nature immunology

Article

https://doi.org/10.1038/s41590-023-01644-5

Cryptic MHC-E epitope from influenza elicits a potent cytolytic T cell response

Received: 8 November 2022

Accepted: 8 September 2023

Published online: 12 October 2023

Check for updates

Michael J. Hogan[1] ✉, Nikita Maheshwari[1,2,9], Bridget E. Begg[3], Annalisa Nicastri[4], Emma J. Hedgepeth[1], Hiromi Muramatsu[5], Norbert Pardi[5], Michael A. Miller[6,10], Shanelle P. Reilly[7], Laurent Brossay[7], Kristen W. Lynch[3], Nicola Ternette[4] & Laurence C. Eisenlohr[1,8] ✉

The extent to which unconventional forms of antigen presentation drive T cell immunity is unknown. By convention, CD8 T cells recognize viral peptides, or epitopes, in association with classical major histocompatibility complex (MHC) class I, or MHC-Ia, but immune surveillance can, in some cases, be directed against peptides presented by nonclassical MHC-Ib, in particular the MHC-E proteins (Qa-1 in mice and HLA-E in humans); however, the overall importance of nonclassical responses in antiviral immunity remains unclear. Similarly uncertain is the importance of 'cryptic' viral epitopes, defined as those undetectable by conventional mapping techniques. Here we used an immunopeptidomic approach to search for unconventional epitopes that drive T cell responses in mice infected with influenza virus A/Puerto Rico/8/1934. We identified a nine amino acid epitope, termed M-SL9, that drives a co-immunodominant, cytolytic CD8 T cell response that is unconventional in two major ways: first, it is presented by Qa-1, and second, it has a cryptic origin, mapping to an unannotated alternative reading frame product of the influenza matrix gene segment. Presentation and immunogenicity of M-SL9 are dependent on the second AUG codon of the positive sense matrix RNA segment, suggesting translation initiation by leaky ribosomal scanning. During influenza virus A/Puerto Rico/8/1934 infection, M-SL9-specific T cells exhibit a low level of egress from the lungs and strong differentiation into tissue-resident memory cells. Importantly, we show that M-SL9/Qa-1-specific T cells can be strongly induced by messenger RNA vaccination and that they can mediate antigen-specific cytolysis in vivo. Our results demonstrate that noncanonical translation products can account for an important fraction of the T cell repertoire and add to a growing body of evidence that MHC-E-restricted T cells could have substantial therapeutic value.

CD8 T cells conventionally recognize pathogen-derived epitopes presented by the hyperpolymorphic major histocompatibility complex (MHC)-Ia molecules, including HLA-A, HLA-B and HLA-C in humans and H2-D, H2-K and H2-L in mice. However, protection from some viral[1–6] or bacterial[7] diseases has also been demonstrated by CD8 T cells recognizing epitopes presented by the MHC-Ib proteins, which exhibit limited polymorphism and are typically associated with more specialized functions[8]. An important subgroup of the MHC-Ib family are the MHC-E

A full list of affiliations appears at the end of the paper. ✉e-mail: hoganm@chop.edu; eisenlc@pennmedicine.upenn.edu

proteins. These are best known for downregulating the cytolytic activity of natural killer (NK) cells by presenting an MHC-Ia-derived self peptide (called the 'Qa-1 determinant modifier' or Qdm in mice) to inhibitory NKG2/CD94 receptors[9,10]; thus, NK cells are able to specifically kill cells when the MHC-Ia presentation machinery is antagonized or lost due to mutation. Recently, it has become apparent that MHC-E can also stimulate robust CD8 T cell responses in multiple species against various infections and cancer[11–13]. Most notably, the Picker laboratory elicited MHC-E-restricted CD8 T cells in rhesus macaques using a cytomegalovirus (CMV) vector encoding antigens from simian immunodeficiency virus (SIV), and, remarkably, these T cells were uniquely capable of clearing a pathogenic SIV infection[1,2,4]. A protective capacity of MHC-E-restricted CD8 T cells has also been observed in a handful of other infectious disease settings[5–7]. Despite its striking therapeutic potential, research into the biology of MHC-E-restricted CD8 T cells has been minimal and is hampered by a lack of tractable small animal models, since, to our knowledge, no virally derived Qa-1-restricted epitope sequence has been published so far.

Another problem in T cell biology is understanding the full diversity of MHC-binding peptides, collectively known as the immunopeptidome. Efforts to map T cell responses to viruses are typically limited to peptides from annotated viral proteins[14–16]. However, many reports have shown that T cell epitopes can originate from cryptic sources, including polypeptides created by noncanonical translation[17–26] and posttranslational peptide splicing[27–30]. While recent advancements in mass spectrometry have allowed for deeper probing of cryptic epitopes, uncertainties remain regarding the fraction of the immunopeptidome that is derived from noncanonical sources and their clinical relevance[31]. As most, if not all, cryptic epitopes described so far have been MHC-Ia restricted, an additional unknown is whether unconventional epitopes can be presented on MHC-Ib and MHC class II (MHC-II).

In this Article, we set out to investigate the possibility that the T cell response against influenza A virus (IAV) may include cryptic epitopes presented on MHC-II. Serendipitously, we discovered an epitope derived from a cryptic viral translation product that is instead restricted to the nonclassical MHC-Ib molecule, Qa-1. Remarkably, this cryptic Qa-1 epitope drives a major fraction of the CD8 T cell response in IAV-infected C57Bl/6 mice and exhibits potent cytolytic capacity after infection or messenger (m)RNA vaccination.

Results

Discovery of a cryptic CD8 T cell response to influenza

We conducted an immunopeptidomic analysis with the initial aim of identifying cryptic MHC-II epitopes in a mouse model of IAV infection. We hypothesized this was possible given that MHC-II presentation in IAV-infected mice is dominated by endogenous epitopes directly presented by infected antigen-presenting cells (APCs)[32]. As outlined in Fig. 1a, we infected an in-house APC line called B6-CIITA-E^d (C57Bl/6 (B6) fibroblasts expressing MHC-II molecules I-A^b and I-E^d)[33] with influenza virus A/Puerto Rico/8/1934 (PR8), immunoprecipitated MHC-II, isolated peptide ligands, and analyzed by liquid chromatography–tandem mass spectrometry (LC–MS^2). Mass spectra were interrogated against a database containing the PR8 genome fully translated in all six reading frames (RFs), as well as the mouse proteome. We identified 7,864 unique peptide species at an estimated false discovery rate of 3.4%, with 95.7% mapping to mouse proteins. Two-thirds of all peptides were 14–18 amino acids in length, as is typical for MHC-II ligands, and 335 unique peptide species mapped to PR8, showing a similar length distribution (Extended Data Fig. 1a,b). Among both PR8 and non-PR8 peptides, a large portion matched the consensus MHC-II-binding motifs available from NetMHCIIpan 4.0 (Extended Data Fig. 1c–f)[34]. Peptides of nine amino acids formed a small local peak in the overall length distribution (Extended Data Fig. 1a), as seen previously[35], and the consensus of these peptides resembled that of the MHC-I molecules expressed by these cells, suggesting minor co-precipitation[36] of peptide–MHC-I complexes and/or free MHC-I ligands along with MHC-II (Extended Data Fig. 1g).

Only one identified PR8 peptide mapped to a potential noncanonical ribosomal product and was named 'M-SL9' based on its origin in the matrix (M) gene segment and its amino acid sequence, SLQGRTLIL. This 9-mer was validated as an epitope by detecting reactive splenocytes from PR8-infected mice by interferon (IFN)-γ enzyme-linked immunosorbent spot (ELISpot) assay (Fig. 1a and Extended Data Fig. 2). M-SL9 is encoded within the +1 RF of the matrix protein 1 (M1)-coding sequence (Fig. 1b and Extended Data Fig. 3) and could not be attributed to any open RF (ORF) known to be translated in wild-type (WT) PR8 infection, and was therefore designated as a cryptic epitope.

Intriguingly, ELISpot analysis revealed M-SL9-specific (IFN-γ-producing) cells within bulk splenocytes but not purified spleen CD4 T cells from C57Bl/6 mice; no M-SL9 response was detectable in BALB/c spleens from a small sample of $N = 4$ mice (Extended Data Fig. 2). We further characterized the M-SL9-specific response by peptide-stimulating lung lymphocytes from PR8-infected C57Bl/6 mice and intracellular cytokine staining of IFN-γ, tumor necrosis factor (TNF) and IL-2. This clearly showed that M-SL9 peptide activated CD8 T cells but not CD4 T cells (Fig. 1c). In fact, over 10% of stimulated lung CD8 T cells were specific for M-SL9, in comparison to ~20% for the well-described H2-D^b-associated PR8 epitopes, nucleoprotein $(NP)_{366–374}$ and polymerase acidic protein $(PA)_{224–233}$ (Fig. 1d and Extended Data Fig. 4). In some mice, the M-SL9 response exceeded the $NP_{366–374}$ or $PA_{224–233}$ responses. Notably, M-SL9-specific CD8 T cells displayed a high level of cytokine polyfunctionality, with three-quarters of cells expressing ≥2 cytokines and one-quarter expressing 3 cytokines. This level of polyfunctionality was roughly double that of the $NP_{366–374}$ response and was comparable to that of $PA_{224–233}$. Cytolytic markers were analyzed similarly, and comparable frequencies of cytolytic CD8 T cells (~10%) were identified as specific for M-SL9 or $NP_{366–374}$ based on simultaneous expression of perforin, granzyme B and the degranulation marker CD107a (Fig. 1e and Extended Data Fig. 4). The effector expression profile of M-SL9-specific CD8 T cells is thus typical of a type 1-polarized antiviral immune response.

M-SL9 is presented by Qa-1

Given that a CD8 T cell epitope was identified via MHC-II immunopeptidomics, we asked which MHC molecule serves as the restricting element for M-SL9. This investigation was aided by a T cell hybridoma clone, called B6.23, which produces β-galactosidase as a sensitive and specific reporter of M-SL9 presentation (Fig. 2a). This clone was discovered fortuitously by screening PR8-specific T cell hybridoma clones previously generated in house. B6.23 recognizes M-SL9 peptide-pulsed and PR8-infected APCs (Fig. 2b) and is activated by the M-SL9 sequence present in our molecular clone of PR8 (ref. 37) as well as the M-SL9-P variant (SLQGRTPIL) (Extended Data Fig. 5) commonly present in other PR8 isolates. To test MHC restriction, we co-cultured B6.23 cells with M-SL9 peptide-pulsed APCs bearing various sets of MHC molecules. B6.23 was activated by the B6-CIITA fibroblast line but not by L929 cells ($H2^k$ haplotype) expressing the C57Bl/6-derived MHC-Ia molecules H2-D^b or H2-K^b; these lines were validated by measuring presentation of $NP_{366–374}$/H2-D^b and chicken $ovalbumin_{257–264}$ (SIINFEKL)/H2-K^b to corresponding T hybridomas (Fig. 2c). Lack of MHC-Ia binding by M-SL9 was corroborated by the observation that M-SL9 peptide was not able to stabilize empty H2-K^b or H2-D^b molecules on the surface of transporter associated with antigen processing (TAP)-deficient RMA-S cells[38] (Extended Data Fig. 6a).

Interestingly, B6.23 was activated by M-SL9 peptide-pulsed bone marrow-derived dendritic cells (BMDCs) from BALB/c mice (Fig. 2c), suggesting usage of an MHC-Ib molecule with low polymorphism. We noted that M-SL9 bears a striking resemblance to Qdm, with matching or favored amino acids at all anchor positions[39,40] (Fig. 2d); therefore, we tested whether Qa-1 was able to present M-SL9 to B6.23 cells.

https://doi.org/10.1038/s41590-023-01644-5

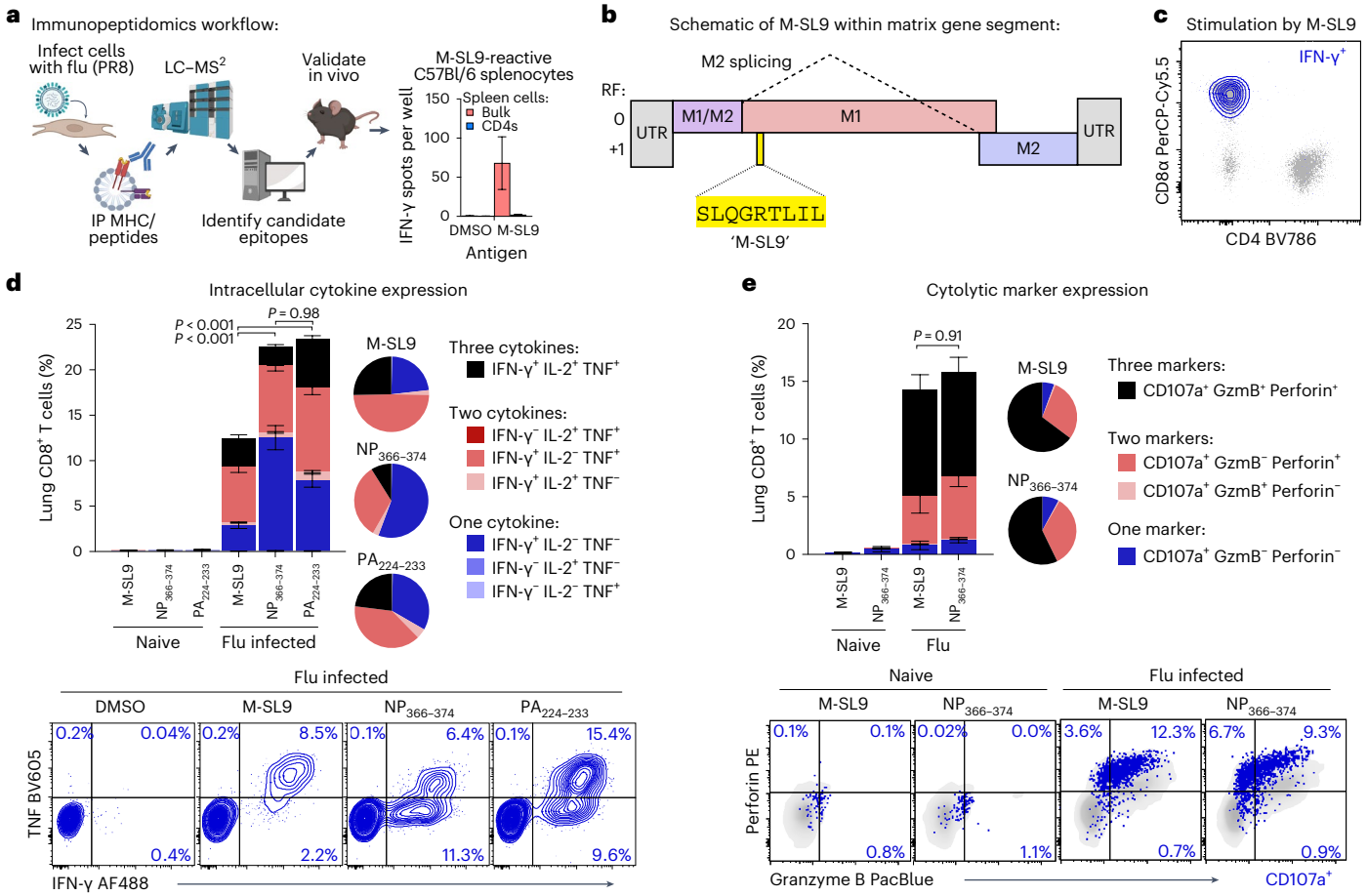


Fig. 1 | Discovery of co-immunodominant CTL response directed against M-SL9, an epitope encoded by a noncanonical IAV matrix ORF. a, Outline of immunopeptidomics and first identification of M-SL9-reactive spleen cells by ELISpot. **b**, Schematic of M-SL9 position within the PR8 flu matrix gene segment. **c**–**e**, C57Bl/6 mice were intranasally infected with 40 FFU (**c** and **d**) or 160 FFU PR8 (**e**); lung lymphocytes were stimulated with indicated peptides and intracellular cytokines (**c** and **d**) or cytolytic markers were detected within the $CD8^+$ $CD4^-$ compartment (**e**). Pie charts represent mean proportions in infected mice, and flow plots are representative. In **c** and **d**, $N = 34$ infected mice and $N = 21$ naive mice across six independent experiments. In **e**, $N = 11$ infected mice and $N = 4$ naive mice across two independent experiments. In **e**, percentages on flow plots describe the frequency of $CD107a^+$ (blue) cells in each quadrant out of the whole $CD8^+$ T cell population. Error bars are s.e.m. A one-way ANOVA was performed on all cytokine$^+$ (**d**) or triple-marker$^+$ cells (**e**). IP: immunoprecipitate, UTR: untranslated region, GzmB: granzyme B.

When HeLa cells were transduced with the *H2-T23* gene encoding Qa-1^b (expressed in C57Bl/6), they gained the ability to present synthetic M-SL9 peptide (Fig. 2e and Extended Data Fig. 6b–e). This was not true for HeLa cells expressing H2-T11, a related MHC-Ib protein originating from a duplication of *H2-T23*. Next, we tested whether Qa-1^b expression was required for C57Bl/6 cells to present M-SL9. We obtained BMDCs from WT C57Bl/6 or various mutant mouse strains, pulsed with peptides, and tested for activation of B6.23 and control hybridomas. M-SL9 presentation was high for WT BMDCs and those deficient for I-A^b (MHC-$II^{-/-}$) or both H2-K^b and H2-D^b ($K^bD^{b-/-}$) but was abrogated for BMDCs that were also deficient for Qa-1^b ($K^bD^bQa\text{-}1^{-/-}$) or for β_2 microglobulin (β_2m) (Fig. 2f). These studies revealed that Qa-1 is both necessary and sufficient for M-SL9 presentation and that the β_2m component of all MHC-I is also required, as expected. This finding was extended in vivo by measuring M-SL9/Qa-1^b tetramer-binding CD8 T cells in the lungs of PR8-infected mice. These cells were robustly detected and accounted for about 5% of all lung CD8 T cells, while $NP_{366–374}$/H2-D^b tetramer-specific CD8 T cells accounted for roughly 18% (Fig. 2g). We noted that the exact frequencies of each T cell specificity can differ on the basis of the virus dose and the detection method.

To our knowledge, M-SL9 is the first sequenced viral CD8 T cell epitope shown to be restricted to Qa-1. While we cannot rule out the possibility that M-SL9 has some affinity for MHC-II in addition to Qa-1, this seems unlikely based on weak M-SL9/MHC-II binding prediction (Supplementary Table 2) and the lack of M-SL9-specific CD4 T cells (Fig. 1c and Extended Data Fig. 2b,c). A more likely explanation is that M-SL9 peptide or M-SL9/Qa-1^b complexes co-precipitated during the immunoenrichment of MHC-II, consistent with previous findings that MHC-Ia, MHC-Ib and MHC-II ligands can all immunoprecipitate nonspecifically and can be detectable by MS[36]. This idea is also supported by the fact that the sequence motif of identified 9-mers resembles that of MHC-Ia and, to a lesser extent, Qa-1 ligands. Additionally, we note that our mouse-derived peptide identifications also included the canonical ligand for Qa-1, Qdm ($-\log_{10}P = 31.77$). Co-precipitation therefore seems to be a general property of immunopeptidomics, highlighting the need to confirm MHC restriction during such studies.

