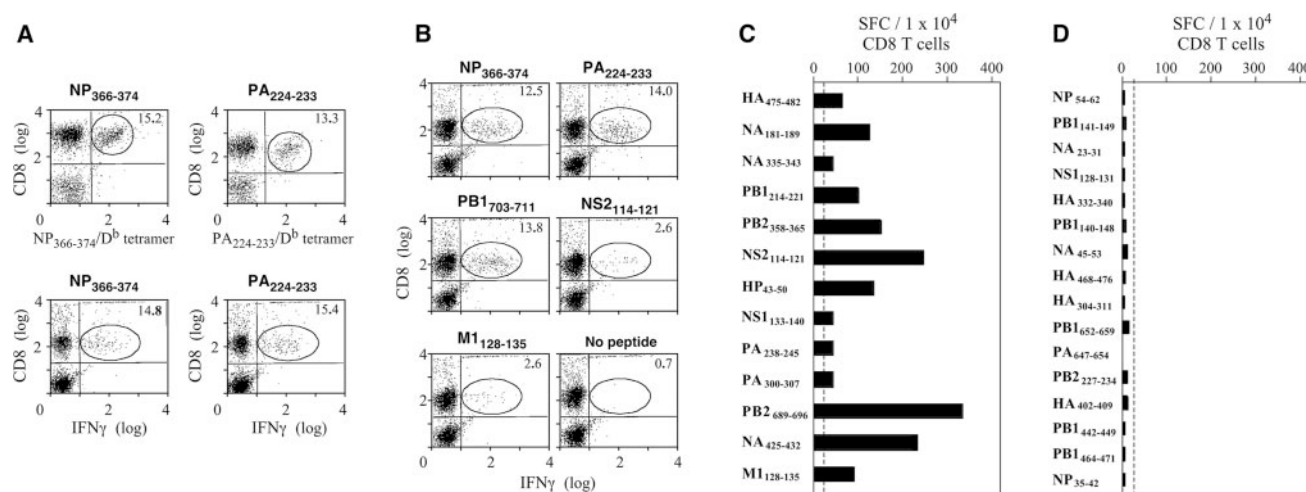


FIG. 1. Genome-wide identification of immunogenic CTL epitopes of the influenza PR8 virus in H-2^b mice. B6 mice were inoculated intranasally with PR8 virus. 10 days after infection, BAL cells were collected from the infected animals and examined either by specific tetramer and intracellular IFN γ staining (A and B) or single cell ELISPOT (C and D). A, comparison of BAL cells stained with either NP₃₆₆₋₃₇₄/D^b or PA₂₂₄₋₂₃₃/D^b tetramer versus intracellular IFN γ staining. The numbers in the upper right quadrants of the fluorescence-activated cell sorter dot plots indicate the percentage of epitope-specific or IFN γ ⁺ CD8⁺ T cells among the total CD8⁺ T cells. B, CD8 and intracellular IFN γ staining of T cells recovered directly *ex vivo* from lung airways of virus-infected B6 mice. Shown is single cell IFN γ -ELISPOT assay of PR8 virus-derived peptides that were either immunogenic in the intracellular IFN γ staining assay (C) or “non-immunogenic” but manifest strong MHC binding in RMA-S assay (D). CD8⁺ T cells enriched from the lung of B6 mice infected with PR8 virus for 10 days were incubated with irradiated splenocytes in the presence of 10 μ g/ml of indicated peptides and 10 units/ml of rIL-2 for 24 h. IFN γ -positive dots in each well were detected by developing the plates with biotinylated anti-mouse IFN γ mAb followed by streptavidin-horseradish peroxidase and substrate. The results were expressed as the number of antigen-specific SFC/1 $\times 10^4$ CD8⁺ T cells. A count was considered positive if it was three times higher than that of control wells (CD8⁺ T cells and APCs without peptides), as indicated by the dotted lines.

TABLE III
Identification of murine H-2 D^b- and K^b-restricted immunogenic CTL epitopes of influenza PR8 virus by intracellular IFN γ staining

Peptide ^a	Sequence/MHC restriction	MHC binding ^b	Proteasomal cleavage ^c	% of responding CD8 ⁺ T cells ^d		
				Exp. 1	Exp. 2	Exp. 3
NP ₃₆₆₋₃₇₄	ASNENMETM/D ^b	+++	+	14.8	12.5	10.9
PA ₂₂₄₋₂₃₃	SSLENFRAYV/D ^b	+++	+	13.2	14.0	7.3
HA ₄₇₅₋₄₈₂	KEIGNGCFEF/D ^b	++	+	0.5	3.2	1.5
NA ₁₈₁₋₁₈₉	SGPDNGAVAV/D ^b	++	–	1.7	2.8	3.1
NA ₃₃₅₋₃₄₃	YRYGNGVWI/D ^b	++	+	1.3	1.7	2.9
PB1 ₇₀₃₋₇₁₁	SSYRRPVG/K ^b	+++	Not considered	5.5	13.8	10.0
PB1 ₂₁₄₋₂₂₁	RSYLIRAL/K ^b	++	+	2.2	3.9	2.7
PB2 ₃₅₈₋₃₆₅	GYEFTMV/K ^b	++	+	2.2	3.3	1.2
NS2 ₁₁₄₋₁₂₁	RTFSFQLI/K ^b	++	+	2.1	2.6	3.9
HP ₄₃₋₅₀	GGLPFSLL/K ^b	++	+	1.5	2.5	4.0
NS1 ₁₃₃₋₁₄₀	FSVIFDRL/K ^b	+++	+	2.1	0.7	3.9
PA ₂₃₈₋₂₄₅	NGYIEGKL/K ^b	++	+	1.4	2.9	1.5
PA ₃₀₀₋₃₀₇	GIPLYDAI/K ^b	++	+	1.7	2.9	1.1
PB2 ₆₈₉₋₆₉₆	VLRGFLIL/K ^b	++	+	1.1	1.2	3.2
NA ₄₂₅₋₄₃₂	SSISFCGV/K ^b	+++	+	0.9	1.0	2.5
M1 ₁₂₈₋₁₃₅	MGLIYNRM/K ^b	+++	–	0.0	2.6	2.1

^a Peptides in bold indicate known immunogenic CD8⁺ T cell epitopes.

^b D^b and K^b binding affinity determined by RMA-S assay. +++, strong binder; ++, intermediate binder.

^c Proteasomal cleavage prediction as described in Table II.

^d Percentage of IFN γ -secreting CD8⁺ T cells among total CD8⁺ T cells, as determined by intracellular IFN γ staining of CD8⁺ T cells from BAL of influenza PR8 virus-infected B6 mice on day 10 after infection. A total of 148 synthetic peptides were examined. A staining was considered positive, when the percentage of CD8⁺ IFN γ ⁺ was three times higher than that of the background staining (without peptides). Only peptides that resulted in positive results in more than one of three independent experiments are listed. The background staining were 0.02, 0.72, and 0.15%, respectively.

quency of CD8⁺ T cells responding to this epitope was variable in this study, ranging from 5.5 to 13.8% in three independent measurements. In addition, three known minor epitopes (NS2₁₁₄₋₁₂₁, NS1₁₃₃₋₁₄₀, and M1₁₂₈₋₁₃₅) were detected in the intracellular IFN γ assay. Most importantly, 10 previously unrecognized CTL epitopes were identified in this manner. The numbers of IFN γ -secreting CD8⁺ T cells in response to stimulation with these peptides were above background values in virtually all independent experiments conducted (Table III). In sum, CD8⁺ T cells recovered by BAL during the peak of a primary influenza virus infection responded to 16 of the 148 predicted CTL epitopes, with NP₃₆₆₋₃₇₄, PA₂₂₄₋₂₃₃, and PB1₇₀₃₋₇₁₁ specificities being most prominent.

Identification of Subdominant CD8 Epitopes within the PR8 Virus by Single Cell ELISPOT Assay—Although efficient for screening large numbers of antigenic peptides, intracellular IFN γ staining assay has limited detection sensitivity (typically ~1%). For more precise quantitation, we used a single cell ELISPOT assay that readily detects a responsive cell frequency of 0.002%. The numbers of IFN γ -producing spots induced by the 13 peptides tested were significantly above background (3-fold or higher) with the NS1₁₃₃₋₁₄₀ responsive T cells being the lowest (~50 SFC per 10⁴) and PB2₆₈₉₋₆₉₆ being the highest (320 SFC/10⁴) (Fig. 1C). These data establish the clear immunogenicity of these epitopes in H-2^b mice infected with the influenza A virus PR8. In contrast, 16 other PR8 virus-derived

strong D^b and K^b binders did not induce IFN γ above background, demonstrating their lack of immunogenicity in the context of native viral proteins (Fig. 1D). Thus, 16 of the 22 strong MHCI binding peptides failed to elicit *in vivo* T cell responses.

Stability of D^b- and K^b-Peptide Complexes—MHC-peptide complex stability has been reported to contribute to the immunogenicity of human immunodeficiency virus- and Epstein-Barr virus-derived CTL epitopes (20, 35). Therefore, we next compared the stability of the MHCI-PR8 peptide complexes. As shown in Table II, no direct correlation between stability of the MHC-peptide complexes and their immunogenicity *in vivo* is evident. Immunodominant NP_{366–374} and PA_{224–233}, for example, do not form the most stable D^b-peptide complexes (5.17 and 4.49 h, respectively). Conversely, the non-immunogenic NA_{45–53} formed the most stable D^b-peptide complex, with a half-life of 8.14 h. Peptides NP_{54–62}, NA_{23–31}, PB1_{140–148}, PB2_{227–234}, and HA_{402–409}, although non-immunogenic *in vivo*, formed MHCI-peptide complexes at least equally stable as those formed with the three immunodominant CTL epitopes. Although the half-lives of MHCI-peptide complexes formed with four D^b-restricted peptides (PB1_{141–149}, NS1_{128–136}, HA_{332–340}, and HA_{468–476}) and six K^b-restricted peptides (PB1_{652–659}, PA_{647–654}, M1_{128–135}, PB1_{442–449}, NS1_{133–140}, and NP_{35–42}) were particularly short (an average $t_{1/2}$ being only about 3 h), it is unknown whether this contributes to their lack of immunogenicity *in vivo*. Nonetheless, neither strength of peptide binding to MHCI nor pMHCI complex half-life alone predicts immunogenicity.

Ability of Strong MHC Binders to Induce CTL Responses *in Vivo*—If the TCR repertoire of B6 mice lacks specificities capable of recognizing certain PR8-derived peptides, this may explain the non-immunogenicity of a fraction of the H-2^b binding viral products. Consequently, we investigated the ability of strong MHCI binders to elicit a CTL response *in vivo*. Peptides chosen for this analysis were those non-immunogenic after infection with PR8 virus, even with the highly sensitive ELISPOT assay (see Table II and Fig. 1D). Mice were immunized with the selected strong D^b and K^b binding peptides. Cytotoxic activity of splenic CD8⁺ T cells from immunized animals was then measured against peptide-pulsed target cells. Initial evaluation showed that 5 of 12 effector populations tested were able to mediate weak lysis of the corresponding peptide-pulsed targets but not an irrelevant PR8 virus-derived peptide-pulsed target (data not shown). The 12 populations were then stimulated repeatedly *in vitro* with their respective peptides to expand the antigen-specific CD8⁺ T cells. Over 90% of the enriched effectors were CD8⁺ T cells after three-four rounds of restimulation (data not shown). All 12 short term CTL lines showed strong cytotoxic activity against peptide-pulsed targets (Supplemental Table S2). These data clearly indicate that CD8⁺ T cell precursors capable of responding to otherwise non-immunogenic potential CTL epitopes of influenza PR8 virus exist in the naive CD8⁺ T cell repertoire of B6 mice.