Phenotype of M-SL9-specific CD8 T cells

We next tested whether the immunophenotype and functional characteristics of M-SL9/Qa-1^b-specific T cells resemble that of classical IAV-specific CD8 T cells. We identified CD8 T cells recognizing M-SL9 or $NP_{366–374}$ from PR8-infected mouse lungs by tetramer staining. While some Qa-1-dependent CD8 T cells have been shown to be T cell receptor (TCR) γδ positive[41], M-SL9-specific CD8 T cells were

https://doi.org/10.1038/s41590-023-01644-5

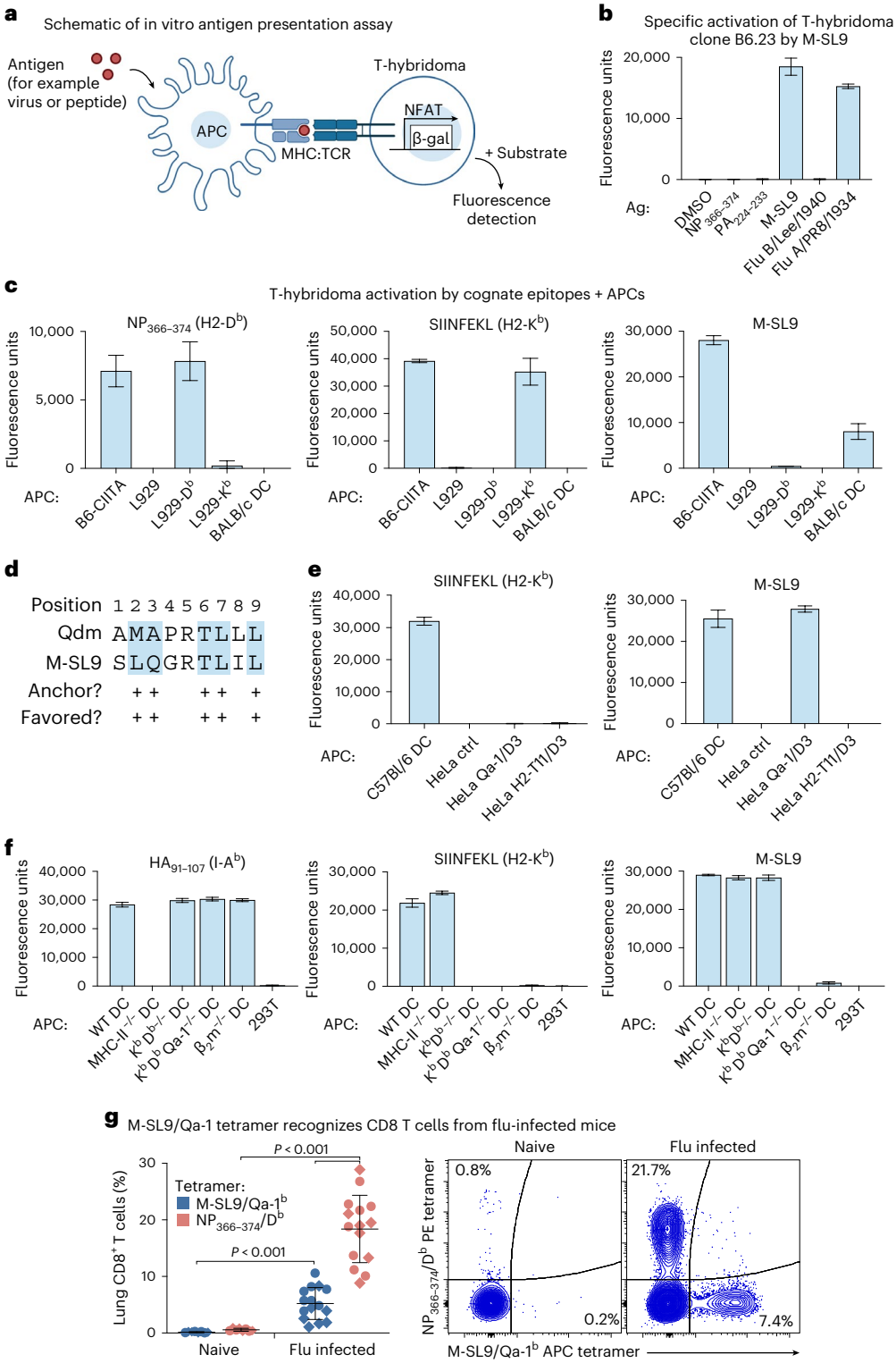


Fig. 2 | M-SL9-specific CD8 T cells are restricted to the nonclassical MHC class Ib, Qa-1. a, Diagram of antigen presentation assay using T-hybridoma. **b**, To validate M-SL9-specific T-hybridoma clone B6.23, B6-CIITA cells were incubated with indicated peptides or infected with live viruses, then co-cultured with B6.23. **c**, Various APCs, including BALB/c BMDCs ('DC') and L929 cells ($H\text{-}2^k$ origin) transduced with the genes encoding $H2\text{-}D^b$ (D^b) or $H2\text{-}K^b$ (K^b), were peptide pulsed and tested for their ability to present to indicated T-hybridomas. **d**, Alignment of M-SL9 with Qdm, showing favored amino acids at anchor positions. **e**, HeLa cells were transduced with genes encoding $Qa\text{-}1^b$ or H2-T11 containing the $H2\text{-}D^b$ α3 domain (D3) to facilitate antibody staining, and peptide presentation was tested as in **c**. **f**, C57Bl/6 BMDCs with one, two, or three knockout mutations were tested for their ability to present M-SL9 and control peptides (including the $I\text{-}A^b$-restricted $HA_{91–107}$epitope) to specific T-hybridomas, with 293T cells as a negative control. **g**, Tetramers were used to detect epitope-specific CD8 T cells in mouse lungs 9 days after infection with 40 FFU PR8 ($N = 15$) or no virus (naive, $N = 8$); diamonds are males, circles are females and flow plots are representative. In **b**–**f**, the background from DMSO-only conditions was subtracted from T-hybridoma data; error bars of technical triplicates are s.d.; results are representative of three independent experiments. *P* values reflect one-way ANOVA. β-gal: β-galactosidase, Ag: antigen.

https://doi.org/10.1038/s41590-023-01644-5

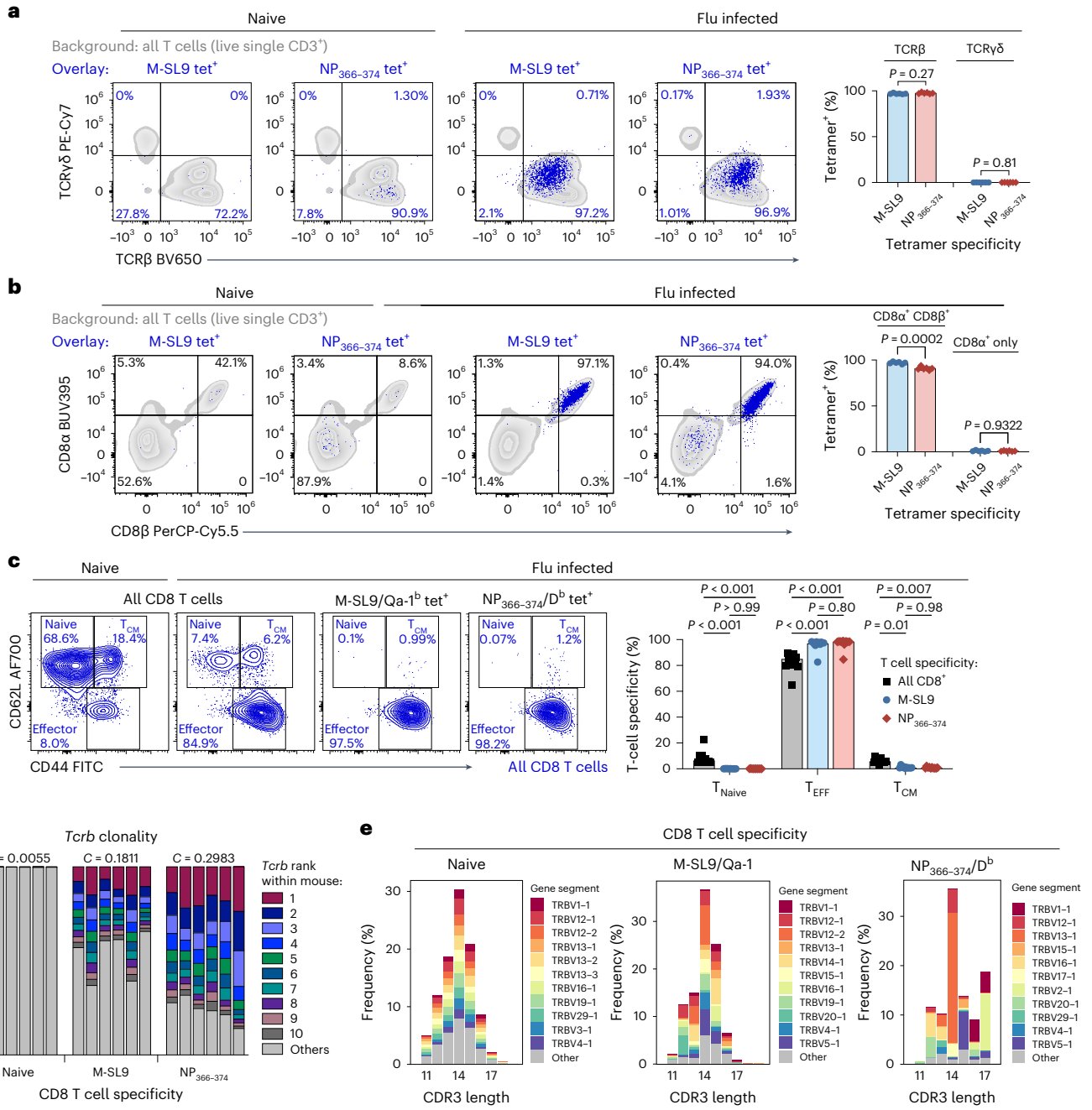


Fig. 3 | Immunophenotype and clonal diversity of M-SL9/Qa-1-specific CD8 T cells. a–**c**, Lung lymphocytes from mice 9 days after intranasal infection with 160 FFU of IAV PR8 or from naive controls were evaluated for expression of TCRβ versus TCRγδ (**a**), CD8α versus CD8β (**b**) and CD44 (marker of antigen experience) versus CD62L (marker of lymphoid homing) (**c**). Antigen specificities were revealed by tetramer staining. **d**,**e**, *Tcrb* genes from sorted naive or tetramer-specific lung CD8 T cells were sequenced and ranked by frequency within each of $N = 6$ mice (**d**). The mean Simpson clonality index (C) is shown above each group. 'Spectratype' histograms (**e**) show the average frequency of *Tcrb* V gene usage and CDR3 lengths for each T cell specificity. In **a**, $N = 10$ and in **b** and **c**, $N = 6$ across ≥2 independent experiments each. All studies used female and male mice. In **a**–**c**, the dots show individual mice and bars show the mean. P values are shown from one-way ANOVAs. tet: tetramer, EFF: effector, CM: central memory.

uniformly TCRβ$^+$ and TCRγδ$^-$, consistent with αβ T cells (Fig. 3a). Qa-1 has also been associated with regulatory CD8 T cells, which can express Qa-1-binding CD8αα homodimers[42] and even the transcription factor FoxP3 (ref. 43), normally a hallmark of regulatory CD4 T cells. Here, we found that M-SL9-specific CD8 T cells only expressed typical CD8αβ heterodimers (Fig. 3b), and all PR8-induced CD8 T cells were negative for FoxP3 (Extended Data Fig. 4e). Additionally, the M-SL9-specific population was virtually all within the CD44$^+$ CD62L$^-$ effector population, as expected of effector CD8 T cells (Fig. 3c).

Given the strong sequence homology between M-SL9 and Qdm (Fig. 2d), we asked whether M-SL9/Qa-1^b might serve as a ligand for NKG2A/CD94 or the related NK receptors NKG2C/CD94 and NKG2E/CD94, all recognized by the same antibody clone used here. We observed that NKG2A/C/E$^+$ cells from naive mouse spleens were efficiently bound by Qdm/Qa-1^b tetramer but not by M-SL9/Qa-1^b or negative control

$NP_{366–374}$/H2-D^b tetramers (Extended Data Fig. 7a). No other $CD3^-$ $CD19^-$ spleen cells were stained by M-SL9/Qa-1^b tetramer, indicating that M-SL9/Qa-1^b does not interact with any subset of NK cells.

Next, we tested the related question of whether Qdm/Qa-1^b tetramer would bind to M-SL9-specific CD8 T cells. Consistent with the literature[44], we found that PR8-induced CD8 T cells stained with Qdm/Qa-1^b tetramer regardless of antigen specificity (Extended Data Fig. 7b). This can be explained by the fact that CD8 T cells normally express NKG2A/CD94 in response to activation[45], allowing interaction with Qdm/Qa-1^b as part of a negative feedback or 'immune checkpoint' pathway[46,47]. Indeed, we found that Qdm/Qa-1^b tetramer mainly co-stained with anti-NKG2A/C/E. These data suggest that Qdm/Qa-1^b tetramer binds to PR8-induced CD8 T cells of multiple specificities in a TCR-independent manner.

Some reports of MHC-Ib-restricted CD8 T cells have shown them to have limited clonality or restricted TCR gene usage[48]. We examined this by sequencing recombined *Tcrb* genes in M-SL9/Qa-1^b and $NP_{366–374}/D^b$ tetramer-specific cells sorted from the lungs of PR8-infected mice. The top ten expanded *Tcrb* clones accounted for ~50% and ~80% of epitope-specific *Tcrb* sequences for M-SL9- and $NP_{366–374}$-specific CD8 T cells, respectively, with a similar fold difference in the Simpson clonality index, indicating that M-SL9-specific cells were moderately more clonally diverse than $NP_{366–374}$-specific cells (Fig. 3d). Both cell populations showed a variety of *Tcrb* variable (V) and joining (J) gene segment usage and recombined V–diversity (D)–J sequences, with a limited number of V genes that were particularly enriched, for example, *Tcrb* V12-2 in M-SL9-specific cells and V13-1 in $NP_{366–374}$-specific cells (Fig. 3e, Extended Data Fig. 8a,b and Supplementary Table 1). The gene usage was more similar within each T cell population across different mice than within each mouse across different T cell populations (Extended Data Fig. 8c,d). Length distributions of complementary determining region 3 (CDR3) were roughly similar with a mode of 14 amino acids in each case (Fig. 3e). Taken together, we conclude that M-SL9/Qa-1^b-specific CD8 T cells are largely similar to conventionally restricted IAV-specific CD8 T cells in cytokine and cytolytic functionality, major surface markers and clonality.

M-SL9-specific T cells home to the lung during PR8 infection

Next, we compared the frequency of CD8 T cells specific for M-SL9/Qa-1 versus the classical epitopes $NP_{366–374}/D^b$ and $PA_{224–233}/D^b$ within various tissues from 6 to 56 days after respiratory PR8 infection using tetramer staining, and we observed striking differences in both anatomical distribution and kinetics. At 6 days, M-SL9 was the dominant specificity in lungs, representing a median of 2.5% of lung CD8 T cells, while the responses to $PA_{224–233}$ and $NP_{366–374}$ were significantly lower, at 0.9% and an undetectable level, respectively (Fig. 4a and Extended Data Fig. 9a). By day 9, the $NP_{366–374}$ and $PA_{224–233}$ responses had surged to arrive at or near their peak levels of 15% and 10%, respectively, while the M-SL9 response increased more slowly and did not peak until day 14, with a median of 10%. Between days 14 and 56, all IAV-specific CD8 T cells waned in the lung until reaching a similar level of roughly 5% of CD8 T cells for each specificity.

The CD8 T cell kinetics followed a similar pattern in bronchoalveolar lavage fluid (BALF) (Fig. 4b) and the lung-draining mediastinal lymph node (MLN) (Fig. 4c), although the responses in MLN were about one order of magnitude lower than in lungs or BALF. At day 6, M-SL9 was either the dominant response (BALF) or co-dominant with $PA_{224–233}$ (MLN), while the $NP_{366–374}$ responses were negligible. At both sites, classical CD8 T cell responses reached their peak at approximately day 9, while the M-SL9 response peaked at day 14. Thereafter, the frequency of each CD8 T cell specificity remained stable or waned slightly over time, and the infiltration of all T cells into BALF dropped significantly by day 31 (Extended Data Fig. 9b).

In blood (Fig. 4d) and spleen (Fig. 4e), classical CD8 T cells peaked at day 9, when the $NP_{366–374}$ response accounted for 15% (blood) and 6% (spleen) of CD8 T cells, and the $PA_{224–233}$ response was at about half these levels. These frequencies diminished rapidly by day 14, in comparison to the slow waning of responses in the respiratory tract and draining lymph node. In marked contrast, M-SL9-specific T cells were largely absent from the blood or spleen. The M-SL9 response on days 6–9 was barely above background at either site and peaked at low frequencies of 1.5% (blood) and 0.7% (spleen) on day 14 before returning to background levels by day 31. Of note, we originally detected M-SL9-specific T cells in spleen (Fig. 1a) by IFN-γ ELISpot, a much more sensitive method of T cell detection than tetramer staining.

Induction of M-SL9-specific tissue-resident memory T cells

Given the remarkable degree of skewing of M-SL9-specific CD8 T cells toward the lungs and BALF (~10% of CD8 T cells at peak) and nearly complete absence from the blood and spleen after PR8 infection, we analyzed the expression of trafficking molecules by different CD8 T cell specificities. We measured CD103, an integrin associated with homing to the respiratory epithelium and possibly the lamina propria of mucosal sites[49,50], and CD69, an inhibitor of lymphocyte egress from the lung and other organs[51,52]. When present on lung T cells ≥30–40 days after IAV infection, CD103 and CD69 serve as markers for noncirculating tissue-resident memory T cells (T_{RM}), which are associated with protection from heterosubtypic IAV challenge[53,54]. In the lung and BALF of PR8-infected mice, CD103 and CD69 expression increased over time on all T cell specificities, but to the greatest extent on M-SL9-specific cells (Fig. 5a,b and Extended Data Fig. 9c). At day 56, T cells specific for M-SL9 showed a significantly greater level of differentiation (67%) into T_{RM} compared with $PA_{224–233}$ (54%) and $NP_{366–374}$ (24%) (Fig. 5c). This supports a model in which M-SL9-specific CD8 T cells skew preferentially to the lung due to increased expression of trafficking molecules that prevent egress from the PR8-infected lung and, eventually, enhanced differentiation into T_{RM} cells.

Recall of the M-SL9 response by heterosubtypic infection

Prior studies have detailed the disparate ways in which $NP_{366–374}$-specific and $PA_{224–233}$-specific CD8 T cells are recalled upon secondary antigen exposure in C57Bl/6 mice. Specifically, 'PR8→X31' mice that are infected with PR8, rested for 1 month, and then infected with X31 (a reassortant virus encoding H3N2 surface proteins and PR8 internal proteins) experience a profound boost in their $NP_{366–374}$ response but little to no increase in the $PA_{224–233}$ response[55]. Here, we reproduced these findings and tested whether the M-SL9 response can be efficiently recalled in the same scenario (Fig. 6a). In all examined anatomical sites, M-SL9-specific CD8 T cells were elevated in PR8→X31 mice in comparison to mice that received PR8 infection alone (Fig. 6b–f). However, PR8→X31 M-SL9 responses in the lung, BALF and MLN were not elevated in comparison to the peak response after X31 infection alone, suggesting that the additional expansion or 'boost' from the secondary infection was minimal in these sites. A trend toward a larger boost was observed in blood and spleen, where the median frequencies of M-SL9-specific CD8 T cells in PR8→X31 mice were increased at least threefold relative to either virus alone, although *P* values (up to 0.15) did not reach statistical significance (Fig. 6e,f). This is in contrast to the $NP_{366–374}$ response, which was significantly increased by two- to fivefold in PR8→X31 mice compared with mice infected with either virus alone. This was also distinct from the $PA_{224–233}$ response, which fell to a significantly lower percentage after secondary infection compared with primary infection with either virus. Interestingly, we noted that X31 infection alone was able to consistently elicit M-SL9-specific T cells in circulation; these responses were low compared with $NP_{366–374}$ and $PA_{224–233}$ but substantially greater than after PR8 alone. This suggests that properties of the antigen exposure can affect the tissue homing of nonclassical CD8 T cells. Accordingly, an alternative explanation for the slight boost in M-SL9-specific T cells in the blood and spleen of PR8→X31 mice is that secondary infection with X31 induced greater circulation of M-SL9-specific T cells rather than a true expansion of memory T cells.

https://doi.org/10.1038/s41590-023-01644-5

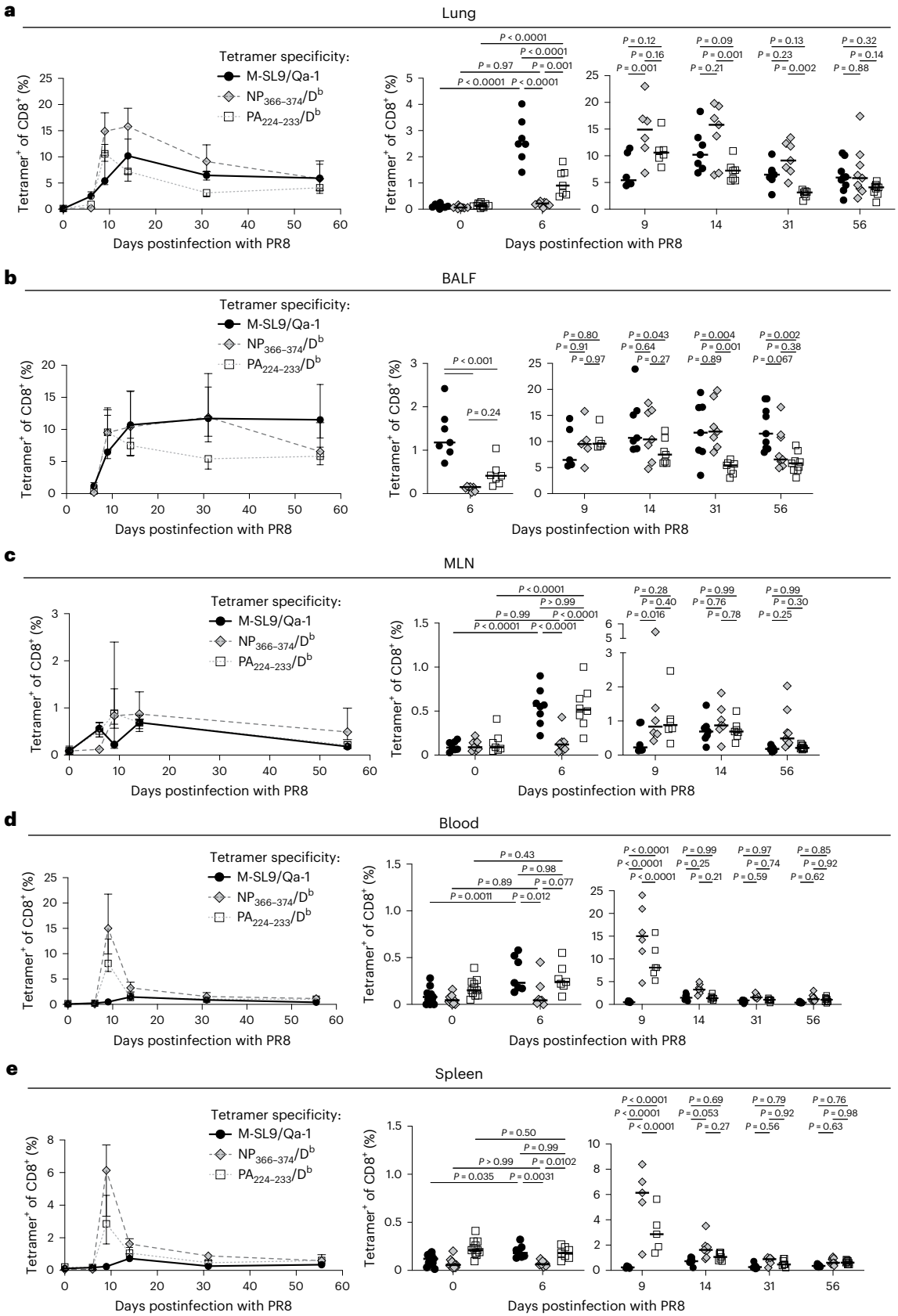


Fig. 4 | Distinct anatomical distribution and kinetics of M-SL9/Qa-1-specific CD8 T cells. a–**e**, C57Bl/6 mice were intranasally infected with 160 FFU of IAV PR8 and were killed at days 6 ($N = 7$), 9 ($N = 5$–6), 14, 31 ($N = 7$) or 56 ($N = 9$) to collect the indicated tissues/fluids from the lung (**a**), BALF (**b**), MLN (**c**), blood (**d**) and spleen (**e**). Uninfected mice were used as day 0 controls ($N = 7$–11). Samples were stained with tetramers to track antigen-specific CD8 T cell frequencies. Day 31 and 56 analyses were performed within 1–2 days of the reported times. Line graphs display the median ± interquartile range, and dot plots show the underlying data for individual mice and median. Data are combined from two independent experiments, with a mix of male and female mice. BALF from naive mice did not contain sufficient cell numbers for quantitation. *P* values reflect two-way ANOVA.