Relative Avidity of CTL Lines to Peptide-pulsed Targets—Although the above results exclude complete absence of T cells in B6 mice directed at the non-immunogenic/high affinity H-2^b MHCI binders, it is possible that the avidity of those TCRs for their respective pMHCI ligands is too low to trigger activation and differentiation of CD8⁺ T cells *in vivo*. To investigate the notion of a “functional” rather than absolute T cell deficit, the short term CTL lines described above were used to lyse targets coated with various amounts of each respective peptide. The viral peptide concentrations necessary to lyse 50% of target cells (EC₅₀) were used to probe the relative avidity of each of the TCRs from different CTLs to their corresponding MHCI-

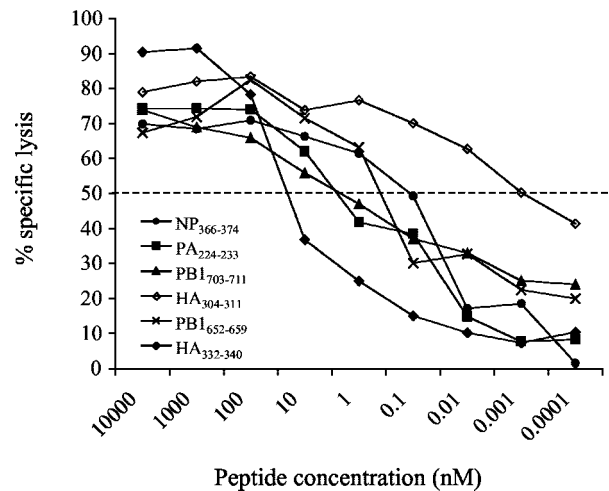


FIG. 2. **Relative avidity of TCR-based immune recognition by antigen-specific CTL lines.** PR8 virus-specific CTL lines were generated as described under “Experimental Procedures” and used to lyse EL-4 targets pulsed with titrated amounts of the indicated peptides. Peptide concentrations required to achieve 50% of the maximal lysis of the targets (EC₅₀) were used to compare the relative avidity of each functional set of pMHCI-TCR interactions. One of two experiments is shown.

peptide ligands. Variation in EC₅₀ values between individual experiments was less than 25% (data not shown).

As shown in Fig. 2, the EC₅₀ for the three CTL lines directed at immunodominant peptides, NP_{366–374}, PA_{224–233}, and PB1_{703–711}, were 0.04×10^{-9} , 0.9×10^{-9} , and 0.9×10^{-9} M, respectively. Surprisingly, the TCR of the CTL line specific for the “non-immunogenic” HA_{304–311} showed much greater avidity (100–1,000-fold) to the corresponding MHCI-peptide ligand, with a 0.0005×10^{-9} M EC₅₀. Moreover, a second non-immunogenic peptide, PB1_{652–659}, stimulated its respective CTL line in the picomolar EC₅₀ range comparable with that of the three immunodominant CTL epitopes. Conversely, the CTL line specific for the non-immunogenic HA_{332–340} showed the weakest avidity to its corresponding D^b-peptide-pulsed target with a 9.0×10^{-9} M EC₅₀. The data clearly show that a relatively high avidity TCR-pMHCI ligand requirement *per se* is not the dominant factor determining immunogenicity of a potential CTL epitope nor is an absolute or relative functional absence of TCR specificity the basis.

Processing and Presentation of PR8 Virus CTL Epitopes *in Vitro*—Failure to detect an epitope-specific CD8⁺ T cell population also may be attributed to absent or inefficient natural processing, resulting in lack of sufficient pMHCI surface complexes to induce a CTL response *in vivo*. We used virus-specific CTL lines to probe the availability of otherwise non-immunogenic, strong D^b and K^b binders of PR8 virus on the surface of EL-4 cells following *in vitro* viral infection. A major advantage of this *in vitro* system is the ability to control the parameters influencing the copy numbers of the virus-derived peptides on the surface of the infected target cells (22, 36, 37).

We first tested CTL-mediated lysis of EL-4 targets infected with a standard dose of the PR8 virus *in vitro* ($10 \text{ HAU}/2 \times 10^6$ cells). Approximately 25% of EL-4 cells were infected by the virus under this condition, as judged by immunostaining of the HA viral glycoprotein with anti-HA mAb (kindly provided by Dr. Jonathan W. Yewdell, NIAID, National Institute of Health; data not shown). As shown in Fig. 3, lytic effect was readily observed when NP_{366–374}, NS2_{114–121}, or PB1_{703–711}-specific CTL lines were used as effectors. In contrast, the CTL line specific for the PA_{224–233} epitope was only weakly lytic for the virus-infected target, although such CTL manifest strong cyto-