https://doi.org/10.1038/s41590-023-01644-5

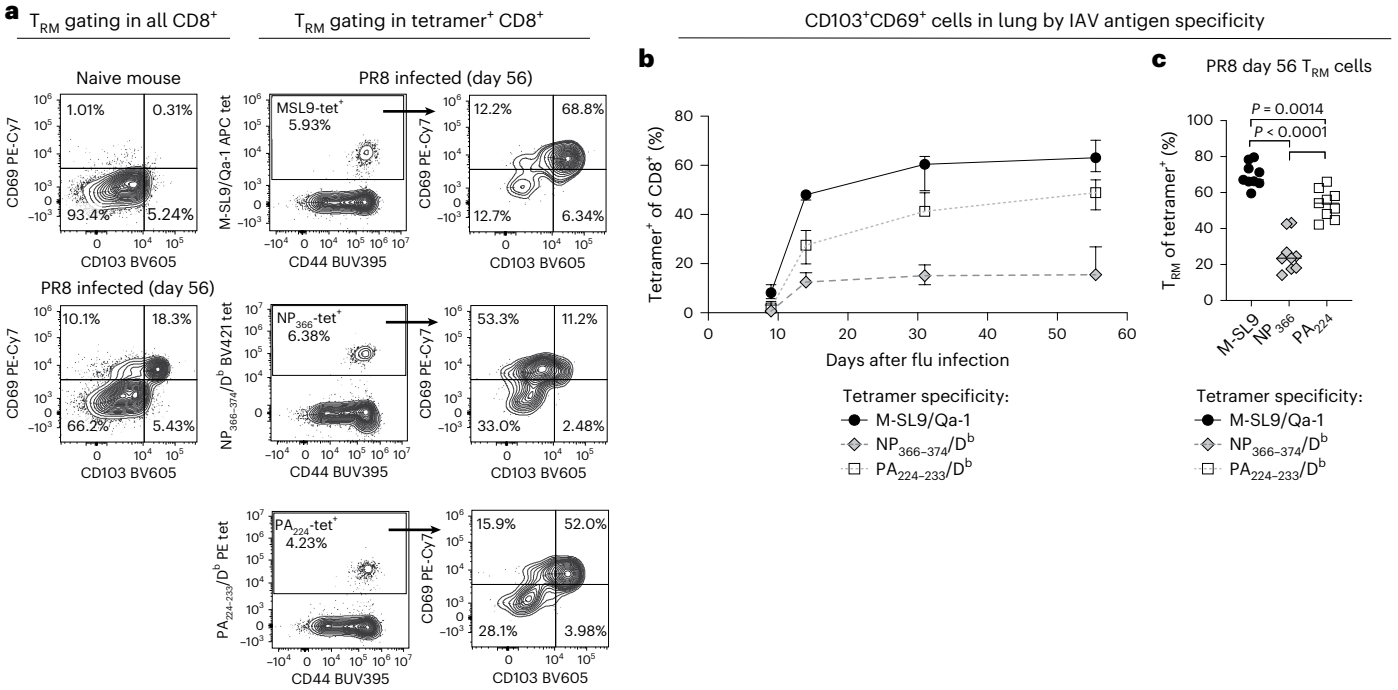


Fig. 5 | Efficient lung-resident memory cell differentiation of M-SL9/Qa-1-specific CD8 T cells. C57Bl/6 mice were intranasally infected with 160 FFU of IAV PR8 (the same mice examined in Fig. 4) and lungs were analyzed for expression of the trafficking molecules and lung T_{RM} cell markers, CD103 and CD69. **a**, Representative gating strategy showing tetramer$^+$ antigen-specific CD8 T cells and staining of CD103$^+$ CD69$^+$ cells. **b**, The line graph shows median ± interquartile range. **c**, The dot plot shows data for individual mice and median. $N = 9$ mice across two independent experiments and includes a mix of females and males. P values reflect one-way ANOVA.

Taken together, these results demonstrate that the M-SL9-specific CD8 T cell response is modestly enhanced by heterosubtypic reinfection, reaching a similar or slightly higher level compared with primary infection, depending on the tissue. Compared with the classical CD8 T cell specificities, the extent of M-SL9 recall during secondary infection was intermediate between that of $NP_{366–374}$ and $PA_{224–233}$.

M-SL9 presentation depends on AUG2

Given the cryptic origin of M-SL9, the mode in which this epitope is translated is not obvious. The in-frame codon immediately preceding M-SL9 is the second AUG (AUG2) in the matrix gene segment, located 85 nucleotides after the primary start codon (AUG1) shared by M1 and the matrix protein 2 (M2) ion channel (Extended Data Fig. 3). Therefore, we hypothesized that translation of M-SL9 is initiated at AUG2. Support for this theory was published by Machkovech et al., who performed ribosomal profiling of human cells infected with an IAV strain closely related to PR8 and found evidence that the ribosome can initiate on AUG2 (ref. 56). If translation starts at AUG2 and proceeds to the nearest in-frame stop codon, the resulting 16-mer peptide would be a novel gene product that we term M-MG16 (Fig. 7a). A less likely possibility is raised by a reported RNA splicing event that produces an alternative ion channel, called M42, which contains M-SL9 in its N-terminus joined to the second exon of M2; however, the M42 transcript was previously shown to not be expressed by WT PR8 (ref. 57).

To investigate these possibilities, we encoded M-MG16 and M42 using mRNA vectors[58] and tested their potential to present M-SL9 when introduced into B6-CIITA fibroblasts. Interestingly, a similar efficiency of M-SL9 presentation per mole of input mRNA was observed for M-MG16 and M42, and these were both slightly more efficient than a positive control mRNA encoding the full matrix gene segment (Fig. 7b), suggesting both constructs would be processed into M-SL9 if they were transcribed and translated during infection. To examine M-SL9 presentation during viral infection, we designed a panel of infectious PR8 variants with mutations in the matrix gene segment that are synonymous in the M1 ORF (Fig. 7c). The 'M42-up' virus was mutated to make M42 splicing more efficient, as confirmed by primer extension assay here (Fig. 7d) and in a prior study[57]. We found that an increase in M42 transcript from undetectable to abundant levels had no effect on the efficiency of M-SL9 presentation, suggesting that M42 is not an important source of M-SL9 presentation in PR8 infection (Fig. 7e). In fact, the sole RNA species encoding M-SL9 that our assay detected was the unspliced M1 transcript. PR8 was also mutated to ablate AUG2 ('ΔAUG2') (ref. 57) to test its role in M-SL9 presentation, and this largely eliminated activation of B6.23 cells but not of a control hybridoma specific for another epitope from PR8 (Fig. 7e). A stop codon was then introduced three codons upstream of M-SL9 ('preSTOP') to test whether M-SL9 is translated via ribosomal frameshifting[20] from the M1 ORF into the +1 RF between AUG1 and the mutated codon. Frameshifting in this region was ruled out, as M-SL9 presentation was not abrogated by this mutation (Fig. 7e); if anything, presentation was increased globally, possibly due to slightly increased M2 transcript expression[59] by this mutant (Fig. 7d).

To test the requirement for AUG2 in vivo, C57Bl/6 mice were infected with PR8 WT or ΔAUG2 viruses and the CD8 T cell response was evaluated. By tetramer staining and intracellular cytokine staining, M-SL9-specific responses were robust in WT-infected mice and undetectable in ΔAUG2-infected mice, indicating that AUG2 is strictly required for M-SL9 immunogenicity (Fig. 7f). As a control, $NP_{366–374}$ responses were comparably high in both cases. Collectively, these findings support a model in which M-SL9 presentation principally derives from 'leaky ribosomal scanning'[17,60] where the ribosome binds to the unspliced M1 transcript but scans past AUG1 and initiates at AUG2, resulting in M-MG16 peptide as the likely precursor to M-SL9.

To understand what other IAV strains might express and present M-SL9, we examined the conservation of this epitope and the preceding AUG2 codon across publicly available human and avian IAV sequences

https://doi.org/10.1038/s41590-023-01644-5

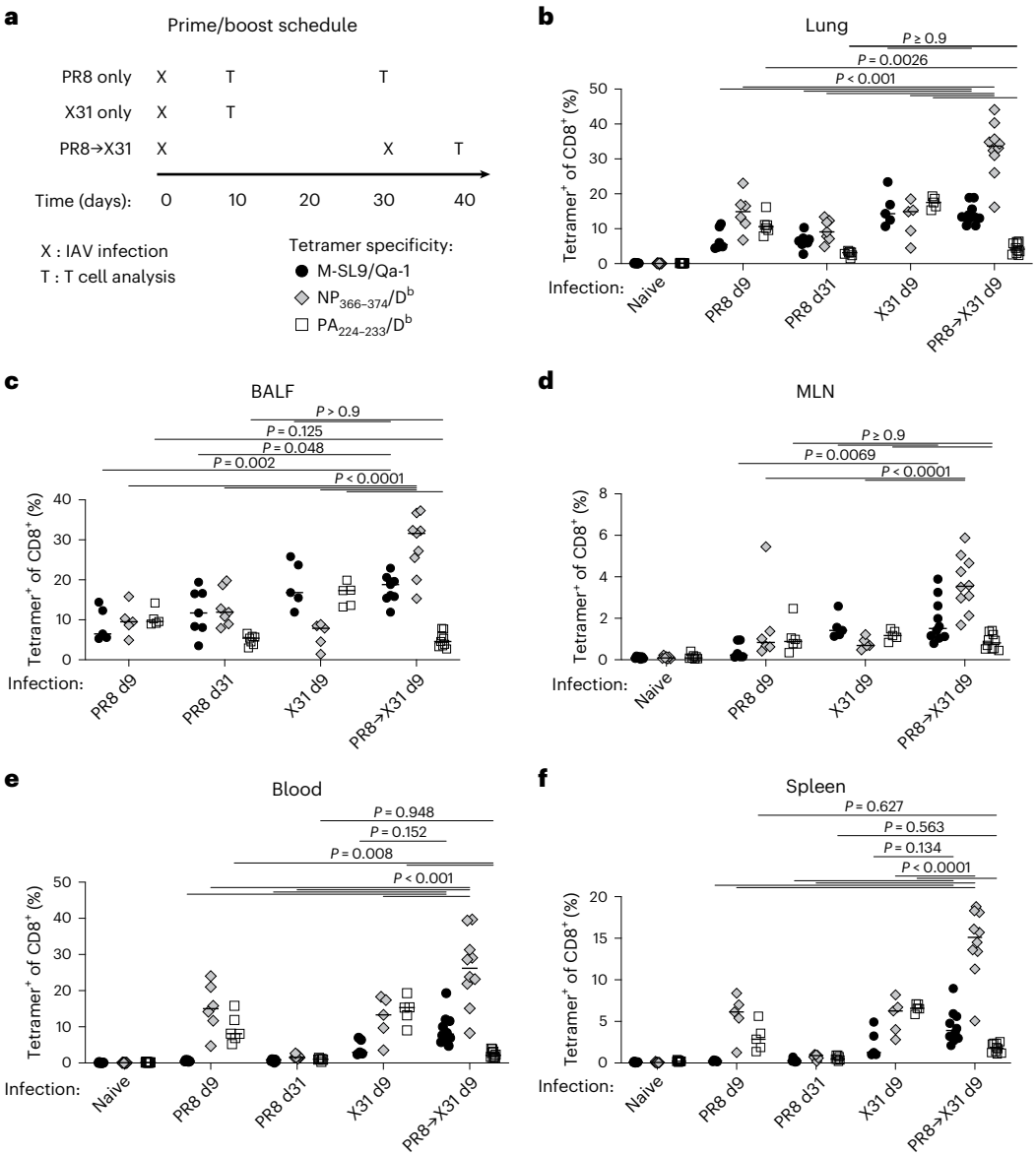


Fig. 6 | Recall of M-SL9/Qa-1-specific CD8 T cells during a secondary IAV infection. a–**f**, C57Bl/6 mice were intranasally infected with 160 FFU of IAV PR8 and/or 1×10^5 FFU of IAV X31 on the indicated schedule (**a**). The lung (**b**), BALF (**c**), MLN (**d**), blood (**e**) and spleen (**f**) were recovered and analyzed for antigen-specific CD8 T cell frequencies by tetramer staining at 9 days after PR8 ($N = 6$) or X31 ($N = 5$) primary infection only, 31 days after PR8 infection only ($N = 7$) or 9 days after primary PR8 followed by secondary X31 ($N = 10$). Day 0 controls (naive, $N = 11$) are identical to those shown in Fig. 4. Data are combined across two independent experiments with a mix of males and females, and horizontal bars are the median. *P* values reflect two-way ANOVA.

from 1980 until the present. The AUG2 codon was intact in ~100% of human H3N2 and pre-2009 H1N1 sequences but was replaced with valine in all post-2009 (pandemic) H1N1 strains (Extended Data Fig. 10a,b), and all lineages encoded a sequence with homology to M-SL9. Representative sequences since 1999 showed slow amino acid evolution but are all predicted to bind Qa-1 and HLA-E with moderate or strong confidence (Supplementary Table 2). Strikingly, avian H5N1 sequences from the current outbreak uniformly encode an intact AUG2 codon followed by a peptide with 100% identity to the M-SL9-P variant (Extended Data Fig. 10c) present in some isolates of PR8. These findings raise the possibility that M-SL9-like epitopes could be immunogenic in H3N2 and H5N1 strains of IAV currently circulating in humans and birds.

Cytolytic M-SL9 response is induced by mRNA vaccination

The mechanisms governing the induction of nonclassical CD8 T cell responses by infection and vaccination are poorly understood. In recent years, mRNA vaccines have proven to be a versatile and potent vaccine platform with outstanding clinical efficacy against coronavirus disease 2019 (COVID-19)[61]. However, the ability of this vaccine class to elicit non-classical CD8 T cells has not yet been explored. Here, we asked whether an mRNA vaccine would be able to elicit an M-SL9/Qa-1^b-specific CD8 T cell response and whether this would exhibit physiologically relevant effector functions in vivo.

M-SL9 was genetically fused to the C-terminus of a highly immunogenic[62] antigen, the surface glycoprotein (GP) of lymphocytic choriomeningitis virus (LCMV), and nucleoside-modified mRNA encoding this fusion protein ('GP:M-SL9') was encapsulated in lipid nanoparticles (LNP) following the design used by the Moderna COVID-19 vaccine[61] (Fig. 8a). C57Bl/6 mice were vaccinated with GP:M-SL9 mRNA–LNP or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike mRNA–LNP as a negative control, and a robust circulating M-SL9/Qa-1^b-specific CD8 T cell response was detected by tetramer staining

https://doi.org/10.1038/s41590-023-01644-5

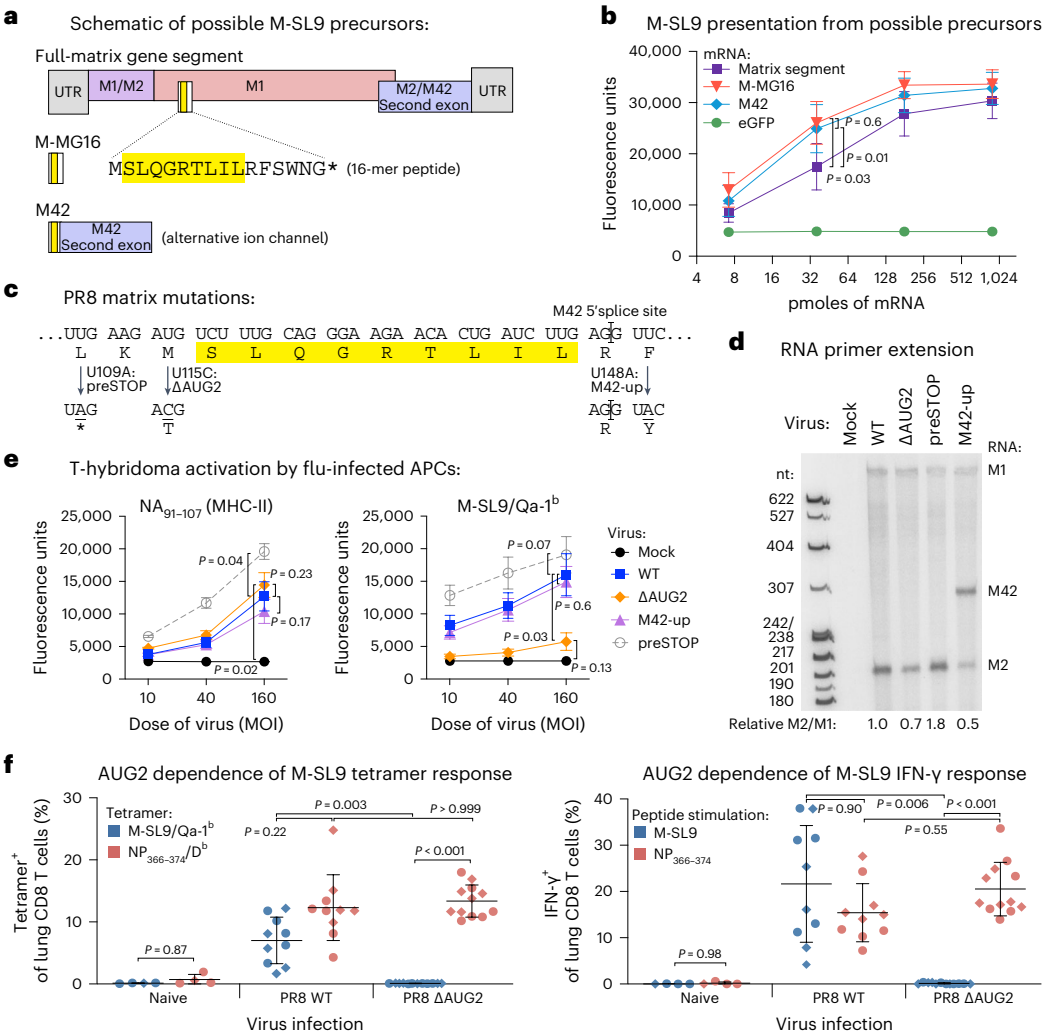


Fig. 7 | M-SL9 presentation depends on the second AUG codon of the matrix gene segment. a, Schematic of possible M-SL9 precursors. **b**, B6-CIITA fibroblasts were transfected with mRNA encoding the constructs in **a** and co-cultured with B6.23 hybridoma to read out M-SL9 presentation. **c**–**e**, PR8 was mutated as shown (**c**), then used to infect B6-CIITA cells to visualize matrix RNA species (**d**) and tested for M-SL9 and control epitope presentation to hybridomas (**e**). **f**, C57Bl/6 mice were infected with 160 FFU of the same viruses and day 9 lung CD8 T cell responses were quantified by tetramer and intracellular cytokine staining. $N = 10$ or 12 infected mice per group across two independent experiments, with female and male mice shown as circles and diamonds, respectively. Horizontal bars are mean ± s.d. In **a** and **c**, the M-SL9 epitope is highlighted in yellow. In **b** and **e**, data points represent the means of three independent experiments normalized to background ± s.e.m. In **b**, **e** and **f**, a one-way ANOVA was performed for doses of mRNA (**b**) or virus (**e**) yielding the greatest range of values, excluding negative controls, and *P* values are shown for comparisons of interest.

(Fig. 8b). Interestingly, M-SL9-specific T cells were abundant in the blood following intraperitoneal mRNA vaccination, in contrast to intranasal PR8 infection. Ten days after vaccination, in vivo cytolysis activity was tested by intravenously infusing vaccinated mice with naive donor spleen cells pulsed with M-SL9 or irrelevant peptide, which were labeled with high or low concentrations of carboxyfluorescein succinimidyl ester (CFSE) dye (Fig. 8c). After 18 h, spleens were recovered from recipient mice and analyzed by flow cytometry. M-SL9-pulsed cells were cleared in an epitope- and vaccine-specific manner at a mean rate of 46 ± 5.1% (s.e.m.) (Fig. 8d), and clearance of M-SL9-pulsed cells was significantly correlated with the magnitude of the M-SL9/Qa-1^b tetramer$^+$ response (Fig. 8e). To our knowledge, this is the first demonstration of a nonclassical cytotoxic T cell response elicited by an mRNA vaccine.

Discussion

In this study, we employed an immunopeptidomic approach and identified a dominant CD8 T cell response against an IAV PR8 epitope, M-SL9, with cryptic genomic origin and unexpected Qa-1 restriction in C57Bl/6 mice. The discovery of M-SL9 in some ways resembles that of another PR8 epitope, NS1-ARF2$_{1–8}$, which was found to elicit a CD8 T cell response in BALB/c mice and was mapped to a noncanonically translated alternative RF of the NS1 protein-coding sequence, believed to be translated into a nonfunctional 14-mer peptide[26]. In both cases, the cryptic CD8 T cells accounted for a major fraction of the T cell response against IAV and demonstrated a cytolytic phenotype. Interestingly, NS1-ARF2$_{1–8}$ was also discovered serendipitously, as it was the dominant specificity of CD8 T cells from PR8-infected mice after repeated in vitro stimulation and expansion by PR8-infected APCs. This work and ours raise the intriguing possibility that noncanonically translated epitopes, defective ribosomal products[63] and other kinds of cryptic epitopes[64] may drive a larger fraction of the T cell response than previously thought, arguing for the use of comprehensive methods to explore the immunopeptidome rather than relying on traditional approaches such as screening with overlapping peptide libraries. This may be especially important in autoimmune disease and cancer, where T cell-based therapies are particularly valuable but where the epitopes

https://doi.org/10.1038/s41590-023-01644-5

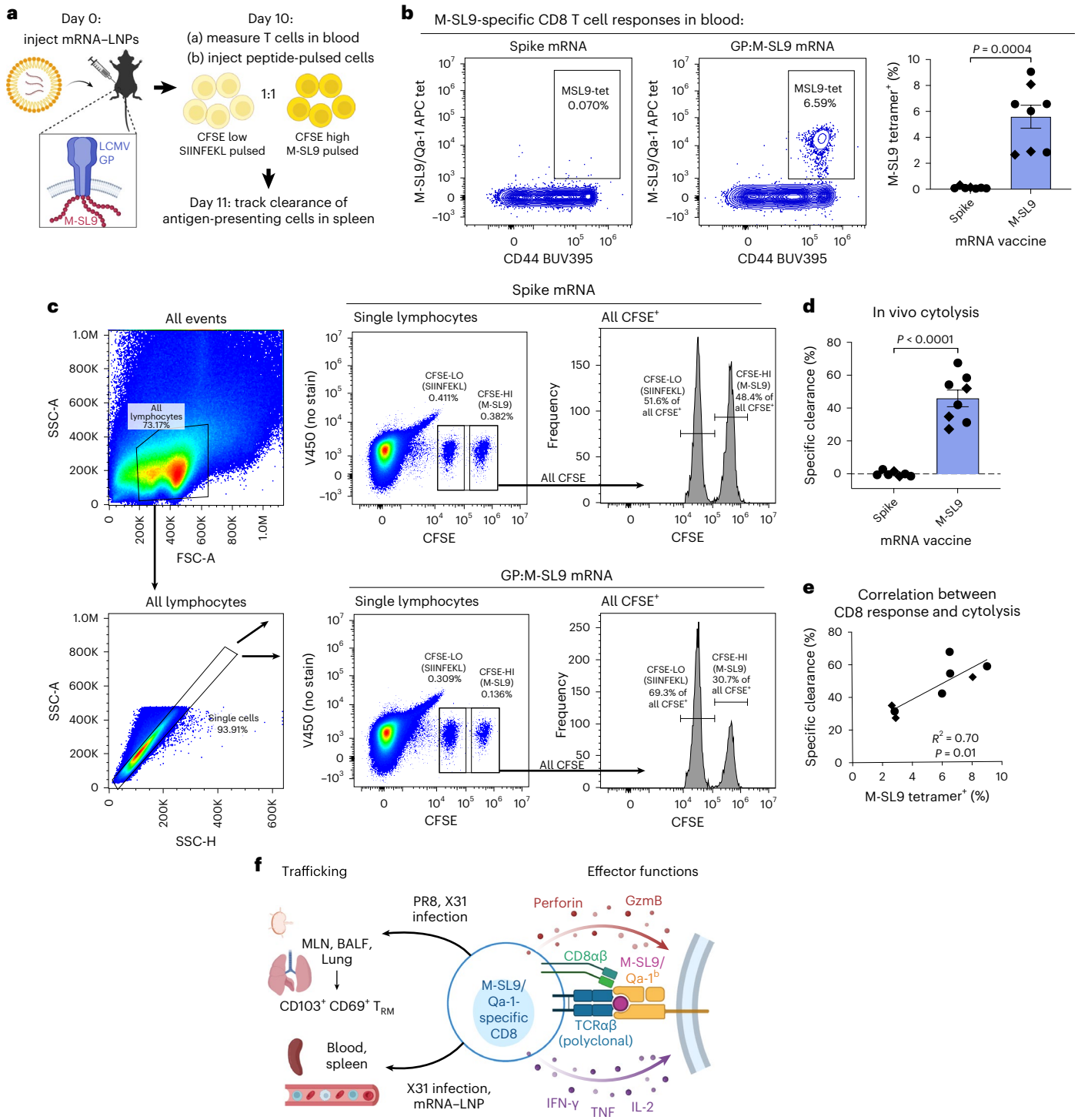


Fig. 8 | M-SL9/Qa-1-specific CD8 T cells are raised by mRNA vaccination and mediate cytolysis in vivo. a, C57Bl/6 were immunized intraperitoneally on day 0 with mRNA–LNP vaccines encoding M-SL9 fused to the C-terminus of the LCMV surface GP (GP:M-SL9; $N = 8$ mice) or SARS-CoV-2 spike protein ($N = 7$ mice), as a negative control, across two independent experiments. **b**, On day 10, M-SL9/Qa-1-specific CD8 T cells in blood were quantified by tetramer staining, with a representative plot shown. **c**,**d**, In vivo cytolysis assay: later on day 10, vaccinated mice were intravenously injected with ~5×10^6 spleen cells from naive donor mice in a 1:1 ratio of M-SL9 peptide-pulsed and SIINFEKL peptide-pulsed cells that were labeled with high and low concentrations of CFSE, respectively. After 18 h, spleens were recovered from the recipient mice and the relative frequencies of CFSE-high and CFSE-low cells were determined by gating (**c**) to compute the specific killing of M-SL9-pulsed cells (**d**). **e**, Correlation between M-SL9-specific CD8 T cell response and specific killing of M-SL9-pulsed cells. The bars represent mean ± s.e.m. Circles and diamonds represent female and male mice, respectively. **f**, Visual summary of M-SL9/Qa-1-specific CD8 T cell properties. The *P* values in **b** and **d** are from two-tailed Student's *t*-tests and in **e** from Pearson's correlation analysis. FSC-A: forward scatter-area, SSC-A/H: side scatter-area/height.

are often unknown and may be cryptic in origin[65–67]. It is even possible that cryptic epitopes are overrepresented in the immunopeptidome[68], since nonfunctional translation products are expected to skew toward short, unstructured polypeptides that are rapidly degraded, and MHC-I presentation appears to be dominated by rapidly degraded polypeptides[69–71].

Several aspects of the M-SL9-specific T cell response appear to be exceptional among IAV-specific CD8 T cell responses. In particular, M-SL9 was by far the dominant CD8 T cell specificity in the lung at 6 days postinfection—a critical time point associated with falling viral loads and rising disease symptoms in mice[72,73]. This finding prompts the hypothesis that M-SL9-specific and possibly other Qa-1-restricted CD8 T cells may play a role as 'early responders' in the lung during IAV infection, giving the classical CD8 T cell specificities such as $NP_{366–374}$ and $PA_{224–233}$ sufficient time to reach protective levels. There is precedent for just such a niche being filled by other nonclassical CD8 T cells restricted to the MHC-Ib molecule H2-M3 responding to *Listeria monocytogenes*[74]. Another notable property of M-SL9-specific CD8 T cells is their retention in the lung and efficient differentiation into lung T_{RM} cells. As noted previously, T_{RM} cells are considered an important component of cross-strain immunity to IAV[53,54], and therefore it will be valuable to ascertain whether Qa-1-restricted CD8 T cells contribute to IAV protection both at early time points (day 6) and at memory time points. Interestingly, M-SL9-specific cells also appeared to be amenable to 'flexible programming' in terms of their trafficking properties, since infection with X31 and an mRNA vaccine encoding M-SL9 were both able to induce M-SL9/Qa-1-specific cells to enter the circulation. This suggests the possibility that a vaccine or immune therapy harnessing MHC-E-restricted T cells might be able to direct tissue trafficking with some degree of specificity. For all the unique properties discussed above, it remains to be determined whether these are generalizable properties of Qa-1-restricted CD8 T cells or unique to M-SL9. It is also unclear whether Qa-1 restriction is a common feature of cryptic T cell responses or whether it was just a coincidence that M-SL9 is both cryptic and Qa-1-restricted. Future immunopeptidomic studies that analyze the full breadth of Qa-1 ligands presented on virus-infected cells will be critical to shed light on these questions, offering the possibility to analyze multiple Qa-1 epitope specificities and thus discern more general properties of MHC-E-restricted CD8 T cells.

The importance of Qa-1 (MHC-E) restriction by M-SL9 is augmented by the fact that MHC-E-restricted CD8 T cells are the only type of immune response so far that has been shown to protect against a pathogenic primate lentivirus infection. The Picker laboratory found that MHC-E-restricted CD8 T cells elicited by a CMV vector were able to control and eventually clear a severely pathogenic SIVmac239 infection in 50% of rhesus macaques, while classical CD8 T cells were totally unprotective[3]. The regulation of classical versus nonclassical CD8 T cell responses is largely undefined, but recent data implicate myeloid cell tropism of the vector as a contributor to the nonclassical response[4]. Since M-SL9 represents the first published viral epitope restricted to Qa-1 in mice, this provides a much-needed mouse model in which to investigate underlying mechanisms. We note that CD8 T cells are necessary but not sufficient for protection from IAV challenge in mice[75,76], and therefore it was outside the scope of this study to establish an experimental system in which to test for protection by M-SL9-specific T cells, but this will be an important area of follow-up investigation.

T cell-based vaccines and immune therapies that harness MHC-E-restricted CD8 T cells present at least two major advantages over classically restricted CD8 T cells. First, HLA-E is ubiquitously expressed and minimally polymorphic across the human population, having only two major allomorphs (HLA-E*01:01 and HLA-E*01:03) with identical peptide-binding grooves[77]. Therefore, vaccines and therapies could be designed for the general population as opposed to the personalized approaches that are often necessary when raising MHC-Ia-restricted T cells[78,79]. Second, the evidence from Picker and colleagues suggests that MHC-E-restricted T cells have unique protective properties compared with classically restricted T cells, leading to interest in targeting MHC-E epitopes in vaccine strategies against infectious pathogens and cancer[80]. In the present study, we demonstrated that a strong, circulating cytotoxic T cell response could be raised against M-SL9/Qa-1^b by an mRNA vaccine using the same approach as the highly successful COVID-19 vaccine from Moderna. This provides a proof-of-concept that mRNA vaccines can be used to explore the regulation and protective capabilities of MHC-E-restricted CD8 T cells, and it provides optimism that this safe and effective platform could be used to elicit these therapeutically promising T cell responses in humans.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-023-01644-5.

References

1. Hansen, S. G. et al. Immune clearance of highly pathogenic SIV infection. *Nature* **502**, 100–104 (2013).
2. Hansen, S. G. et al. A live-attenuated RhCMV/SIV vaccine shows long-term efficacy against heterologous SIV challenge. *Sci. Transl. Med.* **11**, eaaw2607 (2019).
3. Malouli, D. et al. Cytomegaloviral determinants of $CD8^+$ T cell programming and RhCMV/SIV vaccine efficacy. *Sci. Immunol.* **6**, eabg5413 (2021).
4. Hansen, S. G. et al. Myeloid cell tropism enables MHC-E–restricted $CD8^+$ T cell priming and vaccine efficacy by the RhCMV/SIV vaccine. *Sci. Immunol.* **7**, eabn9301 (2022).
5. Anderson, C. K., Reilly, E. C., Lee, A. Y. & Brossay, L. Qa-1-restricted $CD8^+$ T cells can compensate for the absence of conventional T cells during viral infection. *Cell Rep.* **27**, 537–548 (2019).
6. Chen, L., Jay, D. C., Fairbanks, J. D., He, X. & Jensen, P. E. An MHC class Ib-restricted $CD8^+$ T cell response to lymphocytic choriomeningitis virus. *J. Immunol.* **187**, 6463–6472 (2011).
7. Hansen, S. G. et al. Prevention of tuberculosis in rhesus macaques by a cytomegalovirus-based vaccine. *Nat. Med.* **24**, 130–143 (2018).
8. Rodgers, J. R. & Cook, R. G. MHC class Ib molecules bridge innate and acquired immunity. *Nat. Rev. Immunol.* **5**, 459–471 (2005).
9. Braud, V. M. et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature* **391**, 795–799 (1998).
10. Vance, R. E., Kraft, J. R., Altman, J. D., Jensen, P. E. & Raulet, D. Mouse CD94/NKG2A is a natural killer cell receptor for the nonclassical MHC class I molecule Qa-1b. *J. Exp. Med.* **188**, 1841–1848 (1998).
11. Anderson, C. K. & Brossay, L. The role of MHC class Ib-restricted T cells during infection. *Immunogenetics* **68**, 677–691 (2016).
12. Chen, X.-R. et al. A signal peptide derived from Hsp60 induces protective cytotoxic T lymphocyte immunity against lymphoid malignancies independently of TAP and classical MHC-I. *Cancer Lett.* **494**, 47–57 (2020).
13. Malouli, D. et al. Cytomegalovirus-vaccine-induced unconventional T cell priming and control of SIV replication is conserved between primate species. *Cell Host Microbe* **30**, 1207–1218.e7 (2022).
14. Grifoni, A. et al. SARS-CoV-2 human T cell epitopes: adaptive immune response against COVID-19. *Cell Host Microbe* **29**, 1076–1092 (2021).
15. Mahajan, S. et al. Immunodominant T-cell epitopes from the SARS-CoV-2 spike antigen reveal robust pre-existing T-cell immunity in unexposed individuals. *Sci. Rep.* **11**, 13164 (2021).
16. Saini, S. K. et al. SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial $CD8^+$ T cell activation in COVID-19 patients. *Sci. Immunol.* **6**, eabf7550 (2021).
17. Bullock, T. N. & Eisenlohr, L. C. Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. *J. Exp. Med.* **184**, 1319–1329 (1996).

18. Bullock, T. N. J., Patterson, A. E., Franlin, L. L., Notidis, E. & Eisenlohr, L. C. Initiation codon scanthrough versus termination codon readthrough demonstrates strong potential for major histocompatibility complex class I–restricted cryptic epitope expression. *J. Exp. Med.* **186**, 1051–1058 (1997).
19. Schwab, S. R., Li, K. C., Kang, C. & Shastri, N. Constitutive display of cryptic translation products by MHC class I molecules. *Science* **301**, 1367–1371 (2003).
20. Zook, M. B., Howard, M. T., Sinnathamby, G., Atkins, J. F. & Eisenlohr, L. C. Epitopes derived by incidental translational frameshifting give rise to a protective CTL response. *J. Immunol.* **176**, 6928–6934 (2006).
21. Starck, S. R. et al. Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* **336**, 1719–1723 (2012).
22. Apcher, S. et al. Translation of pre-spliced RNAs in the nuclear compartment generates peptides for the MHC class I pathway. *Proc. Natl Acad. Sci. USA* **110**, 17951–17956 (2013).
23. Goodenough, E. et al. Cryptic MHC class I-binding peptides are revealed by aminoglycoside-induced stop codon read-through into the 3′ UTR. *Proc. Natl Acad. Sci. USA* **111**, 5670–5675 (2014).
24. Yang, N. et al. Defining viral defective ribosomal products: standard and alternative translation initiation events generate a common peptide from influenza A virus M2 and M1 mRNAs. *J. Immunol.* **196**, 3608–3617 (2016).
25. Sanz, M. A., Almela, E. G., García-Moreno, M., Marina, A. I. & Carrasco, L. A viral RNA motif involved in signaling the initiation of translation on non-AUG codons. *RNA* **25**, 431–452 (2019).
26. Zanker, D. J. et al. Influenza A virus infection induces viral and cellular defective ribosomal products encoded by alternative reading frames. *J. Immunol.* **202**, 3370–3380 (2019).
27. Hanada, K., Yewdell, J. W. & Yang, J. C. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427**, 252–256 (2004).
28. Delong, T. et al. Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science* **351**, 711–714 (2016).
29. Paes, W. et al. Contribution of proteasome-catalyzed peptide cis-splicing to viral targeting by $CD8^+$ T cells in HIV-1 infection. *Proc. Natl Acad. Sci. USA* **116**, 24748–24759 (2019).
30. Tran, M. T. et al. T cell receptor recognition of hybrid insulin peptides bound to HLA-DQ8. *Nat. Commun.* **12**, 5110 (2021).
31. Purcell, A. W. Is the immunopeptidome getting darker?: a commentary on the discussion around Mishto et al., 2019. *Front. Immunol.* **12**, 720811 (2021).
32. Miller, M. A., Ganesan, A. P. V., Luckashenak, N., Mendonca, M. & Eisenlohr, L. C. Endogenous antigen processing drives the primary $CD4^+$ T cell response to influenza. *Nat. Med.* **21**, 1216–1222 (2015).
33. Tewari, M. K., Sinnathamby, G., Rajagopal, D. & Eisenlohr, L. C. A cytosolic pathway for MHC class II-restricted antigen processing that is proteasome and TAP dependent. *Nat. Immunol.* **6**, 287–294 (2005).
34. Reynisson, B. et al. Improved prediction of MHC II antigen presentation through integration and motif deconvolution of mass spectrometry MHC eluted ligand data. *J. Proteome Res.* **19**, 2304–2315 (2020).
35. Parker, R. et al. Mapping the SARS-CoV-2 spike glycoprotein-derived peptidome presented by HLA class II on dendritic cells. *Cell Rep.* **35**, 109179 (2021).
36. Partridge, T. et al. Discrimination between human leukocyte antigen class I-bound and co-purified HIV-derived peptides in immunopeptidomics workflows. *Front. Immunol.* **9**, 912 (2018).
37. Martínez-Sobrido, L. & García-Sastre, A. Generation of recombinant influenza virus from plasmid DNA. *J. Vis. Exp.* https://doi.org/10.3791/2057 (2010).
38. Ljunggren, H.-G. et al. Empty MHC class I molecules come out in the cold. *Nature* **346**, 476–480 (1990).
39. Kraft, J. R. et al. Analysis of Qa-1b peptide binding specificity and the capacity of CD94/NKG2A to discriminate between Qa-1–peptide complexes. *J. Exp. Med.* **192**, 613–624 (2000).
40. Ying, G., Wang, J., Kumar, V. & Zajonc, D. M. Crystal structure of Qa-1a with bound Qa-1 determinant modifier peptide. *PLoS ONE* **12**, e0182296 (2017).
41. Davies, A. et al. Infection-induced expansion of a MHC class Ib-dependent intestinal intraepithelial γδ T cell subset. *J. Immunol.* **172**, 6828–6837 (2004).
42. Tang, X. et al. Regulation of immunity by a novel population of Qa-1-restricted $CD8\alpha\alpha^+$ $TCR\alpha\beta^+$ T cells. *J. Immunol.* **177**, 7645–7655 (2006).
43. Niederlova, V., Tsyklauri, O., Chadimova, T. & Stepanek, O. $CD8^+$ Tregs revisited: a heterogeneous population with different phenotypes and properties. *Eur. J. Immunol.* **51**, 512–530 (2021).
44. Miller, J. D. et al. CD94/NKG2 expression does not inhibit cytotoxic function of lymphocytic choriomeningitis virus-specific $CD8^+$ T cells. *J. Immunol.* **169**, 693–701 (2002).
45. Borst, L. et al. NKG2A is a late immune checkpoint on CD8 T cells and marks repeated stimulation and cell division. *Int. J. Cancer* **150**, 688–704 (2022).
46. Van Montfoort, N. et al. NKG2A blockade potentiates CD8 T cell immunity induced by cancer vaccines. *Cell* **175**, 1744–1755.e15 (2018).
47. Borst, L., Van Der Burg, S. H. & Van Hall, T. The NKG2A–HLA-E axis as a novel checkpoint in the tumor microenvironment. *Clin. Cancer Res.* **26**, 5549–5556 (2020).
48. Lepore, M. et al. Parallel T-cell cloning and deep sequencing of human MAIT cells reveal stable oligoclonal TCRβ repertoire. *Nat. Commun.* **5**, 3866 (2014).
49. Yuan, R. et al. The roles of tissue-resident memory T cells in lung diseases. *Front. Immunol.* **12**, 710375 (2021).
50. Schön, M. P. et al. Mucosal T lymphocyte numbers are selectively reduced in integrin alpha E (CD103)-deficient mice. *J. Immunol.* **162**, 6641–6649 (1999).
51. Takamura, S. et al. Specific niches for lung-resident memory $CD8^+$ T cells at the site of tissue regeneration enable CD69-independent maintenance. *J. Exp. Med.* **213**, 3057–3073 (2016).
52. Walsh, D. A. et al. The functional requirement for CD69 in establishment of resident memory $CD8^+$ T cells varies with tissue location. *J. Immunol.* **203**, 946–955 (2019).
53. Laidlaw, B. J. et al. $CD4^+$ T cell help guides formation of $CD103^+$ lung-resident memory $CD8^+$ T cells during influenza viral infection. *Immunity* **41**, 633–645 (2014).
54. Zens, K. D., Chen, J. K. & Farber, D. L. Vaccine-generated lung tissue-resident memory T cells provide heterosubtypic protection to influenza infection. *JCI Insight* **1**, e85832 (2016).
55. Belz, G. T., Xie, W., Altman, J. D. & Doherty, P. C. A previously unrecognized H-2Db-restricted peptide prominent in the primary influenza A virus-specific $CD8^+$ T-cell response is much less apparent following secondary challenge. *J. Virol.* **74**, 3486–3493 (2000).
56. Machkovech, H. M., Bloom, J. D. & Subramaniam, A. R. Comprehensive profiling of translation initiation in influenza virus infected cells. *PLoS Pathog.* **15**, e1007518 (2019).
57. Wise, H. M. et al. Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS Pathog.* **8**, e1002998 (2012).
58. Pardi, N., Hogan, M. J. & Weissman, D. Recent advances in mRNA vaccine technology. *Curr. Opin. Immunol.* **65**, 14–20 (2020).