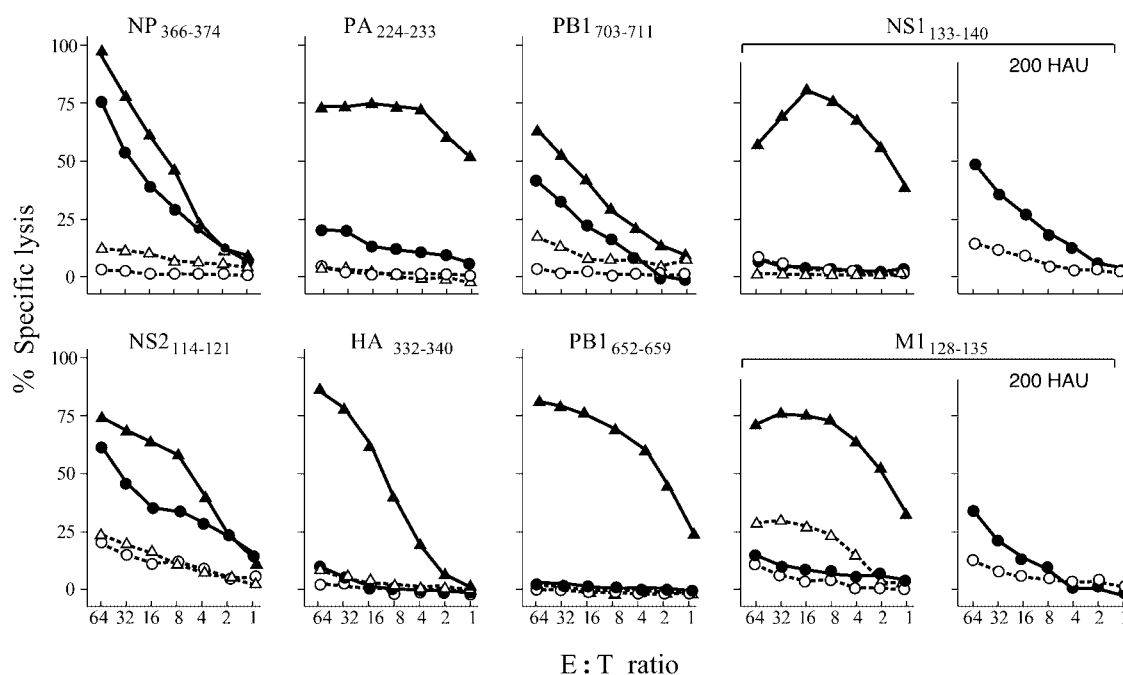


FIG. 3. **Endogenously processed and presented CTL epitopes of influenza PR8 virus.** PR8 virus-specific epitopes that are processed and presented on the surface of EL-4 targets after infection with a standard dose of the viral particles ($10 \text{ HAU}/2 \times 10^6$ cells) were detected by ^{51}Cr release assay using the corresponding CTL lines as effectors. $\text{NS1}_{133-140}$ - and $\text{M1}_{122-135}$ -specific CTL lines were further tested against EL-4 targets infected with a high dose of the viral particles ($200 \text{ HAU}/2 \times 10^6$ cells), as shown in the labeled panels on the right. The symbols used are as follows: filled circles, PR8 virus-infected EL-4 targets; open circles, EL-4 targets without PR8 infection; filled triangles, peptide-pulsed EL-4 targets; open triangles, EL-4 targets without peptide pulsing.

toxic activity against $\text{PA}_{224-233}$ peptide-pulsed targets. Given that synthetic $\text{PA}_{224-233}$ peptide possesses a high D^b binding affinity and can form a relatively stable complex with D^b molecules (Table II), the observed weak lysis suggests that this viral protein may be insufficiently processed and/or presented by EL-4 cell line *in vitro* (see below). CTL lines specific for another two known subdominant CTL epitopes ($\text{NS1}_{133-140}$, $\text{M1}_{128-135}$) as well as those generated to the eight non-immunogenic strong binders, were capable of lysing the corresponding peptide-pulsed targets with strong lytic activity, but their lytic effect on the virus-infected targets were minimal (see Fig. 3 and Supplemental Table S2).

We then increased the infecting *in vitro* viral dose to $200 \text{ HAU}/2 \times 10^6$ cells. This measured the frequency of virally infected EL-4 targets to 75%, as judged by expression of the HA glycoprotein (data not shown). Importantly, when $\text{NS1}_{133-140}$ - and $\text{M1}_{128-135}$ -specific CTL lines were tested against the heavily infected EL-4 cells, a detectable and specific cytotoxicity was observed (Fig. 3). In contrast, $\text{HA}_{332-340}$ - and $\text{PB1}_{652-659}$ -specific CTL lines showed no significant cytolysis against EL-4 cells infected at either viral multiplicity of infection (see Fig. 3 and Supplemental Table S2). These data strongly suggest that immunogenic peptides are naturally processed during viral infection but that the degree of viral infection may influence processivity, thereby determining lysis susceptibility. On the other hand, non-immunogenic peptides may be processed too inefficiently to target infected cells for lysis but allow CTL to lyse cells when pulsed with exogenous peptide.

Detection of a Limited Number of Presented Peptides by Nanospray MS/MS—The finding that the $\text{PA}_{224-233}$ epitope is immunodominant *in vivo* (see Fig. 1B and Table III) but results in weak lysis for virally infected targets (Fig. 3) suggests that this epitope may be insufficiently processed and/or presented by EL-4 cells. This may be true for other peptide epitopes of the PR8 virus, as well. A direct measure of the peptide density displayed on the surface of infected EL-4 cells by a physical method would address this issue. Therefore, an approach based

on nanospray MS/MS of unfractionated or partially fractionated mixtures using a quadrupole time of flight spectrometer was developed. Compared with conventional capLC (38), nanospray of appropriately cleaned peptide mixtures will produce higher counts per mole because of the lower flow rate ($<20 \text{ nl}/\text{min}$) and the ability to optimize the static spray.