59. Liu, X. et al. MARCH8 inhibits influenza A virus infection by targeting viral M2 protein for ubiquitination-dependent degradation in lysosomes. *Nat. Commun.* **12**, 4427 (2021).
60. Kozak, M. Adherence to the first-AUG rule when a second AUG codon follows closely upon the first. *Proc. Natl Acad. Sci. USA* **92**, 2662–2666 (1995).
61. Hogan, M. J. & Pardi, N. mRNA vaccines in the COVID-19 pandemic and beyond. *Annu. Rev. Med.* **73**, 17–39 (2022).
62. Knudson, C. J., Hartwig, S. M., Meyerholz, D. K. & Varga, S. M. RSV vaccine-enhanced disease is orchestrated by the combined actions of distinct CD4 T cell subsets. *PLoS Pathog.* **11**, 1–23 (2015).
63. Wei, J. & Yewdell, J. W. Flu DRiPs in MHC class I immunosurveillance. *Virol. Sin.* **34**, 162–167 (2019).
64. Lodha, M., Erhard, F., Dölken, L. & Prusty, B. K. The hidden enemy within: non-canonical peptides in virus-induced autoimmunity. *Front. Microbiol.* **13**, 840911 (2022).
65. Kracht, M. J. L. et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat. Med.* **23**, 501–507 (2017).
66. Marcu, A. et al. Natural and cryptic peptides dominate the immunopeptidome of atypical teratoid rhabdoid tumors. *J. Immunother. Cancer* **9**, e003404 (2021).
67. Chong, C. et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat. Commun.* **11**, 1293 (2020).
68. Ruiz Cuevas, M. V. et al. Most non-canonical proteins uniquely populate the proteome or immunopeptidome. *Cell Rep.* **34**, 108815 (2021).
69. Croft, N. P. et al. Kinetics of antigen expression and epitope presentation during virus infection. *PLoS Pathog.* **9**, e1003129 (2013).
70. Wu, T. et al. Quantification of epitope abundance reveals the effect of direct and cross-presentation on influenza CTL responses. *Nat. Commun.* **10**, 2846 (2019).
71. Yewdell, J. W., Dersh, D. & Fåhraeus, R. Peptide channeling: the key to MHC class I immunosurveillance? *Trends Cell Biol.* **29**, 929–939 (2019).
72. Rutigliano, J. A. et al. Highly pathological influenza A virus infection is associated with augmented expression of PD-1 by functionally compromised virus-specific $CD8^+$ T cells. *J. Virol.* **88**, 1636–1651 (2014).
73. Vogel, A. J., Harris, S., Marsteller, N., Condon, S. A. & Brown, D. M. Early cytokine dysregulation and viral replication are associated with mortality during lethal influenza infection. *Viral Immunol.* **27**, 214–224 (2014).
74. Seaman, M. S., Wang, C.-R. & Forman, J. MHC class Ib-restricted CTL provide protection against primary and secondary *Listeria monocytogenes* infection. *J. Immunol.* **165**, 5192–5201 (2000).
75. Laidlaw, B. J. et al. Cooperativity between $CD8^+$ T cells, non-neutralizing antibodies, and alveolar macrophages is important for heterosubtypic influenza virus immunity. *PLoS Pathog.* **9**, e1003207 (2013).
76. LaMere, M. W. et al. Contributions of antinucleoprotein IgG to heterosubtypic immunity against influenza virus. *J. Immunol.* **186**, 4331–4339 (2011).
77. Kanevskiy, L. et al. Dimorphism of HLA-E and its disease association. *Int. J. Mol. Sci.* **20**, 5496 (2019).
78. Sahin, U. et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* **547**, 222–226 (2017).
79. Rojas, L. A. et al. Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer. *Nature* **618**, 144–150 (2023).
80. Voogd, L., Ruibal, P., Ottenhoff, T. H. M. & Joosten, S. A. Antigen presentation by MHC-E: a putative target for vaccination? *Trends Immunol.* **43**, 355–365 (2022).

[1]Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [2]School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, USA. [3]Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [4]The Jenner Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [5]Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [6]Department of Microbiology and Immunology, Thomas Jefferson University, Philadelphia, PA, USA. [7]Department of Molecular Microbiology and Immunology, Brown University, Providence, RI, USA. [8]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [9]Present address: Department of Pathology, University of Chicago, Chicago, IL, USA. [10]Present address: Century Therapeutics, Philadelphia, PA, USA. ✉e-mail: mihogan@upenn.edu; eisenlc@pennmedicine.upenn.edu

Methods

Mice

WT C57Bl/6 mice (strain 000664) and BALB/c mice (strain 000651) were purchased from the Jackson Laboratory and for some experiments were bred at the Children's Hospital of Philadelphia. C57Bl/6-background MHC-II$^{-/-}$ mice (strain 003584) and $\beta_2m^{-/-}$ mice (strain 002087) were purchased from the Jackson Laboratory. For collection of bone marrow, C57Bl/6-background K$^{b-/-}$ D$^{b-/-}$ mice were purchased from Taconic Biosciences, and K$^{b-/-}$ D$^{b-/-}$ Qa-1$^{b-/-}$ mice were derived previously[5] and housed at Brown University. Mice were matched for age and breeding location for all studies and 7–12-week-old mice were used for all experiments. Male and female mice were included in all experiments and the sex is indicated in key experiments in the figures and legends. For all infection and immunization experiments, mice were maintained in a specific pathogen-free facility at the Children's Hospital of Philadelphia at an ambient temperature of 68–74 °F with 30–70% humidity, with a 12 h light/dark cycle (lights on from 06:15 to 18:15). All animal procedures complied with all ethical guidelines for animal testing and research from our institutions, which are accredited by the American Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International. Animal protocols were approved by the Institutional Animal Care and Use Committees (IACUC) at the Children's Hospital of Philadelphia and Brown University.

Cell lines

The B6-CIITA and B6-CIITA-E^d cell lines were derived from an in-house C57Bl/6 skin fibroblast cell line by gamma retroviral transduction with human class II transactivator (*Ciita*), driving I-A^b expression, plus *Gfp* (B6-CIITA cells only) or BALB/c-derived *I-E^d* (B6-CIITA-E^d cells only; described previously[81]). HEK 293T/17 ('293T') cells (CRL-11268) were obtained from the American Type Culture Collection (ATCC). L929-K^b and L929-D^b cell lines were derived from L929 cells by retroviral transduction of *H2-D^b* or *H2-K^b* genes. All the above adherent cell lines were maintained in Dulbecco's modified eagle medium (DMEM) containing 10% fetal bovine serum (FBS), 2 mM L-glutamine, penicillin and streptomycin. Madin–Darby canine kidney (MDCK) cells were originally obtained from the National Institutes of Health (NIH) and were maintained in minimum essential medium (MEM) with 10% FBS. RMA-S, DC2.4 and A20 cells were maintained in 'R10 medium': Roswell Park Memorial Institute (RPMI) medium containing 10% FBS, 2 mM L-glutamine, penicillin and streptomycin. Mouse lymphocytes were cultured in R10 medium supplemented with 50 µM 2-mercaptoethanol in R10 medium.

The M-SL9-specific B6.23 hybridoma T cell line was created by fusing splenocytes from PR8-infected C57Bl/6 mice with a fusion partner, BWZ.36/CD8a, expressing a nuclear factor of activated T cells (NFAT)-inducible *lacZ* reporter gene, as previously described[82]. Resulting hybridoma clones were screened for reactivity to PR8-infected APCs. Several clones were initially unmappable and were later identified as reactive to M-SL9 peptide, with the strongest signal from clone B6.23. Other hybridoma T cell lines were made similarly and described previously, including those specific for HA$_{91–107}$/I-A^b ('HA-16' (ref. 32)), NA$_{437–451}$/I-A^b ('NA-110' (ref. 32)), NP$_{366–374}$/H2-D^b ('DBFZ.25' (ref. 82)) and SIINFEKL/H2-K^b ('B3Z' (ref. 82)). All hybridomas were maintained in R10 medium supplemented with 50 µM 2-mercaptoethanol.

HeLa cell lines used in Fig. 2 were a kind gift from the laboratory of Peter E. Jensen (University of Utah); previously, HeLa ctrl, Qa-1/D3 and H2-T11/D3 cells were referred to as HeLa MigR1, T23D3 and T11D3 cells[83]. The HeLa cell lines used in Extended Data Fig. 6 were a kind gift from Thorbald van Hall (Universiteit Leiden).

All cell lines used in this study were confirmed free of mycoplasma contamination by a PCR test kit (abm cat. no. G238) and the expected pattern of MHC expression on APC lines was confirmed functionally and by flow cytometry.

Viruses and infections

The PR8 virus was an infectious molecular clone derived from a set of eight ambisense plasmids in the pDZ vector[37], which encode each of the PR8 gene segments and are transcribed into both negative sense genomic RNA and positive sense mRNA when transfected into cells. WT PR8 virus stock and pDZ plasmids were a kind gift from Scott E. Hensley (University of Pennsylvania). Mutant PR8 viruses were created by site-directed mutagenesis of the pDZ M gene segment, making the following mutations based on the numbering used in Extended Data Fig. 3: U148A for M42-up, G145A for M42-down, U115C for ΔAUG2 (ref. 57) and U109A for preSTOP. These viruses and WT PR8 were launched in a co-culture of highly transfectable 293T cells and infection-susceptible MDCK cells and expanded in MDCK cells with OptiMEM medium containing 0.3% bovine serum albumin (BSA), 0.01% FBS, 0.1 mg ml^{-1} $CaCl_2$, penicillin–streptomycin and 3 µg ml^{-1} of trypsin treated with L-(tosylamido-2-phenyl) ethyl chloromethyl ketone (Worthington LS003750). PR8 plasmids and matrix RNA were sequence verified, and WT PR8 was identical to GenBank accession AF389121. Influenza virus B/Lee/1940 was propagated in the allantoic fluid of day 10 embryonated chicken eggs.

Influenza viruses were titered by focus-forming assay. Briefly, viruses were serially diluted on a monolayer of MDCK cells and cultured in MEM containing 1.25% Avicel (FMC RC-591 NF), fixed with 4% paraformaldehyde for 1 h, permeabilized with 0.5% Triton X-100, blocked with 5% nonfat dry milk in phosphate-buffered saline (PBS) and stained with chicken anti-IAV NP mAb IC5-1B7 (BEI NR-43899) or anti-B/Lee/1940 immune serum (NIH reference reagent no. V302-501-552), and developed using standard methods. The focus-forming unit (FFU) titers of mutant PR8 viruses were further corroborated by measuring the 50% tissue culture infectious dose.

Mice were infected with 25–200 FFU of WT or mutant PR8 by instilling 20–40 µl of PBS-diluted virus dropwise in the right nostril under isoflurane anesthesia. Early experiments used lower virus doses (25–40 FFU), while later experiments used higher doses (160–200 FFU) that yielded more consistently strong M-SL9-specific CD8 T cell responses, particularly when comparing male and female mice. Mice were monitored between days 5 and 10 after infection and euthanized if their clinical condition became severe (humane endpoint per the IACUC protocol).

Immunoprecipitation of peptide/MHC complexes

B6-CIITA-E^d cells (200 × 10^6) were either mock infected or PR8 infected in suspension in 0.1% BSA/PBS for 1 h and incubated at 37 °C, 6% CO_2 for ~18 h. Cells were collected with PBS containing 2 mM ethylenediaminetetraacetic acid (EDTA). Cell lysis and immunoprecipitation of peptide/MHC-II complexes were carried out largely as previously described[84], and all chemicals used after cell collection were analytical grade or equivalent when available. Briefly, anti-MHC-II antibody (clone M5/114.5, BioXCell) was cross-linked to Protein G Sepharose beads (GE 17061801). Lysates and beads were co-incubated and rotated overnight at 4 °C and then transferred to cleaned glass columns (Bio-Rad 7371012). After extensive washing, peptide/MHC-II complexes were eluted in 10% acetic acid, and the collected material was lyophilized.

High-performance liquid chromatography for peptide fractionation

After immunoprecipitation, purified peptide/MHC complexes were dissolved in 120 µl loading buffer (0.1% trifluoroacetic acid, 1% acetonitrile (ACN) in water). Using an Ultimate 300 high-performance liquid chromatography system (Thermo Fisher), samples were loaded onto a 4.6 × 50 mm ProSwift RP-1S (Thermo Fisher) monolith column, and peptides were eluted by applying a 10 min linear gradient from 3% to 30% ACN in 0.1% trifluoroacetic acid at a flow rate of 500 nl min^{-1}. Fractions were collected in 1 min intervals. Alternate fractions

(odd and even) were combined in two final fractions, dried, resuspended in 20 µl of loading buffer, and analyzed by LC–MS^2.

LC–MS^2

Peptides were analyzed by LC–MS^2 using an UltiMate 3000 RSLCnano System supplemented with a PepMap C18 column (2 µm particle size, 75 µm × 50 cm, Thermo Fisher) directly interphased with an Orbitrap Fusion Lumos Tribrid™ Mass Spectrometer (Thermo Fisher). A 60 min linear gradient from 3% to 25% ACN in 5% dimethyl sulfoxide (DMSO)/0.1% formic acid at a flow rate of 250 nl min^{-1} was applied for peptide elution. Peptide ions were introduced to the mass spectrometer using an Easy-Spray Source at 2,000 V. Detection was performed with a resolution of 120,000 for full MS (300–1,500 *m/z* scan range), and precursors were selected using TopSpeed ion selection within a 2 s cycle time, and a quadrupole isolation width of 1.2 a.m.u. for fragmentation. For MS^2 acquisition, the resolution was set at 30,000 and high-energy collisional dissociation energy was set at 28 for peptides with two to four charges and 32 for peptides that were singly charged.

LC–MS^2 data analysis

LC–MS^2 datasets were analyzed using PEAKS v8.5 with the following parameters: precursor mass tolerance: 5 ppm; fragment mass tolerance: 0.03 Da; digestion: none; fixed modifications: none; variable modifications: 313 (PEAKS posttranslational modification search). The search database included the full reference sequences of the pDZ PR8 gene segments translated in all six RFs (positive and negative sense), irrespective of start and stop codons, concatenated with the reviewed mouse entries in Uniprot (Swissprot version 17 July 2019). For immunopeptidomics data, a threshold of −10log*P* =15 was applied across all datasets, and the false discovery rate for peptide/sequence spectrum matches was reported as 3.4% for PR8-infected samples and 2.8% for uninfected control samples. Subsets of peptide identifications were analyzed by sequence logo diagrams (Extended Data Fig. 1).

Synthetic peptides

The following synthetic peptides were used: M-SL9 (SLQGRTLIL), M-SL9-P (SLQGRTPIL), $NP_{366–374}$ (ASNENMETM), $PA_{224–233}$ (SSLENFRAYV), $NS2_{109–121}$ (VEQEIRTFSFQLI) with the core epitope $NS2_{114–121}$ (RTFSFQLI), $OVA_{257–264}$ (SIINFEKL), $HA_{91–107}$ (RSWSYIVETPNSENGIC)[32], $HA_{121–137}$ (QLSSVSSFERFEIFPKE), $NA_{437–451}$ (TVDWSWPDGAELPFT)[32], $NP_{55–71}$ (RLIQNSLTIERMVLSAF), $NP_{264–280}$ (LILRGSVAHKSCLPACV, 'NP-45')[32] and $NP_{306–322}$ (LLQTSQVYSLIRPNENP, 'NP-52')[32]. All peptides were obtained at >85% purity from Genscript and dissolved in DMSO at a stock concentration of 10 mg ml^{-1}.

Preparation of mouse tissues

Mouse bone marrow was collected by flushing it from the femur and tibia with R10 medium. BMDCs were prepared by plating 2 million cells in 10 ml R10 medium in 10 cm tissue culture plates with 50 ng ml^{-1} granulocyte macrophage colony-stimulating factor ('GM-CSF', Shenandoah), with complete replacement of medium on day 3 and partial medium replacement as needed between day 6 and collection on day 9 or 10.

Mouse T cell responses were measured using cell suspensions prepared from various tissues as follows. BALF was collected immediately after killing by carbon dioxide inhalation (before lung collection) by exposing the trachea, inserting a catheter and securing with a suture tie, and using a syringe to gently lavage the lower airway with 1 ml of PBS/0.1 mM EDTA three times. BALF was collected into PBS 1% FBS, and ammonium-chloride–potassium (ACK) lysis was performed when red blood cells were visually detectable.

Lung samples were prepared by first injecting the right cardiac ventricle of mice with 3 ml of PBS 1% FBS to perfuse the lungs; then lungs were collected in gentleMACS C tubes containing 1% FBS in PBS on ice. Digest media was added to achieve a final concentration of 2.25 mg ml^{-1} sterile-filtered Collagenase D (Sigma 11088866001) and 0.15 mg ml^{-1} DNase I in 4 ml of 1% FBS in PBS. Lungs were disrupted using gentleMACS Dissociator program m_spleen_01.01 and incubated for 45 min at 37 °C with shaking. Then 10 ml of R10 medium was added to each tube, followed by further homogenization using gentleMACS Dissociator program m_lung_02.01. Digested lungs were passed through a 70 µm strainer, incubated in ACK lysis buffer to remove red blood cells, resuspended in R10 medium and then passed through a 40 µm strainer to obtain a single-cell suspension.

The MLN was collected after perfusing/expanding the lung with PBS 1% FBS (see above) but before removal of the lung. This was confirmed as the lung-draining lymph node based on marked expansion upon IAV infection. The MLN was collected with forceps into PBS 1% FBS and homogenized similarly to spleens (below), and ACK lysis was performed if red blood cells were visually detectable.

Spleens were collected in PBS and manually homogenized through a 70 µm cell strainer using the hard end of a syringe plunger, followed by ACK lysis and passage through a 40 µm strainer to prepare a single-cell suspension.

Blood was collected from the submandibular vein of live mice into a microcentrifuge tube containing PBS with 25 mM EDTA. Red blood cells were lysed using two incubations in ACK lysis buffer and the remaining cells were resuspended in R10 medium.

All primary cell samples were maintained at 0–4 °C during preparation (except as otherwise specified) and until analysis.

IFN-γ ELISpot for T cell activation

Spleens were collected from C57Bl/6 and BALB/c mice on day 10 postinfection with PR8 (25 FFU). CD4 T cells were purified from the splenocytes by negative isolation using Dynabeads Untouched Kit for mouse CD4 T cells, following the manufacturer's protocol (Thermo Fisher). Mouse IFN-γ ELISpot assays were carried out following the manufacturer's protocol (BD Biosciences). DC2.4 (C57Bl/6) or A20 (BALB/c) cells were used as APCs and cultured in R10 supplemented with 2 units ml^{-1} murine IFN-γ (BEI NR-3081) for 2 days before the assay. 5×10^4 APCs were co-cultured with 1×10^5 purified CD4 T cells, or 3×10^5 bulk splenocytes were cultured alone, with 10 µg ml^{-1} of peptide or DMSO vehicle, overnight at 37 °C before development of the ELISpot plate. IFN-γ spots were imaged and counted on a CTL ImmunoSpot S6 Universal Analyzer.

Flow cytometry and antibodies

All stains of primary mouse cells for flow cytometry utilized Fc receptor blockade with a final concentration of 10 µg ml^{-1} of anti-CD16/CD32 (bioXcell BE0307) and a viability stain using LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Thermo Fisher L34957). The following antibody specificities were used, with key details noted: CD3ε (145-2C11 or 17A2), CD4 (RM4-5), CD8α (53-6.7), CD8β (53-5.8), CD44 (IM7), CD62L (MEL-14), TCRβ (H57-597), TCRγδ (GL3), FoxP3 (FJK-16s), CD19 (6D5), NKG2A/C/E (20d5), IFN-γ (XMG1.2), TNF (MP6-XT22), IL-2 (JES6-5H4), CD107a Alexa Fluor 647 (ID4B, BioLegend 121610), granzyme B Pacific Blue (GB11, BioLegend 515408), perforin PE (S16009A, BioLegend 154305), CD103 (2E7), CD69 (H1.2F3), MHC-II (M5/114), β2 microglobulin (A16041A) and MHC-I H2-D^b α3 domain (28-14-8). Qa-1^b staining was performed with 2 µg ml^{-1} of PE-conjugated antibody clone 6A8.6F10.1A6 (BD 566641). Intracellular stains were performed using the BD Fixation/Permeabilization Solution Kit (554714) or eBioscience Foxp3/Transcription Factor Staining Buffer Set (Thermo Fisher 00-5523-00). In general, a final concentration of 2 µg ml^{-1} was used for all antibody stains. Compensation samples were prepared using UltraComp eBeads (Thermo Fisher 01-2222-42) and ArC Amine Reactive Beads (Thermo Fisher A10346). The CytoFLEX LX (Beckman Coulter) and Cytek Aurora cytometers were used for sample collection, and FlowJo v10.7 was used for analysis.