The MS/MS spectra of the synthetic peptides were measured under collision conditions optimized for each peptide, creating a library of reference MS/MS spectra. The peptide mixtures purified from either peptide-pulsed or virally infected EL-4 cells were loaded into a nanospray tip, and the MS/MS spectra of the m/z windows corresponding to the reference spectra were measured under the pre-established collision conditions. A statistical measure, *e.g.* Poisson process, is used to identify the signature of the reference spectrum in the spectrum of the mixture (see Supplemental Figs. S1–S11). Such a measure is especially robust in the presence of background overlaps with some of the target peptide's MS/MS peaks. Fig. 4 offers the results of such analysis. For example, $\text{PA}_{224-233}$ is readily detected at one copy per cell using a mixture of uninfected EL-4 cells and $\text{PA}_{224-233}$ pulsed RMA-S cells as detailed in the Supplemental Material (Fig. 4A). In uninfected EL-4 cells, no $\text{PA}_{224-233}$ peptide is detected (Fig. 4B) whereas in PR8-infected EL-4 cells, the Poisson statistics reveal its presence although at a level of <1 copy/cell (compare panel C with panel A). This low level of processing explains why the $\text{PA}_{224-233}$ -specific CTL weakly kills PR8-infected EL-4 cells but readily lyses the $\text{PA}_{224-233}$ peptide-pulsed targets. The copy number of the latter is orders of magnitude greater than present in the infected sample (see Supplemental Material). In contrast, whereas $\text{HA}_{332-340}$ is also detectable by mix-in analysis at $10 \text{ HA}_{332-340}$ copies per cell (Fig. 4D), this peptide is not observed in the virally infected EL-4 cells (Fig. 4F) and yields a pattern indistinguishable from uninfected EL-4 cells (Fig. 4E). In addition $\text{PB1}_{140-148}$ could also be detected at 10 copies per cell in the mix-in analysis but not in the virally infected EL-4 cells (Supplemental Figs. S9–S11).