Tetramer staining

Peptide/MHC tetramers were obtained from the NIH Tetramer Core Facility, including $NP_{366-374}$/H2-D^b linked via biotin to

streptavidin-phycoerythrin (SA-PE) or SA-BV421, $PA_{224-233}$/H2-K^b linked to SA-PE, Qdm/Qa-1^b linked to SA-PE and M-SL9/Qa-1^b linked to SA-APC or SA-BV421. Tetramer stains were generally performed at 5 μg ml^{-1} for ~30 min at 4 °C in PBS 0.1% BSA buffer, except stains including Qdm/Qa-1^b tetramer, which were performed at 37 °C, as Qdm/Qa-1^b staining of NK cells is poor at 4 °C.

Intracellular cytokine and cytolytic marker staining for T cell activation

To perform intracellular cytokine staining, 1 million cells per well were stimulated in U-bottom 96-well tissue culture plates in R10 medium in the presence of synthetic peptides (10 μg ml^{-1} final concentration) and anti-CD28 co-stimulation (2 μg ml^{-1}; Tonbo 40-0281). Secretion was inhibited with 5 μg ml^{-1} brefeldin A (Biolegend 420601), and cells were incubated for 14 h at 37 °C, 6% CO_2. A viability stain was performed, followed by Fc receptor blockade, surface stain for CD8α, fixation and permeabilization and intracellular stains for CD3ε, CD4, IFN-γ, TNF and IL-2. Cytolytic markers were analyzed similarly except secretion inhibition was delayed by 1 h after antigen stimulation, 2 μM monensin and 2.5 μg ml^{-1} anti-CD107a were included along with brefeldin A in the 14 h incubation, and intracellular stains included granzyme B and perforin. Mice that did not show ≥2% of granzyme B^+ CD8 T cells in unstimulated (DMSO vehicle) lung samples were excluded from antigen-specific lung T -cell analyses as it was not clear that virus infected the lower respiratory tract. All reported cytokine frequencies are background subtracted, with the DMSO condition serving as background for peptide-stimulated conditions.

Hybridoma activation assay

Hybridoma recognition of cognate peptide/MHC complexes results in β-galactosidase production, detected using a fluorometric substrate, 4-methyl-umbelliferyl-β-D-galactopyranoside (Sigma M1633). Briefly, APCs were treated with antigen in various ways and then co-cultured overnight with hybridoma cells at a 1:2 ratio. Antigen stimulation with peptide was performed by pulsing APCs with mixtures of synthetic peptides (37.5–50 μg ml^{-1} per peptide) for 2 h in PBS and washing three times. In other cases, APCs were infected with influenza virus for 45 min in serum-free DMEM. In other cases, 50,000 APCs per well were transfected with mRNA using *Trans*IT-mRNA (Mirus, MIR 2225), with 0.34 μl of *Trans*IT-mRNA reagent and 0.22 μl of mRNA Boost reagent in a final volume of 200 μl per well. After ~18 h of co-culture, cells were treated with one-fifth volume of lysis/substrate buffer containing PBS with 1.25% Triton X-100, 33 μg ml^{-1} 4-methyl-umbelliferyl-β-D-galactopyranoside, 38.5 μM 2-mercaptoethanol and 9 mM $MgCl_2$, incubated for 3 h at 37 °C, 6% CO_2 and then ~24 h at 4 °C. Fluorescence (excitation 365 nm and emission 445 nm) was detected using an Infinite M200 Pro Plate Reader. APCs were validated for the expected expression of MHC molecules by flow cytometry. All antigen presentation assays contained three technical replicates per experiment, averaged together, and at least three independent experiments.

MHC stabilization assay with RMA-S cells

Peptide binding to MHC-Ia was tested using a method based on RMA-S cells deficient in TAP[38]. Briefly, RMA-S cells were incubated for 18 h at 26 °C to accumulate unstable 'empty' surface MHC-Ia molecules, lacking bound epitopes. The cells were then incubated at 37 °C for 1 h to allow internalization of unstable MHC and then stained for surface H2-D^b and H2-K^b and analyzed by flow cytometry.

TCR analysis of sorted cells

C57Bl/6 mice (n = 3 males and 3 females) were infected with 40 FFU (females) or 200 FFU (males) and confirmed to lose weight as a marker of infection. Lungs were collected and three populations of CD8 T cells were isolated from each lung by fluorescence-activated cell sorting (FACS): naive ($CD44^-$ $CD62L^+$), M-SL9-specific ($CD44^+$ M-SL9/Qa-1^b-APC tetramer$^+$) and $NP_{366-374}$/H2-D^b-PE tetramer$^+$). Roughly 89,000–390,000 cells were collected per specificity from each lung. Genomic DNA was isolated by the QIAamp DNA Micro Kit (Qiagen 56304), and the hypervariable region of the recombined *Tcrb* genes was sequenced, identified and analyzed for abundance by immunoSEQ (Adaptive Biotechnologies). Simpson's clonality was calculated by immunoSEQ as 1−Simpson diversity, a measure of alpha diversity. Further analyses of *Tcrb* V and J gene usage and CDR3 lengths were performed using the Immunarch software package in R, including spectratype graphs, a principal component plot and gene usage correlation plot.

mRNA production

Preclinical-grade mRNA was produced using previously described methods[85–87]. The constructs of interest (full-length PR8 matrix gene segment, M-MG16, M42, enhanced green fluorescent protein, LCMV GP:M-SL9 fusion protein or SARS-CoV-2 spike) were encoded by an mRNA production plasmid used previously[85]. The SARS-CoV-2 spike sequence contained a mutation ablating the furin-cleavage site and was published previously[86]. The GP:M-SL9 sequence was prepared using the LCMV GP protein sequence from GenBank accession NP_694851.1, immediately followed by the M-SL9-coding sequence and then a stop codon. To maximize translation of mRNA vectors, GP:M-SL9 and spike coding sequences were codon optimized. All mRNAs were produced by in vitro transcription (MEGAscript T7 Transcription Kit, Thermo AMB13345) using N1-methylpseudouridine (TriLink N-1081) in place of uridine, co-transcriptionally capped with the cap1 structure via CleanCap (TriLink N-7413) and purified of double-stranded RNA contaminants by adsorption to cellulose[87] (Sigma 11363). The mRNA length and integrity were confirmed by agarose gel electrophoresis.

LNP encapsulation of mRNA

SARS-CoV-2 spike and LCMV GP:M-SL9 mRNA–LNP vaccines were designed to mimic the approach of the Moderna COVID-19 vaccine[61]. mRNA was encapsulated in LNPs containing a mixture of ionizable cationic lipid SM-102 (BroadPharm BP-25499), cholesterol (Avanti 700100P), 1,2-distearoyl-sn-glycero-3-phosphocholine ('DSPC', Avanti 850365P) and 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 ('DMG-PEG-2000', Avanti 880151P) in a 50:38.5:10:1.5 molar ratio. In each encapsulation reaction, 1.1 mg of total lipid was dissolved in ethanol and buffered with 10% volume of aqueous citric acid (pH 4.0), 0.0625 mg of mRNA was prepared in citric acid buffer (pH 4.0), and the lipid and mRNA components were mixed at a 1:1 volume ratio (0.5 ml total volume) by rapid hand pipetting. This was promptly diluted with an equal volume of PBS (pH 7.4) and then dialyzed against PBS overnight at 4 °C. The resulting mRNA–LNPs were used within 2 days.

In vivo cytolysis assay

C57Bl/6 mice were immunized with a dose of spike or GP:M-SL9 mRNA–LNP vaccine containing 18 μg of mRNA by the intraperitoneal route. Ten days later, blood was collected from the submandibular vein and the T cell response was evaluated by M-SL9/Qa-1^b tetramer stain. Also on day 10, a single-cell suspension was made from the spleens of naive donor mice (sex-matched C57Bl/6). Cells were divided and (1) labeled with a low concentration of 0.6 μM CFSE (Thermo Fisher C34554) and pulsed with SIINFEKL peptide, or (2) labeled with a high concentration of 5 μM CFSE and pulsed with M-SL9 peptide. CFSE labeling was performed in PBS for 15 min at 37 °C, followed by washing. Peptide pulsing was performed with 20 μg ml^{-1} peptide in R10 medium for 40 min at 37 °C, followed by washing. These two cell populations were mixed at a 1:1 ratio, suspended in PBS and injected into the vaccinated mice via the retro-orbital sinus. After 18 h, spleens were recovered from these mice and analyzed by flow cytometry. In vivo cytolysis was read out as the percent specific clearance of M-SL9-pulsed ($CFSE^{high}$) cells relative to SIINFEKL-pulsed ($CFSE^{low}$) cells, normalized to the control spike-vaccinated mice, according to the formula: %Specific

killing = 100 × (1 − (%CFSE$^{high}_{Vaccine}$/%CFSE$^{low}_{Vaccine}$)/(%CFSE$^{high}_{ControlAvg}$/%CFSE$^{low}_{ControlAvg}$)), as previously described[88], where %CFSE^{low} and %CFSE^{high} represent percents of all singlet lymphocytes.

Primer extension analysis of viral RNA

Total cellular RNA was isolated from virus-infected (multiplicity of infection (MOI) of 20) or mock-infected B6-CIITA fibroblasts at 9 h postinfection. TRIzol and chloroform were added and RNA was extracted and precipitated from the aqueous layer. Then, 10 µg of each RNA sample was subjected to primer extension assay with ^{32}P-labeled primers targeting the IAV matrix segment (5′-GTATTTAAAGCGACGATAAATGCATTTG-3′) and mouse U1 snRNA (5′-TCAGCACATCCGGAGTGCAATGG-3′). Primer extension and sample denaturing were performed as previously described[89], and samples were resolved on a denaturing 5% polyacrylamide gel (5% polyacrylamide, 1 M urea, 0.5× Tris/borate/EDTA buffer, 0.2% ammonium persulfate and 0.1% *N*,*N*,*N*′,*N*′-tetramethylethylenediamine). Gels were dried on a gel dryer (Bio-Rad, model 573), exposed on a phosphor screen (GE Healthcare), and imaged on an Amersham Typhoon (GE Healthcare). Relative abundances were quantified using ImageJ software. The primer extension gel was performed twice.

IAV sequence analysis

Variation of the M-SL9 sequence was investigated by accessing the Global Initiative on Sharing All Influenza Data (GISAID) database[90] to download matrix gene segment nucleotide sequences isolated from human or avian species during various intervals since 1980, and preparing sequence logo diagrams using the ggseqlogo[91] software package in R (Extended Data Fig. 10). Sequences from IAV strains recommended by the World Health Organization for seasonal influenza vaccines were downloaded from GISAID, Bacterial and Viral Bioinformatics Resource Center (BV-BRC) or the Influenza Research Database and analyzed as representative examples. Binding predictions for Qa-1^b and HLA-E allomorphs are provided as percentiles in Supplementary Table 2 and were obtained with the Immune Epitope Database (IEDB) MHC-I Binding Prediction tool using the NetMHCpan EL 4.1 algorithm.

Statistics

Intracellular cytokine, cytolytic marker, tetramer frequencies and fluorescence units in antigen presentation assays were normally distributed and were analyzed by one-way analysis of variance (ANOVA) (Brown–Forsythe and Welch test with Dunnett T3 method to correct for multiple comparisons) or two-way ANOVA for time course experiments (with Tukey test to correct for multiple comparisons). In vivo cytolysis results were analyzed by a two-tailed Student's *t*-test as there were only two groups. Microsoft Excel was used to manage data. GraphPad Prism v.9.5.1 was used to compute ANOVA and *t*-test results and to plot the four-parameter logistic sigmoidal curve in Extended Data Fig. 5. Immunarch 0.9.0 was used in R 4.3.0 to perform the principal component and Pearson correlation analyses of TCR gene usage, and ggseqlogo version 0.1 was used to create sequence logo diagrams in R.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the identifiers PXD045025 and https://doi.org/10.6019/PXD045025. Data supporting TCR sequence analysis are deposited in GitHub at the URL https://github.com/mhogan240/NatImmuno2023. Data supporting M-SL9 sequence variation analysis are publicly available from BV-BRC. Other data and reagents that support the findings of this study are available from the corresponding authors Michael J. Hogan and Laurence C. Eisenlohr upon request. Source data are provided with this paper.

Code availability

Code supporting analysis of TCR sequences and M-SL9 variant sequences is deposited in GitHub at the URL https://github.com/mhogan240/NatImmuno2023.

References

81. Sinnathamby, G., Maric, M., Cresswell, P. & Eisenlohr, L. C. Differential requirements for endosomal reduction in the presentation of two H2-Ed-restricted epitopes from influenza hemagglutinin. *J. Immunol.* **172**, 6607–6614 (2004).
82. Sanderson, S. & Shastri, N. LacZ inducible, antigen/MHC-specific T cell hybrids. *Int. Immunol.* **6**, 369–376 (1994).
83. Chen, L. et al. Expression of the mouse MHC class Ib H2-T11 gene product, a paralog of H2-T23 (Qa-1) with shared peptide-binding specificity. *J. Immunol.* **193**, 1427–1439 (2014).
84. Purcell, A. W., Ramarathinam, S. H. & Ternette, N. Mass spectrometry–based identification of MHC-bound peptides for immunopeptidomics. *Nat. Protoc.* **14**, 1687–1707 (2019).
85. Pardi, N., Muramatsu, H., Weissman, D. & Karikó, K. In vitro transcription of long RNA containing modified nucleosides. *Methods Mol. Biol.* **969**, 29–42 (2013).
86. Laczkó, D. et al. A single immunization with nucleoside-modified mRNA vaccines elicits strong cellular and humoral immune responses against SARS-CoV-2 in mice. *Immunity* **53**, 724–732.e7 (2020).
87. Baiersdörfer, M. et al. A facile method for the removal of dsRNA contaminant from in vitro-transcribed mRNA. *Mol. Ther. Nucleic Acids* **15**, 26–35 (2019).
88. Freyn, A. W. et al. A multi-targeting, nucleoside-modified mRNA influenza virus vaccine provides broad protection in mice. *Mol. Ther.* **28**, 1569–1584 (2020).
89. Thompson, M. G. et al. Co-regulatory activity of hnRNP K and NS1-BP in influenza and human mRNA splicing. *Nat. Commun.* **9**, 2407 (2018).
90. Shu, Y. & McCauley, J. GISAID: global initiative on sharing all influenza data—from vision to reality. *Eurosurveillance* **22**, 30494 (2017).
91. Wagih, O. ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* **33**, 3645–3647 (2017).

Acknowledgements

We thank F. Tuluc, J. Murray and J. Lora of the Children's Hospital of Philadelphia Flow Cytometry Core Facility for technical advice and services; L. Spruce, H. Fazelinia and S. Seeholzer (formerly) of the Children's Hospital of Philadelphia Proteomics Core Facility for technical guidance and services; the NIH Tetramer Core Facility for providing tetramers for this study; J. R. Melamed and D. Weissman for technical advice on LNP generation; R. Serafin for providing related data; J. J. Rim for assistance with manuscript preparation; and D. F. Jenkins for data management support. We gratefully acknowledge the contributors to the Influenza Research Database, BV-BRC and the GISAID database, including the laboratories and authors responsible for obtaining specimens, generating genetic sequences and sharing data via the GISAID Initiative. M.J.H. was supported by the Cancer Research Institute as a Cancer Research Institute Irvington Fellow and by the Roberts Family–Katalin Karikó Fellowship in Vaccine Development from the Aileen K. and Brian L. Roberts Family Foundation via the University of Pennsylvania Institute for Immunology & Immune Health (I3H). N.M. was supported by the Roy and Diana Vagelos Molecular Life Sciences Program and by a College Alumni Society Research Grant from the University of Pennsylvania.

N.P. was supported by NIH R01AI146101 and R01AI153064. S.P.R. is supported by research supplement 3R01AI046709-18S1 to promote diversity and L.B. is supported by NIH R01AI046709. K.W.L. and B.E.B. are supported by R01AI125524. L.C.E. and N.T. were supported by NIH R21AI153978. This work was funded in part by contract #75N93021C00015 from NIH NIAID. BioRender.com was used to create panels in Figs. 1, 2 and 8.

Author contributions

M.J.H. conceived the project, designed the studies, performed experiments, analyzed and interpreted the data and wrote the paper. N.M. co-conceived the project, co-designed studies, co-performed the immunoprecipitation, ELISpot and antigen presentation experiments and analyzed and interpreted data. B.E.B. performed primer extension assays and provided interpretation together with K.W.L. A.N. and N.T. performed LC–MS^2, analyzed data and provided interpretation. E.J.H. provided essential support for animal studies. H.M. and N.P. contributed mRNA reagents and related expertise. M.A.M. isolated the B6.23 hybridoma. S.P.R. and L.B. provided reagents and expertise regarding MHC-Ia and Qa-$1^{b-/-}$ bone marrow. L.C.E. co-conceived and advised the project and interpreted the data. All authors provided critical scientific feedback, aided in the preparation of the manuscript and agree with the conclusions.

Competing interests

N.T. is or has been a paid consultant to Roche Pharma, Enara Bio, Grey Wolf Therapeutics, T-Cypher Bio and Infinitopes on the topic of cancer antigen discovery. All other authors declare no competing interests.

Additional information

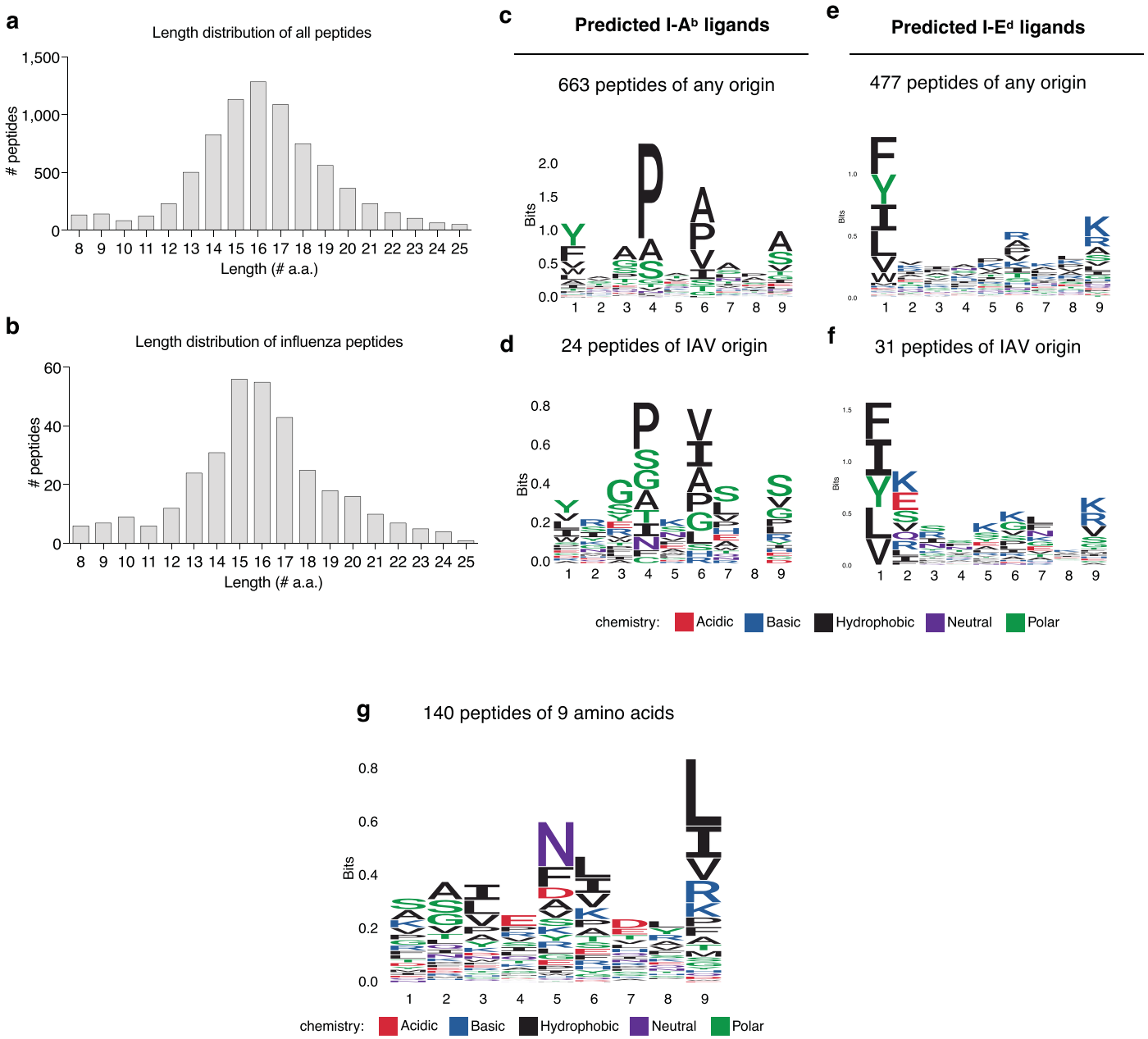
Extended data is available for this paper at https://doi.org/10.1038/s41590-023-01644-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-023-01644-5.

Correspondence and requests for materials should be addressed to Michael J. Hogan or Laurence C. Eisenlohr.

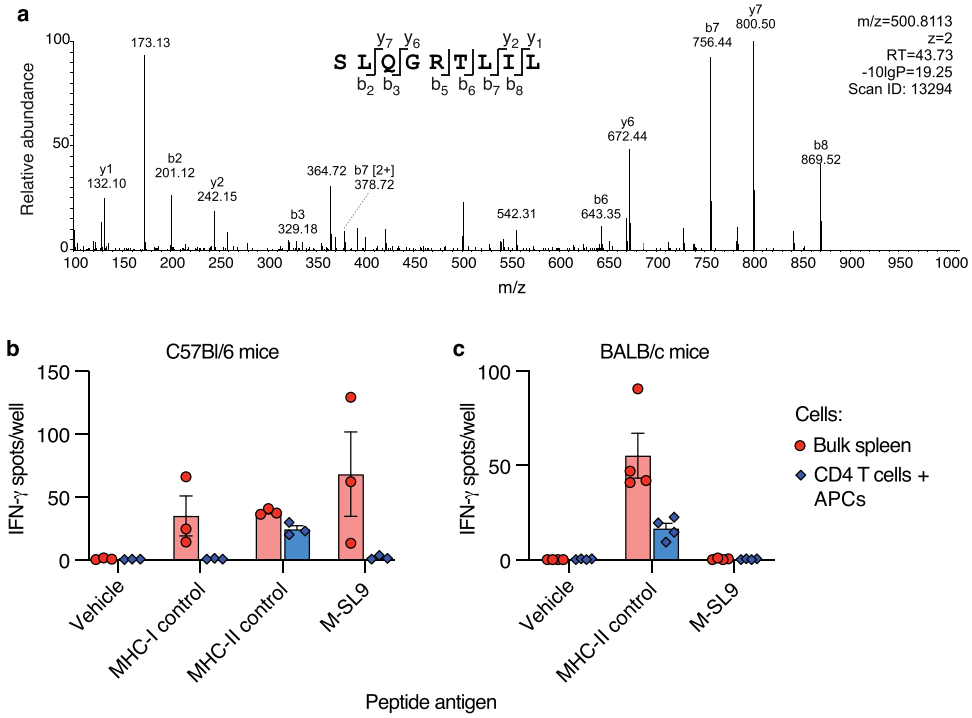
Peer review information *Nature Immunology* thanks Katherine Kedzierska and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: S. Houston, in collaboration with the *Nature Immunology* team.

https://doi.org/10.1038/s41590-023-01644-5



Extended Data Fig. 1 | Peptide analysis by mass spectrometry (MS). **a**, **b**, Histograms of peptide lengths of unique peptide species identified from (a) any origin and (b) IAV origin. **c**–**f**, NetMHCIIpan 4.0 was used to predict peptide:MHC-II affinity (K_D) values for core epitope sequences of 9 amino acids from MS-identified peptides of all lengths. Sequence logo diagrams were prepared using unique core epitopes predicted to bind to (c,d) I-A^b or (e,f) I-E^d with a K_D < 2,000 nM for peptides of (c,e) any origin or K_D < 10,000 nM for peptides of (d,f) IAV origin. These sequence logos clearly exhibit the expected MHC-II ligand sequence motifs (based on ref. 34). **g**, The sequence logo diagram for all unique 9-mer peptide identifications (appearing as a small local peak in panel 'a') shows strong similarity to the H2-K^b and H2-D^b sequence motifs, but not to the Qa-1 sequence motif (motifs available from NetMHCpan 4.0 at ref. 82). In sequence logo diagrams, 'bit' is a unit of relative amino acid frequency that is inversely related to the Shannon entropy of each position.

https://doi.org/10.1038/s41590-023-01644-5

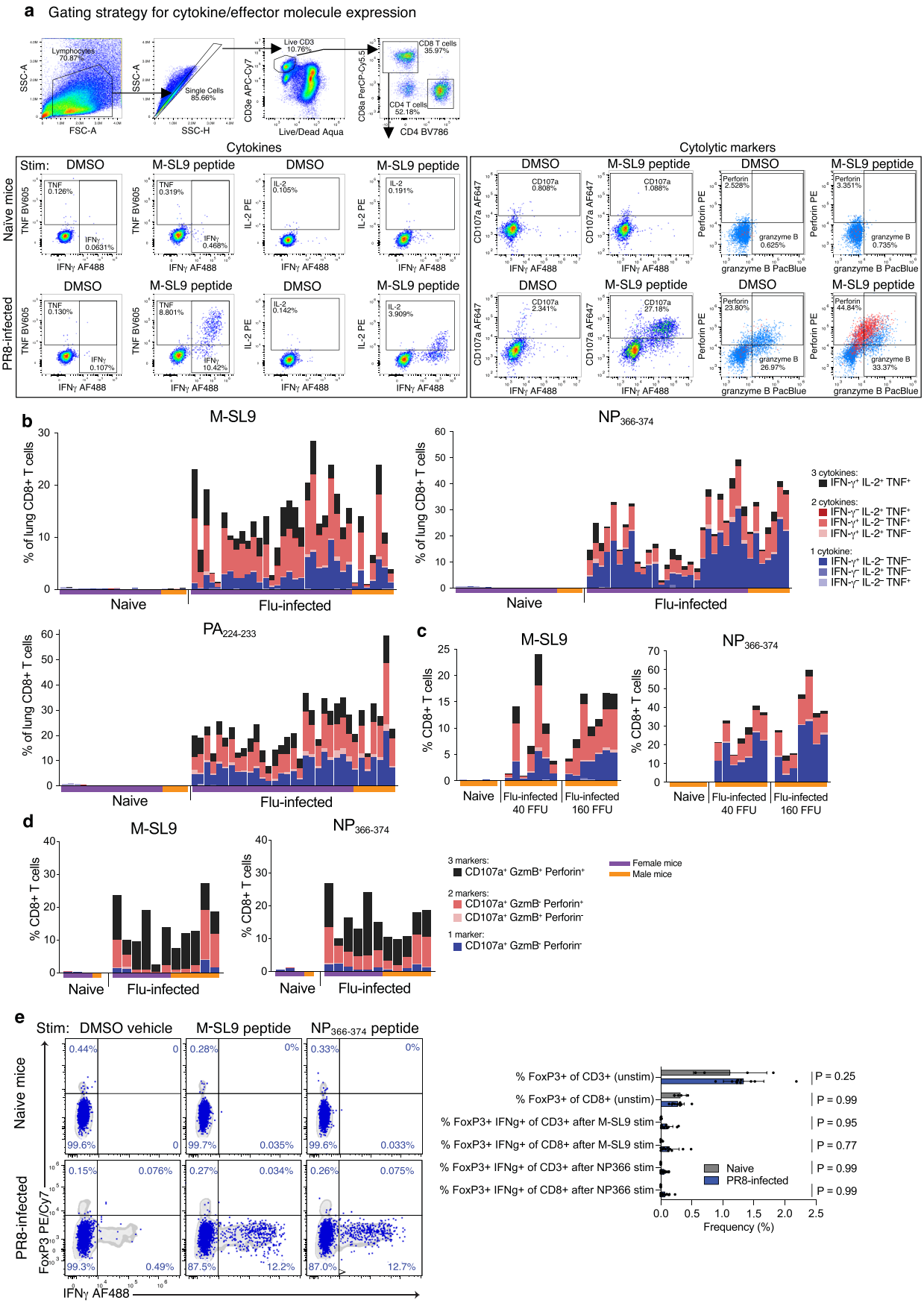


Extended Data Fig. 2 | Identification of M-SL9 by mass spectrometry and validation of immunogenicity in C57Bl/6 mice. a, MS^2 spectrum resulting in M-SL9 identification, with b and y ion fragments indicated, along with mass/charge (m/z), retention time (RT), and *P*-value. **b**, **c**, Spleens were recovered from C57Bl/6 (N = 3) or BALB/c (N = 4) mice 9 days after infection with IAV PR8, and either bulk spleen cells or isolated spleen CD4 T cells were stimulated overnight with synthetic M-SL9 peptide, positive control peptides, or DMSO vehicle, and secreted IFN-γ was detected by ELISpot. **b**, For C57Bl/6 mice, the MHC-Ia control peptide was $NS2_{109\text{-}121}$; the MHC-II control peptides were $NP_{264\text{-}280}$ and $NP_{306\text{-}322}$; and DC2.4 cells were used as the APC to stimulate CD4 T cells. **c**, For BALB/c mice, the MHC-II control peptides were $NP_{55\text{-}71}$ and $HA_{121\text{-}137}$, and A20 cells were used as the APC to stimulate CD4 T cells. Data points represent individual mice in one independent experiment; bars are the mean +/− s.e.m.

https://doi.org/10.1038/s41590-023-01644-5

Extended Data Fig. 3 | Annotation of M-SL9 coding sequence within the full IAV PR8 matrix gene segment. The positive-sense RNA sequence is shown on top. Primary and secondary AUG codons are underlined and labeled, M-SL9 amino acids are highlighted in yellow, the M1 protein sequence is in light blue, M-MG16 nucleotides including stop codon are colored red, and relevant splice sites are labeled. The first nucleotide of each codon is aligned with the single-letter amino acid code and the first digit of the nucleotide number. This sequence was used to generate PR8 for this study and matches the sequence in GenBank accession AF389121.

https://doi.org/10.1038/s41590-023-01644-5

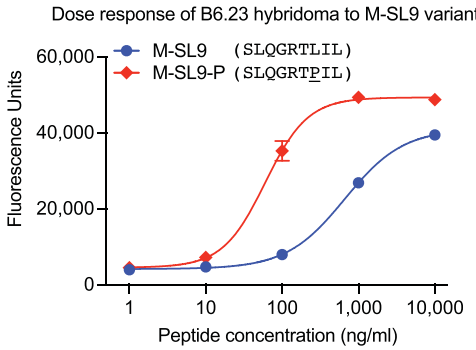


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41590-023-01644-5

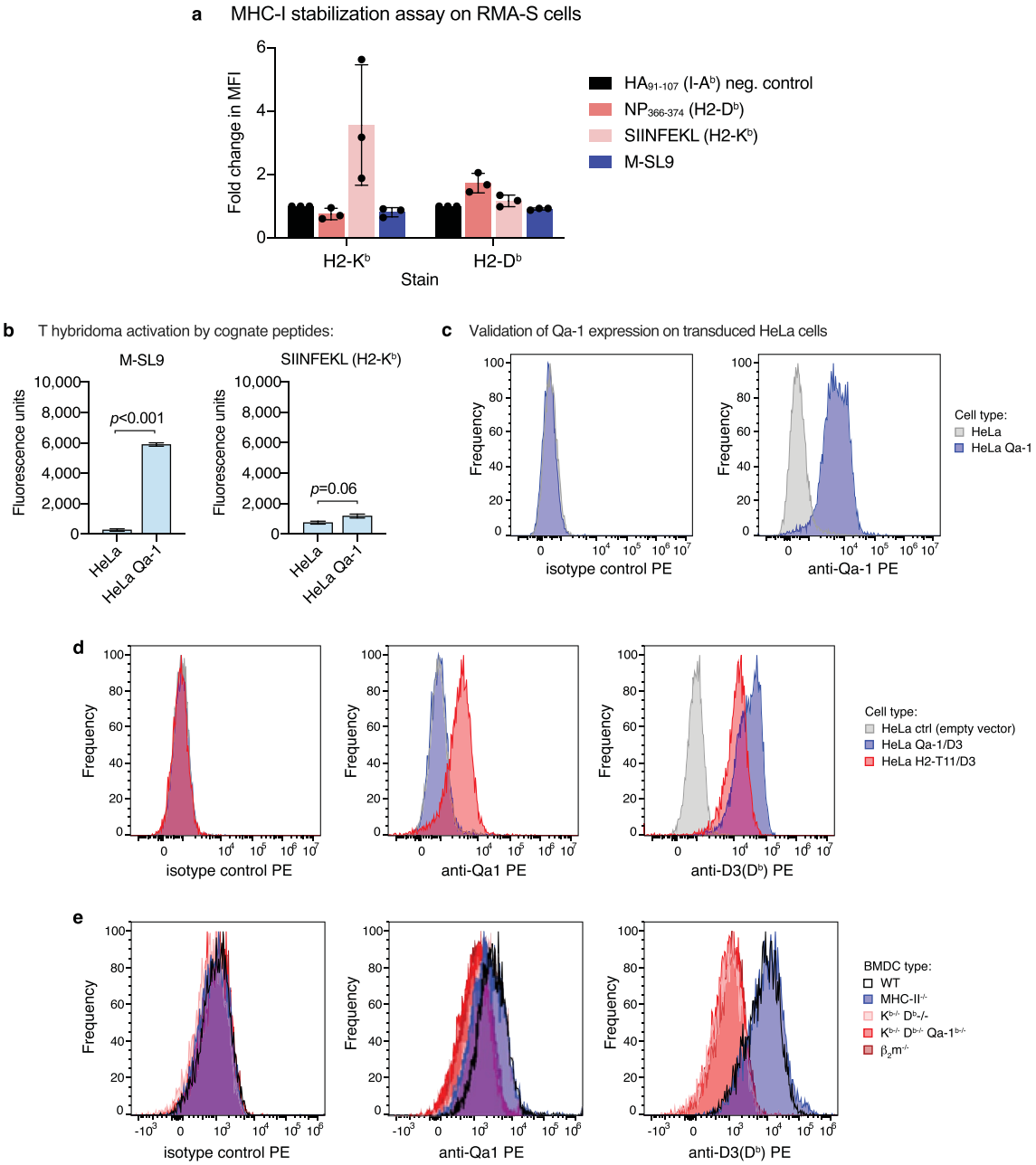
Extended Data Fig. 4 | Intracellular cytokines, cytolytic markers, and FoxP3 expression in individual mice. a–**d**, Representative gating strategy for intracellular cytokine and cytolytic marker staining (lower boxes from two different experiments/stains). **b**–**e**, Lymphocytes were isolated from naïve (N = 21) or PR8 flu-infected (N = 34) C57Bl/6 mouse lungs, stimulated with indicated peptides, and stained for the indicated markers. **b**–**d**, Data for individual mice are shown in the same order for each epitope. **c**, Comparison of intracellular cytokine responses following infection with 40 FFU PR8 (N = 7), 160 FFU PR8 (N = 7), or no virus (naïve; N = 5), showing more consistent M-SL9 responses to 160 FFU. Female and male mice are indicated by purple and orange bars underneath the graphs. **e**, CD8 T cells (both total unstimulated as well as peptide-restimulated IFN-γ^+ cells) from PR8-infected mice (N = 11) stain do not upregulate FoxP3 expression relative to the naïve (N = 4) mouse baseline. Gray events are all $CD3^+$ cells; blue events and blue percentages represent $CD3^+$ $CD8^+$ cells. Bars show the mean +/- s.d. and *P*-values of interest are shown from a two-way ANOVA with Sidak's multiple comparisons test comparing naïve and PR8-infected conditions. GzmB: granzyme B.

https://doi.org/10.1038/s41590-023-01644-5



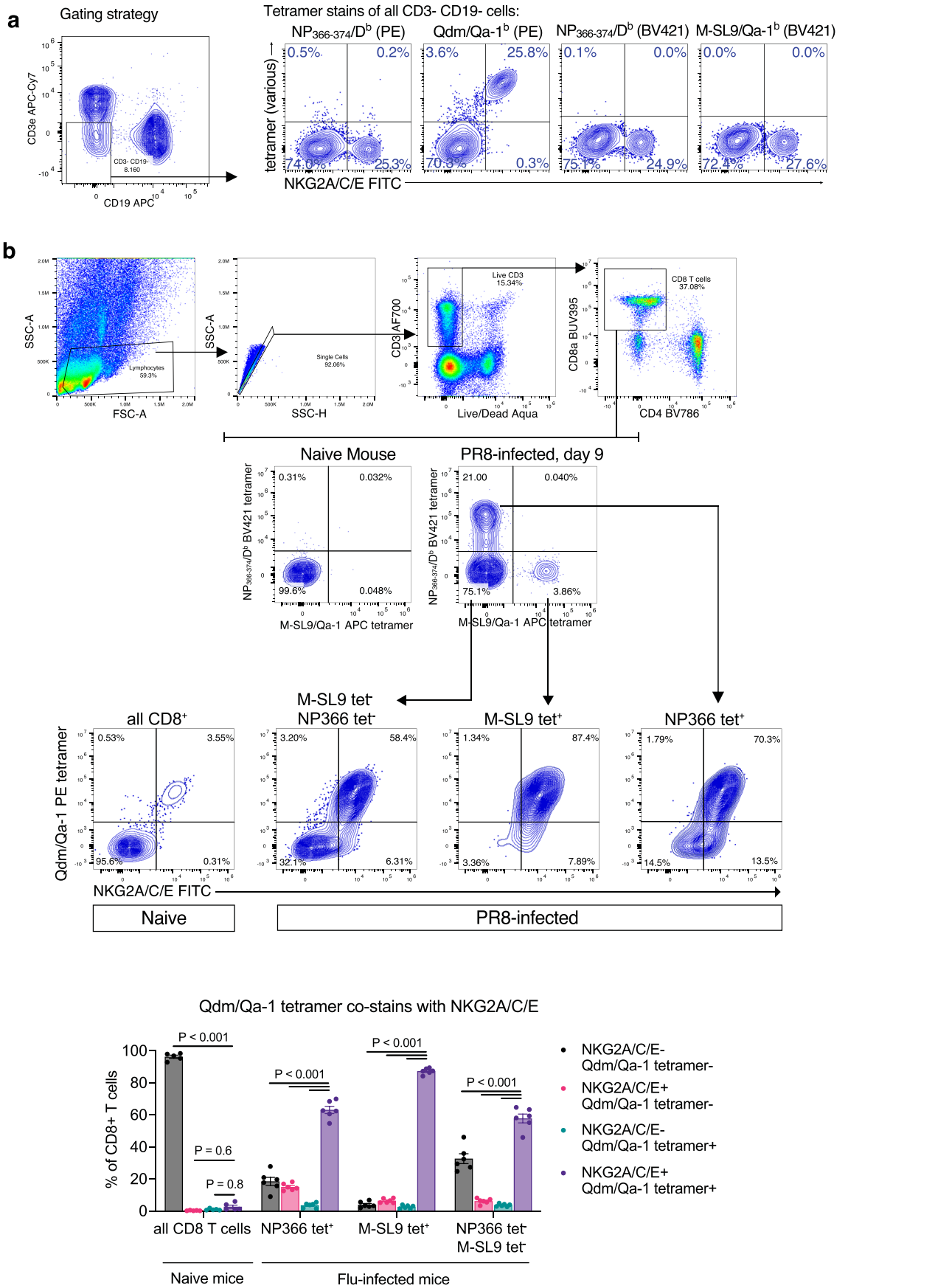
Extended Data Fig. 5 | Hybridoma clone B6.23 recognizes two forms of M-SL9 present in isolates of PR8. Amino acid sequences are shown for the originally identified M-SL9, present in pDZ PR8, and M-SL9-P, present in other PR8 isolates (for example GenBank V01099). B6-CIITA fibroblasts served as APCs and were co-cultured overnight with B6.23 cells in the presence of the indicated peptide concentrations. A sigmoidal curve was fit to the data points above (mean +/− s.d.), representative of three independent experiments. The geometric mean half-maximal effective concentration (EC_{50}) values across all three experiments were computed as 940 ng/ml for M-SL9 and 51 ng/ml for M-SL9-P.

https://doi.org/10.1038/s41590-023-01644-5



Extended Data Fig. 6 | Evidence supporting Qa-1 restriction of M-SL9. a, The MHC-Ia molecules H2-D^b and H2-K^b are not stabilized on RMA-S cells by M-SL9 peptide. RMA-S cells bearing unstable empty MHC-I molecules (due to TAP deficiency) were incubated in the presence of the indicated synthetic peptides, and surface expression of H2-D^b and H2-K^b was measured by flow cytometry. Mean fluorescence intensities of each stain were normalized to the negative control condition using $HA_{91\text{-}107}$, an I-A^b-binding epitope with no known binding to H2-D^b or H2-K^b, and shown as averages +/- s.d from 3 independent experiments. H2-D^b-binding $NP_{366\text{-}374}$ and H2-K^b-binding SIINFEKL were used as positive controls. **b-e**, Validation of HeLa cell lines and BMDCs showing Qa-1 restriction of M-SL9. **b**, The sufficiency of Qa-1^b expression for M-SL9 presentation to its cognate T hybridoma was confirmed using a HeLa cell line transduced with full-length, wild-type Qa-1^b and an untransduced parental HeLa cell line as a control. Bars are mean +/- s.e.m. from triplicate technical replicates, representative of 3 independent experiments, and *P*-values were calculated by Welch's t-test (two-tailed). **c**, Qa-1 expression on cell lines used in **b** was validated by flow cytometry. **d**, The expected staining pattern was confirmed for HeLa cell lines used in Fig. 2; these lines were transduced with retroviruses encoding chimeric MHC-Ib molecules containing the α3 domain (D3) from H2-D^b to allow efficient staining with the H2-D^b D3-specific mAb 28-14-8. **e**, The expected staining pattern was also confirmed for BMDCs used in Fig. 2.