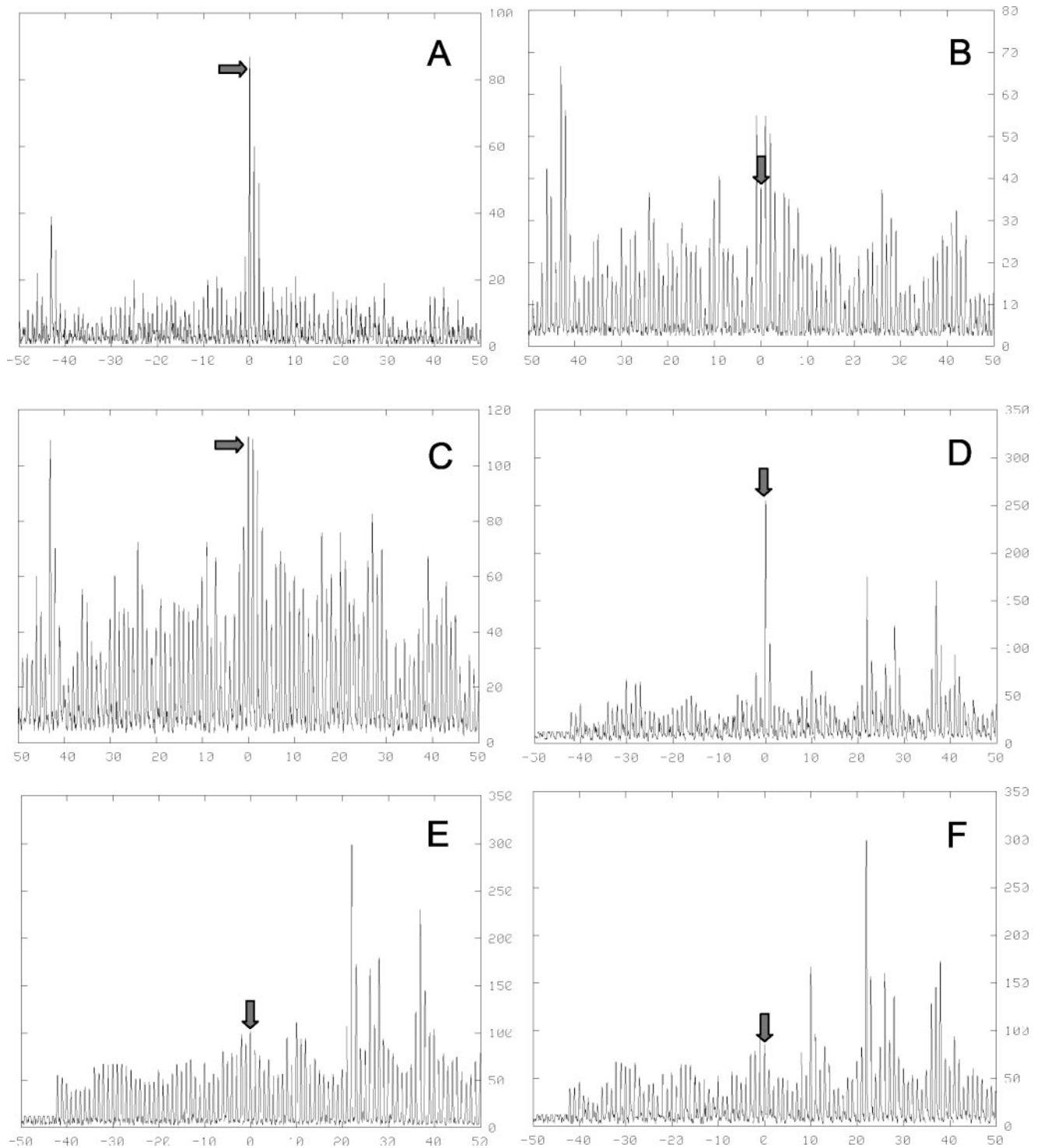


FIG. 4. Detection of viral peptide by nanospray MS/MS. *A*, PA_{224–233} detected in a mixture of peptide-loaded RMA-S and non-loaded EL-4 cells at a total concentration of one PA_{224–233} copy/cell. *B*, no PA₂₂₄ signal in uninfected EL-4 cells. *C*, PA_{224–233} signal associated with infected EL-4 cells. *D*, HA_{332–340} detected in a mixture of RMA-S and EL-4 cells at a total concentration of 10 HA_{332–340} copies/cell as described. *E*, no HA_{332–340} signal in uninfected EL-4 cells. *F*, no HA_{332–340} signal in infected EL-4 cells. The plots show a single contour of a surface in which probability is plotted as a function of m/z shift (x) and event number (y) (see Supplemental Material).

DISCUSSION

MHCI-restricted antigenic peptides are derived from cytosolic precursor proteins. After intracellular protein degradation, mainly by proteasomes, the resulting peptides are translocated to ER by TAP (transporters associated with antigen processing), where they bind to newly synthesized empty MHC I molecules in a MHC allele-specific manner (39). pMHC complexes are transported to the cell surface of professional APC for

recognition by CD8⁺ CTLs via their TCRs. Proteasome cleavage, TAP transport, and MHC binding represent three key processing events in generating immunologically relevant CTL epitopes (27). Within the ER, peptides may be further trimmed by ERAAP, an aminopeptidase associated with antigen processing (40).

We used computer-based algorithms to rapidly search for all potential CTL epitopes from the genome of influenza PR8 vi-

rus. SYFPEITHI identified 148 peptide sequences with potential mouse D^b and K^b binding capacities among the eleven proteins encoded by the virus. Five of six known PR8 CTL epitopes in H-2^b mice were included among the top 1.6% of potential CTL epitopes. The prediction results could be further verified by the RANKPEP algorithm; ~80% of the predicted MHCI binders were identical in both algorithms. When experimentally tested for their MHCI binding abilities by RMA-S assay, 114 of 148 peptides (77%) were able to bind to mouse D^b and K^b molecules, although with a substantial range of binding affinity. We infer that the 148 potential D^b and K^b binders originally predicted by SYFPEITHI algorithm may contain the majority of the natural CTL epitopes of the virus.

The SYFPEITHI algorithm predicts potential MHCI binders using a motif matrix (16). For pathogens with large genomes, such as smallpox virus (~200 open reading frames) and *Chlamydia trachomatis* (894 protein-coding genes), the list of candidate peptides will be too long to be pursued practically. Therefore, additional selection criteria are needed for CTL epitope prediction. Comprehensive analysis of proteasome-digested peptide products from unmodified model proteins has generated abundant experimental data on cleavage specificities of proteasomes (41, 42), making it possible to extract proteasomal cleavage motifs (*e.g.* preferred sequence patterns around cleavage sites) for automated prediction (43). We developed a method based on a probabilist model to predict proteasome cleavage (PMPC), bypassing tedious experimental digestion and biochemical analyses. When the 11 PR8 virus-coding proteins were run on the proteasome program, 101 potential MHC binders (67%) revealed a high processing probability (Table I). Given this large number of proteins, experimental data cannot be obtained to assess the predictive accuracy of this algorithm. However, that 13 of 16 immunogenic CTL epitopes identified could be anticipated among the peptide list of potential proteasomal cleavage (Table III) suggests that our PMPC algorithm may have substantial predictive power. The combination of MHCI binding algorithms with proteasomal cleavage programs represent a powerful search engine for robust, automated prediction of CTL epitopes from infectious genomes.

We first used a high throughput intracellular IFN γ assay to screen the 148 potential T cell epitopes for CD8⁺ CTLs recovered directly *ex vivo* from the lung of PR8 virus-infected B6 mice. Lung represents the inflammatory site harboring the highest frequency of virus-specific CD8⁺ T cells during an acute infection with influenza virus (33). As expected, three known immunogenic CTL epitopes, *e.g.* NP₃₆₆₋₃₇₄, PA₂₂₄₋₂₃₃, and PB1₇₀₃₋₇₁₁, were readily detected by this assay (see Fig. 1 and Table III). Most importantly, 13 other PR8 virus-derived peptides were found to be immunogenic in virally infected B6 mice, 10 of which have not been identified previously (Table III). The immunogenicity of these minor CTL epitopes *in vivo* was further confirmed by single cell ELISPOT assay (Fig. 1C) and collectively establishes that over 70% of CD8⁺ T cells recruited into the lungs of the mice during a primary influenza A infection are virus-specific. The remaining 30% of the CD8⁺ T cells may reflect bystander activation with specificity for other antigens or include a subset of CD8⁺ T cells with specificity for influenza A epitopes distinct from those predicted. Nonetheless, the data suggest that combined use of SYFPEITHI and PMPC predicts the majority of the naturally existing CTL epitopes of the influenza PR8 virus.

These 16 CTL epitopes are directed at 10 of the 11 PR8 viral proteins, products of both early and late genes of the virus (44). The only viral gene that fails to encode demonstrable epitopes is the M2 matrix protein, the smallest of the viral proteins,

predicted to harbor only one possible K^b and one D^b binding peptide. Thus, in contrast to previous observations (28, 29, 31) showing that CD8⁺ T cells target primarily internal proteins of influenza A virus in H-2^b mice, the data presented here clearly establish that T cell recognition includes coverage of the vast majority of viral products.

Three of the 16 D^b- and K^b-restricted peptide epitopes (NP₃₆₆₋₃₇₄, PA₂₂₄₋₂₃₃, and PB1₇₀₃₋₇₁₁) collectively account for the majority of CD8⁺ antigen-specific CTLs from B6 mice infected with PR8 virus (Table III). The precise molecular basis of this immunodominance is not yet clear. Presentation may be a key factor in determining immunogenicity of an epitope and involves the availability of peptides for assembly into pMHC in the ER and their MHCI binding affinities (22, 37). In the present study, we found that 101 of 148 potential CTL epitopes could theoretically be processed by proteasomes. However, the majority were intermediate or weak MHCI binders, as determined experimentally by RMA-S assay. Therefore, it is most likely that these peptides may not productively bind to MHCI molecules in the ER, especially in an environment that places them in competition with strong MHCI binders. The "subdominance" of the 10 newly identified CTL epitopes may be explained by their intermediate MHCI binding affinity (Table III).

Seventeen PR8 virus-related peptides were able to bind to MHCI molecules at least as strongly as the three immunodominant CTL epitopes (Table II), yet they are non-immunogenic (Fig. 1D). Analysis of off-rates revealed that 11 of 17 strong MHC binders dissociated very rapidly from the MHCI molecules (Table II). Furthermore, failure to detect cytolysis of heavily infected target cells by the two non-immunogenic, strong K^b binders tested, *e.g.* PA₆₄₇₋₆₅₄ and PB1₆₅₂₋₆₅₉, implies either a T cell repertoire limitation (see below) or that these peptides are not presented or in too low numbers to induce a detectable CD8⁺ T cell response *in vivo*, as suggested by ELISPOT analysis (Fig. 1D). This latter assumption is further supported by the absence of HA₃₃₂₋₃₄₀ peptide epitopes on the surface of EL-4 cells infected with PR8 *in vitro* using physical measurements (Fig. 4). Because the *in vitro* system may not reflect the physiologic conditions, it is crucial to assess antigen processing and presentation *in vivo*. The nanospray MS/MS method herein can detect as low as one copy number of peptide per cell using 2×10^7 cells, an ~100-fold greater sensitivity (Fig. 4) compared with the conventional HPLC-MS analysis requirement. This new MS methodology should allow us to dissect antigen processing and presentation in various APC compartments from *ex vivo* infectious materials in the future.

The repertoire of responding CTL is another important parameter determining epitope immunogenicity. The observation that 10 non-immunogenic, strong D^b and K^b binders were able to induce epitope-specific CTL *in vivo* when given as peptide immunogens in complete Freund's Adjuvant (Supplemental Table S2) suggests that relevant CTL precursors are present in the pre-immune CD8⁺ T cell repertoire of B6 mice. The lack of responsiveness to the strong D^b and K^b binders does not appear to be because of low avidity between responding TCRs and pMHC ligands, as peptide titration data reveal that the avidities of the respective CTL lines are essentially similar to those of immunodominant CTLs (Fig. 3). Hence, there seems to be no obvious functional defect (*i.e.* holes) in the responding CD8⁺ CTL repertoire for the silent strong D^b and K^b binders of the PR8 virus. The absence of CD8 T cell responses to those strong binders is probably linked to poor antigen processing during influenza A infection *in vivo*.

Although anti-viral CD8⁺ T cells may recognize only several immunodominant CTL epitopes, our data show that the num-

ber of endogenously processed and presented epitopes is larger than thought previously. The strength of immune response to these minor CTL epitopes appears to be regulated by several factors including MHC binding affinity, stability of pMHCI, and the level of epitope presentation. The phenomenon of immunodominance cannot be explained by any single variable. Because multiple viral proteins can serve as potential vaccine candidates for induction of virus-specific CTL, it is tempting to speculate that a CD8⁺ T cell vaccine consisting of multiple minor viral CTL epitopes may induce broad immunity *in vivo*. This may be of particular importance in situations where generation of viral variants during infection represents a major obstacle for immunodominant CTL epitope-mediated viral clearance.

REFERENCES

- Wang, J. H., and Reinherz, E. L. (2002) *Mol. Immunol.* **38**, 1039–1049
- Wilson, I. A., and Garcia, K. C. (1997) *Curr. Opin. Struct. Biol.* **7**, 839–848
- Zinkernagel, R. M. (1996) *Science* **271**, 173–178
- Doherty, P. C., Allan, W., Eichelberger, M., and Carding, S. R. (1992) *Annu. Rev. Immunol.* **10**, 123–151
- Schaible, U. E., Collins, H. L., and Kaufmann, S. H. (1999) *Adv. Immunol.* **71**, 267–377
- Yewdell, J. W., and Bennink, J. R. (1999) *Annu. Rev. Immunol.* **17**, 51–88
- Schulz, M., Zinkernagel, R. M., and Hengartner, H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 991–993
- Kast, W. M., Roux, L., Curren, J., Blom, H. J., Voordouw, A. C., Meloen, R. H., Kolakofsky, D., and Melief, C. J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2283–2287
- Sastry, K. J., Bender, B. S., Bell, W., Small, P. A., Jr., and Arlinghaus, R. B. (1994) *Vaccine* **12**, 1281–1287
- Moskophidis, D., and Zinkernagel, R. M. (1995) *J. Virol.* **69**, 2187–2193
- Price, G. E., Ou, R., Jiang, H., Huang, L., and Moskophidis, D. (2000) *J. Exp. Med.* **191**, 1853–1867
- Cole, G. A., Hogg, T. L., Coppola, M. A., and Woodland, D. L. (1997) *J. Immunol.* **158**, 4301–4309
- Oukka, M., Riche, N., and Kosmatopoulos, K. (1994) *J. Immunol.* **152**, 4843–4851
- Oukka, M., Manuguerra, J. C., Livaditis, N., Tourdot, S., Riche, N., Vergnon, I., Cordopatis, P., and Kosmatopoulos, K. (1996) *J. Immunol.* **157**, 3039–3045
- Fu, T. M., Friedman, A., Ulmer, J. B., Liu, M. A., and Donnelly, J. J. (1997) *J. Virol.* **71**, 2715–2721
- Rammensee, H., Bachmann, J., Emmerich, N. P., Bachor, O. A., and Stevanovic, S. (1999) *Immunogenetics* **50**, 213–219
- Rammensee, H. G., Friede, T., and Stevanovic, S. (1995) *Immunogenetics* **41**, 178–228
- Stolcke, A. (2002) in *Proceedings of the International Conference on Spoken Language Processing* (Ohala, J. J., Nearey, T. M., Derwing, B. L., Hodge, M. M., and Wiebe, G. E., eds) Vol. 2, pp. 901–904, Center for Spoken Language Research, Boulder, CO
- Preckel, T., Grimm, R., Martin, S., and Weltzien, H. U. (1997) *J. Exp. Med.* **185**, 1803–1813
- van der Burg, S. H., Visseren, M. J., Brandt, R. M., Kast, W. M., and Melief, C. J. (1996) *J. Immunol.* **156**, 3308–3314
- Zhong, W., Roberts, A. D., and Woodland, D. L. (2001) *J. Immunol.* **167**, 1379–1386
- Sijts, A. J., Villanueva, M. S., and Pamer, E. G. (1996) *J. Immunol.* **156**, 1497–1503
- Sette, A., Vitiello, A., Reheman, B., Fowler, P., Nayarsina, R., Kast, W. M., Melief, C. J., Oseroff, C., Yuan, L., Ruppert, J., et al. (1994) *J. Immunol.* **153**, 5586–5592
- Lamb, R. A., and Krug, R. M. (1996) in *Fields Virology* (Fields, B. N., Knipe, D. M., and Howley, P. M., eds) Vol. 2, pp. 1353–1395, Lippincott-Raven Publishers, Philadelphia, PA
- Reche, P. A., Glutting, J. P., and Reinherz, E. L. (2002) *Hum. Immunol.* **63**, 701–709
- Monaco, J. J. (1992) *Immunol. Today* **13**, 173–179
- Shastri, N., Schwab, S., and Serwold, T. (2002) *Annu. Rev. Immunol.* **20**, 463–493
- Townsend, A. R., Rothbard, J., Gotch, F. M., Bahadur, G., Wraith, D., and McMichael, A. J. (1986) *Cell* **44**, 959–968
- Belz, G. T., Stevenson, P. G., and Doherty, P. C. (2000) *J. Immunol.* **165**, 2404–2409
- Belz, G. T., Xie, W., Altman, J. D., and Doherty, P. C. (2000) *J. Virol.* **74**, 3486–3493
- Belz, G. T., Xie, W., and Doherty, P. C. (2001) *J. Immunol.* **166**, 4627–4633
- Vitiello, A., Yuan, L., Chesnut, R. W., Sidney, J., Southwood, S., Farness, P., Jackson, M. R., Peterson, P. A., and Sette, A. (1996) *J. Immunol.* **157**, 5555–5562
- Hogan, R. J., Usherwood, E. J., Zhong, W., Roberts, A. A., Dutton, R. W., Harmsen, A. G., and Woodland, D. L. (2001) *J. Immunol.* **166**, 1813–1822
- Flynn, K. J., Belz, G. T., Altman, J. D., Ahmed, R., Woodland, D. L., and Doherty, P. C. (1998) *Immunity* **8**, 683–691
- Busch, D. H., and Pamer, E. G. (1998) *J. Immunol.* **160**, 4441–4448
- Restifo, N. P., Bacik, I., Irvine, K. R., Yewdell, J. W., McCabe, B. J., Anderson, R. W., Eisenlohr, L. C., Rosenberg, S. A., and Bennink, J. R. (1995) *J. Immunol.* **154**, 4414–4422
- Vijh, S., Pilip, I. M., and Pamer, E. G. (1998) *J. Immunol.* **160**, 3971–3977
- Henderson, R. A., Michel, H., Sakaguchi, K., Shabanowitz, J., Appella, E., Hunt, D. F., and Engelhard, V. H. (1992) *Science* **255**, 1264–1266
- Falk, K., Rotzschke, O., Stevanovic, S., Jung, G., and Rammensee, H. G. (1991) *Nature* **351**, 290–296
- Serwold, T., Gonzalez, F., Kim, J., Jacob, R., and Shastri, N. (2002) *Nature* **419**, 480–483
- Emmerich, N. P., Nussbaum, A. K., Stevanovic, S., Priemer, M., Toes, R. E., Rammensee, H. G., and Schild, H. (2000) *J. Biol. Chem.* **275**, 21140–21148
- Toes, R. E., Nussbaum, A. K., Degermann, S., Schirle, M., Emmerich, N. P., Kraft, M., Laplace, C., Zwinderman, A., Dick, T. P., Muller, J., Schonfisch, B., Schmid, C., Fehling, H. J., Stevanovic, S., Rammensee, H. G., and Schild, H. (2001) *J. Exp. Med.* **194**, 1–12
- Kuon, W., Holzthutter, H. G., Appel, H., Grolms, M., Kollnberger, S., Traeder, A., Henklein, P., Weiss, E., Thiel, A., Lauster, R., Bowness, P., Radbruch, A., Kloetzel, P. M., and Sieper, J. (2001) *J. Immunol.* **167**, 4738–4746
- Shapiro, G. I., Gurney, T., Jr., and Krug, R. M. (1987) *J. Virol.* **61**, 764–773