https://doi.org/10.1038/s41590-023-01644-5

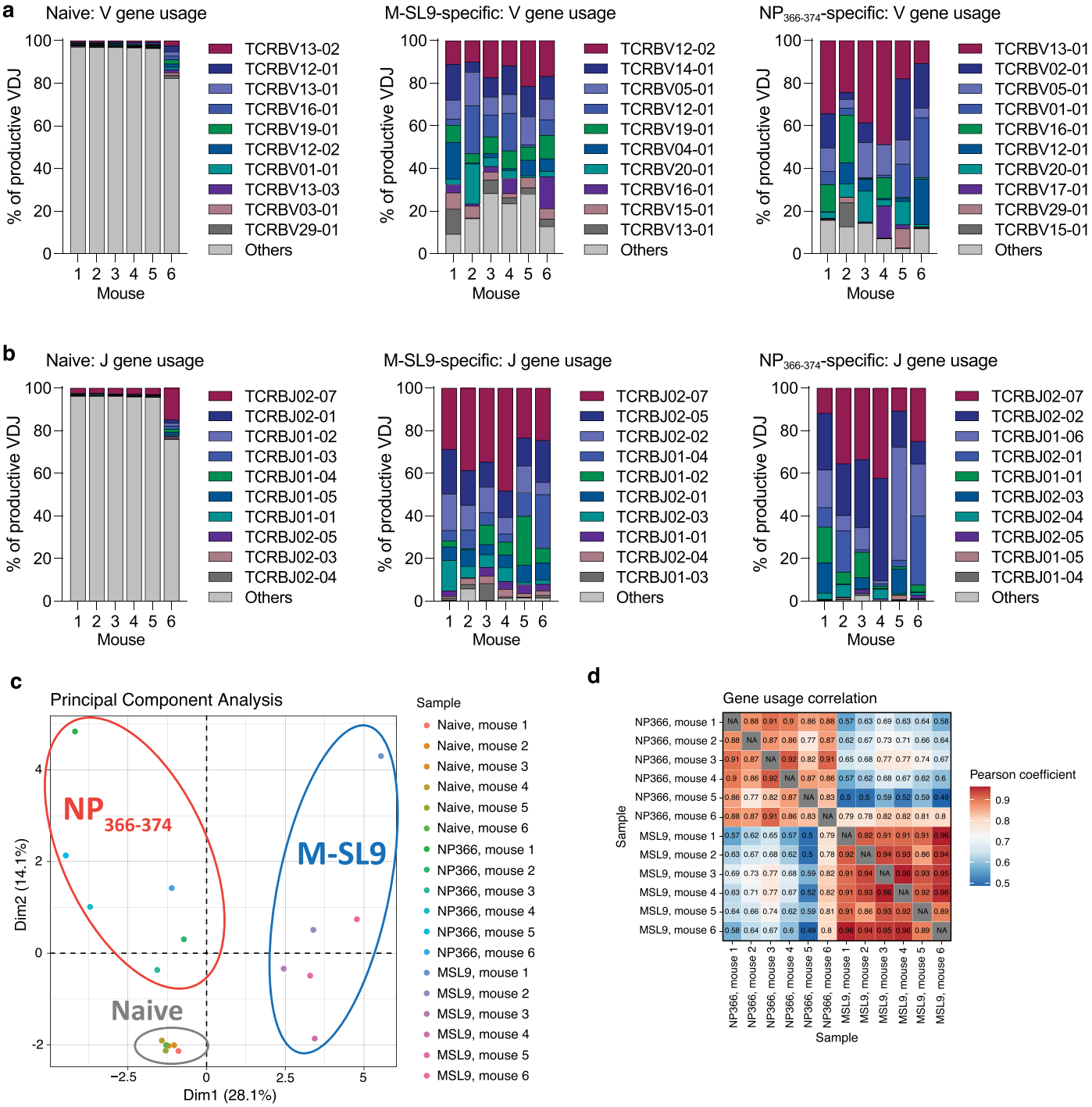


Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s41590-023-01644-5

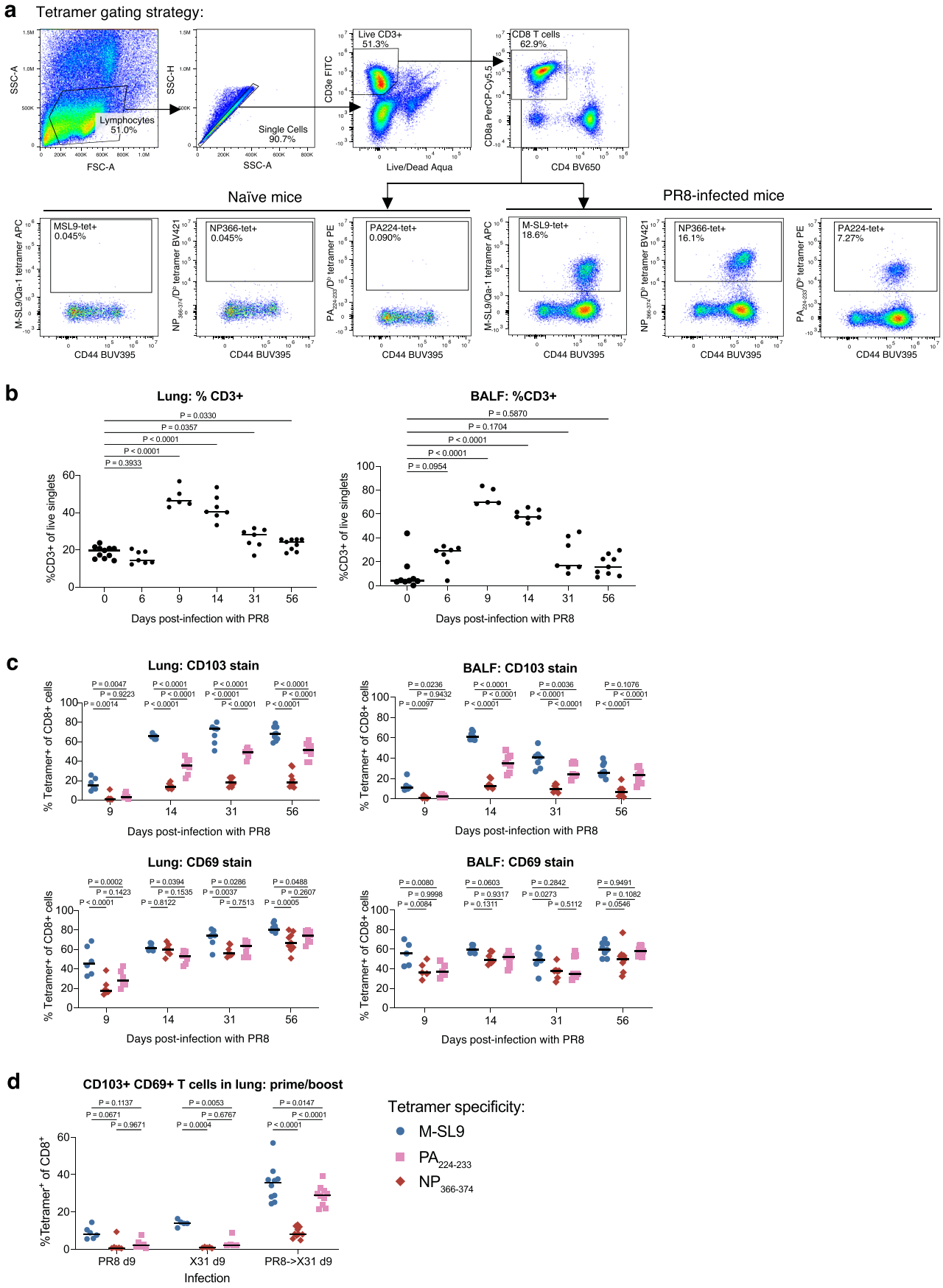
Extended Data Fig. 7 | Qdm/Qa-1^b tetramer co-stains with NKG2A/C/E. a, Lung lymphocytes from naïve C57Bl/6 mice were stained with an anti-NKG2A/C/E mAb and Qdm/Qa-1^b, M-SL9/Qa-1^b, and control $NP_{366-374}/D^b$ tetramers at 37 °C and gated on $CD3^-$ $CD19^-$ cells to interrogate natural killer (NK) cells. NK cells expressing NKG2A/C/E (the natural receptors for Qdm/Qa-1^b) were the only population that stained with Qdm/Qa-1^b tetramer, but neither these nor other NK cells stained with M-SL9/Qa-1^b tetramer. **b**, C57Bl/6 mice were intranasally infected with 160 FFU of PR8 and 9 days later lung lymphocytes were stained with anti-NKG2A/C/E and the indicated tetramers at 37 °C. Qdm/Qa-1^b tetramer generally stained PR8-induced CD8 T cells in a manner that was dependent on NKG2A/C/E but independent of TCR specificity. Flow plots are representative and show the gating strategy used, and bars show mean +/- s.e.m. for (a) N = 4 mice and (b) N = 5 to 6 mice per group across 2 independent experiments each. *P*-values are shown from two-way ANOVA with Dunnett's multiple comparisons test.

https://doi.org/10.1038/s41590-023-01644-5



Extended Data Fig. 8 | Analysis of TCRβ V and J gene usage in sorted CD8 T cell populations. a-d, CD8 T cell populations were sorted by FACS into three populations: naïve ($CD44^-$ $CD62L^+$), M-SL9-specific ($CD44^+$ M-SL9/Qa-1^b $tetramer^+$), and $NP_{366-374}$-specific ($CD44^+$ $NP_{366-374}$/H2-D^b $tetramer^+$). Genomic DNA was isolated, the VDJ region of recombined TCRβ-coding genes was sequenced, and gene usage was analyzed by (a,b) the immunoSEQ Analyzer and (c,d) Immunarch. **a-b**, The frequencies of the top 10 most-used (a) V genes and (b) J genes, on average across all mice, are shown as stacked bar graphs, where each bar represents one mouse. **c**, Principal component analysis (PCA) of the *Tcrb* V and J gene usage showing clustering by T cell population. **d**, Pearson correlation analysis of M-SL9- and $NP_{366-374}$-specific T cells showing greater correlation between mice within each T cell specificity rather than between specificities within each mouse. N = 6 mice, half males and half females.

https://doi.org/10.1038/s41590-023-01644-5

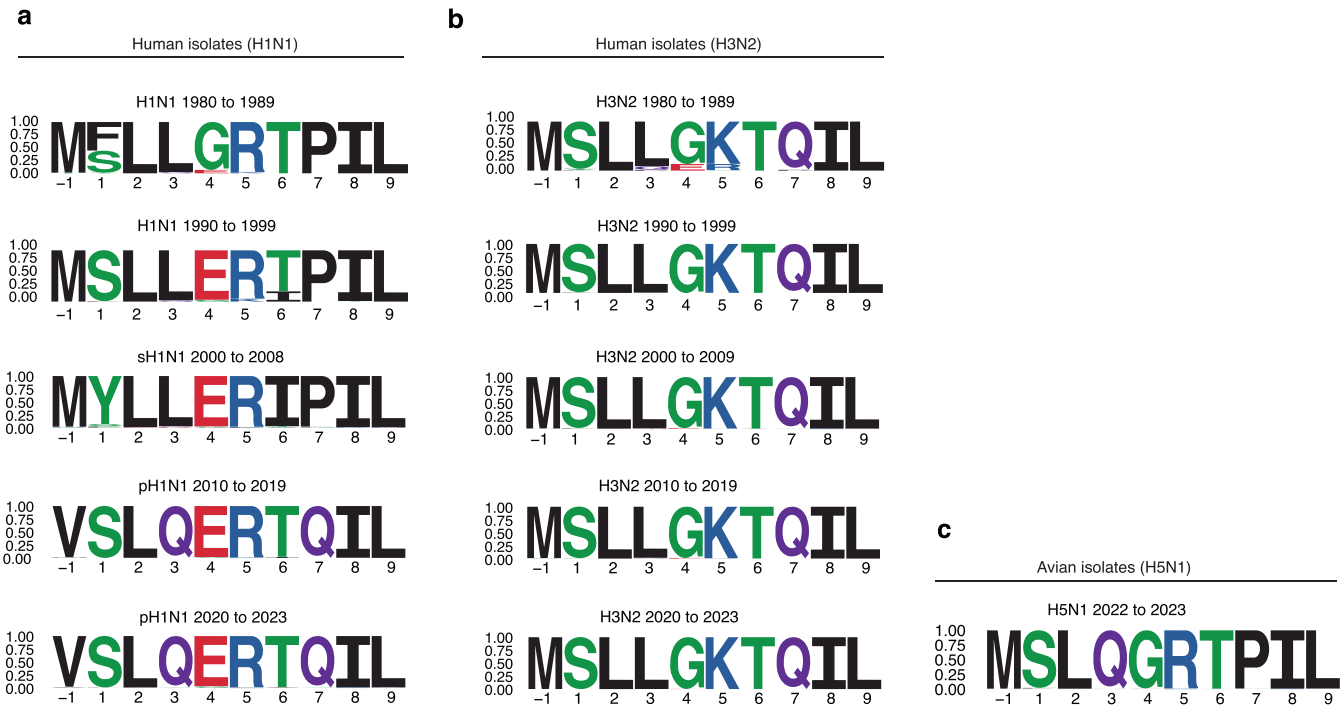


Extended Data Fig. 9 | See next page for caption.

https://doi.org/10.1038/s41590-023-01644-5

Extended Data Fig. 9 | Tracking CD69 and CD103 expression in PR8-infected mice. a-d, C57Bl/6 mice were intranasally infected with 160 FFU of PR8 and were euthanized at day 6 (N = 7), 9 (N = 5-6), 14, 31 (N = 7), or 56 (N = 9) to collect the indicated tissues/fluids. Uninfected mice were used as day 0 controls (N = 7-11). **a**, Gating strategy. **b**, Frequency of all $CD3^+$ T cells in lung and BALF over time, showing the lack of T cell infiltration in uninfected mice (plotted as day 0). Data points were omitted when there were <20 live singlet $CD3^+$ $CD8^+$ T cells collected in total. *P*-values are calculated from Brown-Forsythe and Welch one-way ANOVA with Dunnett T3 multiple comparisons test comparing each condition to day 0 controls. **c**, Frequency of $CD103^+$ and $CD69^+$ CD8 T cells (analyzed separately) in lung and BALF starting from the approximate peak of the T cell response on day 9. **d**, Frequency of $CD103^+$ $CD69^+$ double positive T_{RM} cells in lung at 9 days after PR8 only, X31 only, or PR8 prime and X31 boost. **c, d**, *P*-values were calculated by two-way ANOVA with Tukey's multiple comparisons test.

https://doi.org/10.1038/s41590-023-01644-5



M-SL9-L (molecular clone "pDZ PR8"): M SLQGRTLIL
M-SL9-P (viral isolate/swarm "Lopez PR8"): M SLQGRTPIL

Extended Data Fig. 10 | Sequence evolution and variation of the M-SL9 epitope and open reading frame in IAV strains over time since 1980. **a**–**c**, Sequence logo diagrams were produced from M-SL9-homologous sequences from (a) human H1N1 isolates, (b) human H3N2 isolates, and (c) H5N1 isolates from all avian species, downloaded between April and June 2023 for the indicated sample collection time periods. The BV-BRC database was used for sequences from 1980-1999, while the GISAID database was used for all others. Diagrams were created using the ggseqlogo package in R, and the y-axis units are the probability of each amino acid from 0 to 1. H1N1 sequences after 2009 correspond to the swine-origin pandemic H1N1 lineage, while H1N1 sequences prior to 2009 are from the earlier seasonal H1N1 lineage; sequences from 2009 were omitted to avoid ambiguity. Amino acids are numbered so that position 1 corresponds to the initial serine residue of the M-SL9 epitope, and the preceding residue was designated as position −1 and shown to assess the presence of an initiation codon. The two forms of M-SL9 encoded by PR8 isolates are shown at bottom; note that the avian H5N1 consensus sequence exactly matches the M-SL9-P amino acid sequence.

nature portfolio

Corresponding author(s): M.J. Hogan and L.C. Eisenlohr

Last updated by author(s): 9-1-23 M.J.H.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☑	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☑	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☑	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☑	A description of all covariates tested
☐	☑	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☑	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☑	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☑	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☑	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☑	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data collection is detailed in the Methods.
Data analysis	Data analysis was performed using Excel, GraphPad Prism, FlowJo, and R as detailed in the methods.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw MS data are deposited in the PRIDE database under the accession PXD045025, and TCRB sequences are in GitHub at https://github.com/mhogan240/NatImmuno2023

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Detailed in figure legends.
Data exclusions	IAV-infected mice with <2% granzyme B-positive CD8 T cells were excluded due to presumed lack of infection in t
Replication	All studies performed at least 2 independent times as detailed in the figure legends.
Randomization	Not performed.
Blinding	Not performed, since all measures were objective (e.g. flow cytometry-based).

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing	
Data exclusions	
Non-participation	
Randomization	

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing and spatial scale	
Data exclusions	
Reproducibility	
Randomization	
Blinding	

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	
Location	
Access & import/export	
Disturbance	

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☑ Antibodies
☐	☑ Eukaryotic cell lines
☑	☐ Palaeontology and archaeology
☐	☑ Animals and other organisms
☑	☐ Clinical data
☑	☐ Dual use research of concern
☑	☐ Plants

Methods

n/a	Involved in the study
☑	☐ ChIP-seq
☐	☑ Flow cytometry
☑	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibody clone names and other critical details are listed in the Methods section. Catalog numbers and lot numbers as available for key antibodies are as follows: CD3e APC/Cy7 BD 557596 lot 9142682, CD3e FITC Biolegend 100216 lot B263796, CD3e AF700 Biolegend 100204 lot B351638, CD4 BV786 BD 563727 lot 1187871, CD8 PerCP/Cy5.5 BD 551162 lot 2018402, CD103 BV605 Biolegend 121433 lot B372299, CD69 PE/Cy7 Biolegend 104512 lot B373784, IFN-gamma AF488 Biolegend 505813 lot B282839, TNF BV605 Biolegend 506329 or TNF BV650 Biolegend 506333 lot B374246, IL-2 PE Biolegend 503808 B234716, CD107a AF647 Biolegend 121610 lot B304301, granzyme B PacBlue Biolegend 515408 lot B301155, Perforin PE Biolegend 154306 or Perforin APC Biolegend 154304 lot B347816, and Qa-1 PE BD 566641 lot 8212935.
Validation	Antibodies were validated by detecting expecting staining patterns on positive/negative biological controls. For validation results see flow gating and example stains in Fig. 1c-e, Fig. 2 a-c, 5a, 8b-c, Extended Data Figs. 4a,e, 6c-e, 7, and 9a.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	293T obtained from ATCC, HeLa lines from Jensen and van Hall labs, others in-house.
Authentication	Cell lines were functionally validated in antigen presentation assays.
Mycoplasma contamination	All cell lines confirmed mycoplasma negative
Commonly misidentified lines (See ICLAC register)	None in this study.

Palaeontology and Archaeology

Specimen provenance	
Specimen deposition	
Dating methods	

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	Mus musculus, strains C57Bl/6 and BALB/c, aged 7-12 weeks at start of experiments.
Wild animals	None
Reporting on sex	Males and females were used in roughly equal proportions as detailed in the figure legends.
Field-collected samples	None
Ethics oversight	All animal procedures complied with all relevant ethical guidelines for animal testing and research, including institutional and AAALAC guidelines. Animal protocols were approved by the Institutional Animal Care and Use Committees (IACUC) at the Children's Hospital of Philadelphia and Brown University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	
Study protocol	
Data collection	
Outcomes	

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
X	☐	Public health
X	☐	National security
X	☐	Crops and/or livestock
X	☐	Ecosystems
X	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
X	☐	Demonstrate how to render a vaccine ineffective
X	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
X	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
X	☐	Increase transmissibility of a pathogen
X	☐	Alter the host range of a pathogen
X	☐	Enable evasion of diagnostic/detection modalities
X	☐	Enable the weaponization of a biological agent or toxin
X	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

Novel plant genotypes

Authentication

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

Files in database submission

Genome browser session
(e.g. UCSC)

Methodology

Replicates

Sequencing depth

Antibodies

Peak calling parameters

Data quality

Software

Flow Cytometry

Plots

Confirm that:

☑ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☑ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☑ All plots are contour plots with outliers or pseudocolor plots.

☑ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation: Spleens and lymph node (axillary, iliac, inguinal) were prepared by physical homogenization through a 70 um strainer, ACK lysis (for all spleens and lymph nodes as needed), and 40 um straining. Blood was prepared by ACK lysis two times and 40 um straining. Lungs were prepared by enzymatic digestion with collagenase D and DNase I followed by physical disruption with a GentleMACS machine, ACK lysis, and 40 um straining. BALF was collected by inserting a catheter through the trachea of a euthanized mouse, using a syringe to gently instill PBS EDTA to inflate the lung and then withdrawing the liquid for a total of 3 times; ACK lysis was performed as needed and samples were 40 um strained prior to flow staining. Further details are provided in the Methods.

Instrument: Cytoflex LX and Aurora

Software: FlowJo for analysis

Cell population abundance: In general, 200,000 events collected per sample to ensure populations of interest were sufficiently abu

Gating strategy: In general, samples were gated to maximize separation between negative and positive populations, aiming to achieve a background of approximately 0.1-0.5% across all negative samples in a consistent manner that is determined for each marker - see full gating strategies in Extended Data Figures 4a, 7b, and 9a. Preliminary FSC/SSC gating aimed to gate on all events that were not aggregates or small debris, since lymphocytes were distinguished later by CD3 and other marker expression.

☑ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Design specifications

Behavioral performance measures

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

(See Eklund et al. 2016)

Correction

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis