Character count: 62573 (everything with space)

Influenza A Virus Infection Induces Viral and Cellular DRiPs Encoded by Alternative Reading Frames

Damien J. Zanker[1*], Sara Oveissi[1*], David C. Tscharke[†], Mubing Duan[*], Siyuan Wan[*], Xiaomu Zhang[*], Kun Xiao[*], Nicole A. Mifsud[*], James Gibbs[§], Lenny Izzard[¶], Daniel Dlugolenski[¶], Pierre Faou[*], Karen L. Laurie[‖], Nathalie Vigneron[#], Ian G. Barr[‖], John Stambas[¶], Benoit van den Eynde[#], Jack R. Bennink[§], Jonathan W. Yewdell[§3] and Weisan Chen[*3]

[*]Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, VIC, 3086, Australia.
[†]Research School of Biology, The Australian National University Canberra ACT 0200 Australia.
[‡]Infection and Immunity Program, Monash Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia.
[§]Cellular Biology Section, Laboratory of Viral Diseases, NIAID, NIH, Bethesda, Maryland, 20892, USA.
[¶]School of Medicine, Deakin University, Victoria 3220, Australia.
[‖]WHO Collaborating Centre for Reference and Research on Influenza at the Peter Doherty Institute for Infection and Immunity, Parkville, Victoria 3010, Australia.
[#]Ludwig Institute for Cancer Research - Brussels branch, Avenue Hippocrate 74 B-1200 Brussels Belgium.

[1, 3]These authors contributed equally to this work

Running title: Influenza A virus infection increases DRiPs

Keywords: Influenza A virus, Defective Ribosomal Products (DRiP), alternative reading frames, antigen processing, antigen presentation, $CD8^+$ T cell, BALB/c mice, H-$2L^d$

[2]This project was partly supported by the NHMRC project grants 433608 and 542508; NHMRC Senior Research Fellowship 603104 to WC and the NHMRC program grant 567122. DZ was supported by an NHMRC Biomedical Postgraduate scholarship.

[3] To whom correspondence should be addressed: weisan.chen@latrobe.edu.au or JYEWDELL@niaid.nih.gov

Abstract

The importance of anti-viral $CD8^+$ T cell recognition of alternative reading frame (ARF) derived peptides is uncertain. Here, we describe an epitope (NS1-$ARF2_{1-8}$) present in a predicted 14-residue peptide encoded by the +1 register of NS1 mRNA in the influenza A virus (IAV). NS1-$ARF2_{1-8}$ elicits a robust, highly functional T_{CD8+} response in IAV infected BALB/c mice. NS1-$ARF2_{1-8}$ is presented from unspliced NS mRNA, likely from downstream initiation on a Met residue that comprises the P1 position of NS1-$ARF2_{1-8}$. Derived from a 14-residue peptide with no apparent biological function and negligible impacts on IAV infection, infectivity and pathogenicity, NS1-$ARF2_{1-8}$ provides a clear demonstration of how immunosurveillance exploits natural errors in protein translation to provide anti-viral immunity. We further show that IAV infection enhances a model cellular ARF translation, which potentially has important implications for virus-induced autoimmunity.

Key points

1. Discovery of ARF epitope, NS1-$ARF2_{1-8}$, encoded by NS1 mRNA of influenza A virus.
2. NS1-$ARF2_{1-8}$ elicits robust immunodominant T_{CD8+} response in infected BALB/c mice.
3. NS1-$ARF2_{1-8}$ precursor (14-residue peptide) has no detectable biological function.

Introduction

CD8$^+$ T cells (T_{CD8+}) eliminate virus infected cells by recognizing viral oligopeptides displayed by MHC class I molecules on the infected cell surface. The source of such peptides remains an area of intense interest. Typically, the rapid presentation of viral peptides starkly contrasts with the high stability of their ostensible source proteins (1, 2). This observation led to the defective ribosome hypothesis of antigen presentation, which after a number of updates, posits that errors in transcription, translation, and post-translational protein maturation result in the synthesis of rapidly degraded polypeptides that are exploited for immunosurveillance (3).

The contribution of translational errors to viral peptide generation via frame shifting, and non-canonical or downstream initiation has been elegantly determined in model systems (4-7), but the relative contribution of mis-translation to induction of anti-viral T_{CD8+} responses is unclear. Ironically, a systematic survey for downstream misinitiation of influenza A virus (IAV) DRiPs on Met residues led to the discovery of PB1-F2, at the time, the first IAV gene product discovered in decades (8). A number of immunogenic retroviral peptides of uncertain biological significance have been identified from putative alternative reading frames (9-11), but whether the peptides derive truly from ARFs vs. short open reading frames (ORFs) or random frame shifting is uncertain (the shortest known functional viral proteins is 45 residues (12), while 11 residue "proteins" are known to function in drosophila (13)).

IAV possesses a negative sense single stranded genome whose 8 segments encode 13 known functional proteins (14, 15), and other in-frame truncated proteins that are likely to be functional (14, 16). The synthesis of IAV mRNA in the nucleus enables segments 7 and 8 to encode M1 and NS1, respectively, and M2 and NS2 are created by the host nuclear RNA splicing machinery. PA-X is generated from the PA mRNA by frame-shifting (15).

The mouse IAV infection model has played a key role in deciphering the rules of CD8$^+$ T cell immunodominance (8, 17-22). In studying anti-IAV CD8$^+$ T cell responses in BALB/c mice, we observed an extremely robust response to a naturally processed peptide that we were unable to identify. After an arduous and extended hunt, we have finally caught our Moby Dick peptide, which provides the clearest example of a viral peptide generated from a natural ARF.

Materials & Methods

Mice

BALB/c mice were purchased from Walter Eliza Hall Institute of Medical Research (Kew, Melbourne, Australia). Mice were housed in SPF isolators. Experiments were performed with animals aged at 6-12 weeks and were conducted under the auspices of the Austin Health and La Trobe University Animal Ethics Committee and conformed to the National Health and Medical Research Council Australian code of practice for the care and use of animals for scientific purposes.

Peptide and antibodies

IAV peptides (please see Table S1) were synthesized by Mimotopes at >80% purity (Clayton, Melbourne). PE-labelled anti-IFN-γ, PE-Cy7-labelled anti-TNFα, FITC-labelled anti-CD107α and APC-labelled anti-CD8α were purchased from eBioscience (San Diego, CA, USA). For flow cytometry, antibodies were used at 1/300 dilution in PBS supplemented with 10% FCS. For IAV protein staining, cells were fixed with 1% Paraformaldehyde and stained in the presence of 0.4% Saponin with either anti-NS1 mAb (1A7), anti-NP mAb (HB65), anti-M1 mAb (M2-1C6) or anti-M2 mAb (O19) at 1/1000 dilution, followed by secondary 1/100 anti-mouse-FITC Ab.

Viruses and mouse immunization

IAV A/Puerto Rico/8/34 (H1N1) (PR8), A/Tasmania/2004/2009 (Pan09, H1N1) and NS1-GFP IAV were grown in 10-day embryonic chicken eggs and used as infectious allantoic fluid. The recombinant NS1-GFP IAV (23) was a kind gift from Dr Adolfo García-Sastre (Department of Microbiology Icahn School of Medicine at Mount Sinai, New York). Mice were infected with 100 pfu IAV intranasally (i.n.), or 10^7 pfu IAV intraperitoneally (i.p.).

Mutant virus and reverse genetics

Mutant and wild type IAVs were generated by plasmid-based reverse-genetics, following an adaptation of the method by Hoffmann et al (24). For the transfection, plasmid-clones of the 8 genomic segments were transfected into co-cultured MDCK (4×10^5 cells/well) and HEK293T cells (8×10^5 cells/well) with Fugene 6 (Promega, Australia) according to the manufacturer's protocol and incubated at 37°C**,** 5% CO_2 in a humidified incubator. Transfection medium was carefully removed 6 hours post-transfection and replaced by 1 ml of Easy Flu medium (Opti-MEM®I medium, containing 10U/ml penicillin and 100 µg/ml streptomycin). Following 30 hours incubation, 1 ml of Easy Flu medium containing 1 µg of TPCK-trypsin (TPCK Trypsin, Worthington biochemicals, New Jersey, USA) was added to the cells (total concentration of 0.5 µg/ml of TPCK-trypsin in the cell supernatant). Cytopathic effects (CPE) were obvious 48-72 hours following addition of TPCK-trypsin. Following CPE occurrence, supernatant was collected and clarified by centrifugation (5000 rpm for 5 minutes) and used for infecting 10-day embryonic chicken eggs (100 µl neat/egg) for generating a working viral stock.

In vitro viral infection

For recombinant vaccinia virus **(**rVV) infection, cells were resuspended in 0.1% BSA/PBS and infected with rVV as indicated in figure legend at 10 multiplicity of infection (MOI). For IAV infection, cells were resuspended in FCS-free, acidified RPMI medium and infected with IAV at 10 MOI. Cells were incubated in a 37°C water bath for 1 hour with gentle agitation every 15 mins. Infected cells were topped to 2ml with RF-10, and incubated for a further 4 hours at 37°C. Cells were then washed twice with PBS.

Cell line culture

All cells were cultured at 37°C**,** with 5% CO_2 in a humidified incubator. P815, D2SV, D2SV-transductants and HEK293T-transfectants were cultured in RPMI 1640 containing 10% FCS, 50 µM 2-ME, 2 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin (RF-10). Platinum-E retroviral-producing cell line (Cell Biolabs, USA) was cultured in DMEM medium

containing 10% FCS and above supplements, plus 1 µg/mL puromycin and 10 µg/mL blasticidin (DF-10). For reverse genetics, MDCK cells were maintained in MEM (Life Technologies, Australia) containing 10% FCS, 10U/ml penicillin and 100 µg/ml streptomycin (MEM-10) and HEK293T cells were cultured in Opti-MEM®I (Life Technologies, Australia) containing 5% FCS, 10U/ml penicillin and 100 µg/ml streptomycin. Cells were passaged and media were changed when cells became 80% confluent; cells were used during the exponential growth phase.

T_{CD8+} cultures

For generating polyspecific T_{CD8+} cultures, P815 were infected with PR8 (PR8-P815) at 10 MOI. The infected cells were then washed and irradiated for 10,000 Rads and used as APC to stimulate PR8-primed memory (>30 days) splenic cells. Repeatedly re-stimulated T_{CD8+} cultures were restimulated every 14 days with PR8-P815 cells. For generating peptide-specific T_{CD8+} cultures, 1/10 of memory splenic cells were pulsed with 1 nM peptide of interest for 60min. The cells were then washed with PBS and co-cultured with 9/10 of responder memory splenic cells (19).

Plasmids containing NS and its mutants

Generation of PMIM-NS-Full and PMIM-NS-mut: A 931 bp cDNA containing the wild type IAV PR8 (Mount Sinai) NS gene (898 bp) or its cDNA variant with mutated splicing elements (the 5', 3' splicing sites and the branch point) with 16 bp upstream and 17 bp downstream nucleotide sequences containing restriction enzyme sites was designed (sequence not shown), synthesised and cloned into pCU57 vector (GenScript, New Jersey, USA). NS-Full and NS-mut fragments were PCR amplified from the pCU57-NS-Full and pCU57-NS-mut template vectors respectively using the NS specific primer set (forward primer: tattAGATCTAGGGAGCAAAAGCAGGGTG and reverse primer: acgcCTCGAGTATTAGTAGAAACAAGGG). Both the NS PCR product (Full and mut) and PMIM vector (retroviral vector containing IRES-mCherry) were then subjected to *BglII*$^{5'}$ and *XhoI*$^{3'}$ double digestion to linearize PMIM and to generate matching sticky ends in both the vector and the PCR products for directional cloning the PMIM-NS-Full and PMIM-NS-mut plasmids.

Generation of PMIM-eGFP-NS^{intron}-WT and PMIM-eGFP-NS^{intron}-mut: A 1237 bp eGFP splicing reporter fragment was constructed by inserting either the 473 bp wild type or splicing mutant NS intronic sequence in between 30 and 31 bp of the 744 bp eGFP fragment flanked by 10 bp sequences at either side containing restriction enzyme sites (sequences not shown). The genes were synthesised (GenScript, New Jersey, USA) and cloned into pCU57 vector. eGFP-NS^{intron}-WT and PMIM-eGFP-NS^{intron}-mut fragments were PCR amplified using the eGFP specific primer set (forward primer: acgtGAATTCATGGTGAGCAAGG and reverse primer: atcgCTCGAGTTATGATCTAGAGTC). Both PCR products (eGFP-NS^{intron}-WT and PMIM-eGFP-NS^{intron}-mut) and PMIM vector were then subjected to *EcoRI*$^{5'}$ and *XhoI*$^{3'}$ double digestion to linearize PMIM and generate matching sticky ends in both the vector and the PCR products for directional cloning of the PMIM-eGFP-NS^{intron}-WT and PMIM-eGFP-NS^{intron}-mut plasmids.

Generation of C-terminal truncated NS1 fragments: A 693bp cDNA containing the wild type PR8 NS1 gene was used as a template for PCR amplification of full-length as well as six 3'-terminal truncated fragments subjected to TOPO cloning into the pcDNA™3.1/V5-His TOPO® vector (Life technologies, USA). NS1 truncation variants were PCR amplified using PR8 NS1 specific forward ($^{5'}$**CACC**ATGGATCCAAACACTGTGTCAAGCT$^{3'}$) primer coupling with various reverse primers: NS1/698 ($^{5'}$**GGATCC**TCAAACTTCTGACCTAATTGTTC$^{3'}$) for generation of full-length NS1 cDNA encoding NS1 1-230AA region, NS1/620 ($^{5'}$**GGATCC**GTAGAGTTTCAGAGACTCGAA$^{3'}$) for NS1 1-198AA, NS1/520 ($^{5'}$**GGATCC**GAAGGCAATGGTGAAATTTCG$^{3'}$) for NS1 1-164AA, NS1/420 ($^{5'}$**GGATCC**CTTTCAGTATGATGTTCTTAT$^{3'}$) for NS1 1-131AA, NS1/320 ($^{5'}$**GGATCC**CATTTCCTCAAGAGTCATGTC$^{3'}$) for NS1 1-98AA, NS1/220 ($^{5'}$**GGATCC**ACTATCTGCTTTCCAGCACGT$^{3'}$) for NS1 1-64AA and NS1/120

(5'**GGATCC**TCGGCGAAGCCGATCAAGGAA3') for NS1 1-38AA coding regions, respectively. These plasmids were then used to transfect HEK293T-L^d cells.

Generation of pHW2000-NS-ARF2$_{1-14}$ knockout:

A 931 bp cDNA of IAV PR8 (Mount Sinai) NS gene (898 bp) consisting of mutated NS-ARF2$_{1-14}$ variant (A**C**GC**T**CC**T**TT**T**CT**A**GA**C**CG**C**CT**C**CG**G**CG**G**GA**C**CA**A**AA**G**TC**T**CT**G**A with bold-underlined mutating nucleotide positions) consisting of 16 bp upstream and downstream nucleotide sequences encompassing *BsmbI*$^{5'/3'}$ restriction enzyme sites was designed (sequence not shown), synthesised and cloned into pCU57 vector (GenScript, New Jersey, USA). As incorporated mutations are silent in NS1 and NS2 open reading frames, their intended translational products remain wild-type sequences. The NS-ARF2$_{1-14}$ knockout fragments were agarose gel purified following *BsmbI* restriction enzyme digestion of pCU57-ARF2$_{1-14}$ knockout backbone vector and cloned into *BsmbI* digested and purified pHW2000 vector for construction of pHW2000-NS-ARF2$_{1-14}$ knockout vector which was then used in reverse genetics IAV production.

Transfection and transductions

One day before transfection, 2×10^6 retrovirus-producing cells were plated onto 10cm tissue culture dish in 10 ml DMEM-10 without antibiotics. When cells reached 70-80% confluence FuGENE® HD (Promega, USA) was used for transfecting Platinum-E cells with the bicistronic mammalian expression retroviral vector (PMIM) containing the gene of interest. 48 h later retrovirus-containing supernatant was harvested, filtered (0.45µm) and added to 3×10^5 D2SV target cells in 10 cm tissue culture dish. Polybrene was added into the retrovirus-containing supernatant at 8 µg/mL. 24 hrs later, the medium was replaced and target cells were cultured for three more days. Transduced cells were then analysed for mCherry reporter expression and were sorted using FACSAriaIII™ flow cytometer (Becton Dickinson, USA).

***In vitro* activation of Ag-specific T_{CD8+} assessed by Cr^{51} release**

^{51}Cr-release assays were performed as described (25). Data are expressed as % specific lysis = (T_{CD8+} induced lysis – spontaneous release)/(release by detergent - spontaneous release) x 100%.

***In vitro* activation of Ag-specific T_{CD8+} and enumeration by intracellular cytokine staining (ICS)**

For *ex vivo* isolates, BAL wash and splenic cells were collected in RF-10. The Ag-specific T_{CD8+} were enumerated using ICS for the production of IFN-γ after being stimulated with 1 µM of antigenic peptide in the presence of 10 µg/ml BFA for 5 hours at 37^{o}C unless otherwise stated (22). For BFA kinetics assay, P815 cells were infected with 10 MOI of PR8 for 1 hour at 37°C, washed and then added to monospecificity T_{CD8+} cultures, with BFA added at various time points. Following 4 hours of exposure to BFA, T_{CD8+} were transferred onto ice, fixed with 1% PFA and stained in ICS for IFN-γ in the presence of 0.4% Saponin (26).

Peptide extraction and RP-HPLC fractionation

The method was detailed previously (25). Briefly, 5 x 10^8 cultured cells were infected at 10 MOI with PR8 for 5 hours. Cells were washed, resuspended in TFA/H_2O, homogenized and further sonicated. Peptide-containing supernatant was collected after ultracentrifugation and passed through a 3K cut-off filter. Samples were then dehydrated to <400µl using a speed-vac and the extracted peptides were then fractionated on a RP-HPLC (Agilent 1100). Eluted fractions were collected between 10.0-38.8 mins, at 6 second intervals. Elution times of synthetic peptides were determined under the same running condition after sample collection. To detect naturally presented peptides in the HPLC fractions, P815 cells were low-temperature induced overnight at 26^{o}C, pulsed (10^5) with either synthetic peptide at log fold dilutions or fractions for 1 hour at 26^{o}C in RPMI. 10^5 T_{CD8+} were then added to the pulsed P815 cells in RF-10 containing 10 µg/ml BFA for a further 5 hours and antigen-specific T_{CD8+} activation enumerated by ICS.

RNA extraction and sequencing

Cells were homogenized using the QIAshredder (Qiagen). Total RNA was extracted from 3 x 10^6 cells per sample using the Qiagen RNeasy Mini Kit according to manufacturer instructions, with on-column DNase I digestion incorporated. Following extraction, RNA was quality controlled using the Agilent 2100 bioanalyzer and samples with a RIN score above 7 used for sequencing. Following mRNA purification with Poly-T oligo-attached magnetic beads, strand-specific cDNA libraries were prepared using the NEB Next Ultra RNA library preparation kit according to manufacture instructions and Illumina sequencing was conducted on the Hiseq4000 by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China) and 150 bp paired-end reads were generated.

RNAseq bioinformatics analysis

Quality control was first performed on raw data (FastQ files) using the FastQC software and any reads containing adapter sequences trimmed using TrimGalore. Reads were then aligned with HISAT2 (27) to a modified reference genome comprising the GRCm38.p5 genome and 12 wild-type PR8 influenza virus gene segments (PB1 NC_002021.1, PB2 NC_002023.1, PB1-F2 NC_002021.1, PA NC_002022.1, PAx NC_002022.1, HA NC_002017.1, NP NC_002019.1, NA NC_002018.1, M1 NC_002016.1, M2 NC_002016.1, NS1 NC_002020.1:27-719 and NEP NC_002020.1:27-864). Gene counts were quantified with featureCount and strand-specific anti-sense strands were included in the analysis. For bioinformatic analysis, genes with valid expression were defined as having logCPM (CPM: count per million) > 0.3 in at least 2 different replicates. After multi-dimensional scaling analysis, sample KO infected 3 was removed as an outlier. Differentially expressed genes were calculated using DESEQ2 (28) and by setting logFC ≥ 1 or ≤ -1, p-value ≤ 0.01, FDR ≤ 0.01.

LC-MS/MS

P815 cells were infected with 10 MOI wtPR8 or $ARF2_{1-8}$ KO PR8 IAV for 5 hours, washed, stained and cell sorted for IAV HA expression and 3 x 10^6 sorted cells were dried by vacuum centrifugation. Cell pellets were resuspended in digestion buffer (8 M urea, 50 mM ammonium bicarbonate, 10 mM DTT) before incubation for 5 hours at 25°C. 55 mM iodoacetamide was then added to alkylate thiol groups at 20°C for 35 minutes in the dark. The alkylated preparation was diluted to 1 M urea with 25 mM ammonium bicarbonate (pH 8.5) before sequencing grade trypsin (Promega) was added to 5 µM final concentration. Digests were performed overnight at 37°C. The digests were acidified with 1% (v/v) trifluoroacetic acid (TFA) and the peptides desalted on SDB-XC (Empore) StageTips. Peptides were modified by stable isotope dimethyl labelling for quantitative proteomics. After mixing the peptides 1:1 the samples were fractionated off-line by high pH reversed-phase fractionation on 24 fractions. Peptides from each fraction were reconstituted in 0.1% TFA and 2% acetonitrile (ACN) and loaded onto C_{18} PepMap 100 µm ID × 2 cm trapping column (Thermo-Fisher Scientific) at 5 µl/min for 6 min, and washed for 6 minutes before switching the pre-column in line with the analytical column (Vydac MS C_{18}, 3 µm, 300 Å and 75 µm ID × 25 cm, Grace Pty. Ltd.). The separation of peptides was performed at 300 nl/min using a non-linear ACN gradient of buffer A (0.1% formic acid, 2% ACN) and buffer B (0.1% formic acid, 80% ACN), starting at 5% buffer B to 55% over 60 minutes. Data were collected on an Orbitrap Elite (Thermo-Fisher Scientific) in Data Dependent Acquisition mode using m/z 300–1500 as MS scan range, CID MS/MS spectra were collected for the 20 most intense ions per MS scan. Dynamic exclusion parameters were set as follows: repeat count 1, duration 90 s, the exclusion list size was set at 500 with early expiration disabled. Other instrument parameters for the Orbitrap were: MS scan at 120 000 resolution, maximum injection time 50 ms, AGC target 1 × 10^6, CID at 35% energy for a maximum injection time of 150 ms with AGT target of 5000. The Orbitrap Elite was operated in dual analyser mode with the Orbitrap analyser being used for MS and the linear trap being used for MS/MS.

LC-MS/MS data analysis and statistics

Identification and isotopic quantification of proteins were performed on raw output files from LC-ESI-MS/MS using MaxQuant (Version 1.5.1.6; Cox and Mann, 2008) together with its built-in search engine Andromeda. Uniprot using the Mouse Uniprot FASTA database (September 2015) together with common contaminants was used for this analysis. Carbamidomethylation of cysteines was set as a fixed modification, while acetylation of protein N-termini and methionine oxidation were included as variable modifications and dimethLys0, dimethLys4, dimethNterm0 and dimethNterm4 were included as labels. Parent mass tolerance was set to 5ppm (after refinement by MaxQuant) and fragment mass tolerance to 0.5 Da. Trypsin was set as the digestion enzyme with up to two missed cleavages allowed. Prior to searching, MS/MS spectra were filtered by retaining only the top eight peaks per 100Da. The match between runs feature of MaxQuant was used to transfer peptide identifications from one run to another based on retention time and mass to charge ratio. Both peptide and protein identifications were reported at a false discovery rate (FDR) of 1%.

The normalized protein group ratios values obtained from MaxQuant were then used to perform differential expression analysis with the limma package in R. A moderated t-statistic was calculated for each protein group in order to test for differences in expression between samples. Resulting p-values were corrected for multiple testing using the method of Benjamini and Hochberg et al. (1995), which effectively converts to a false discovery rate. Missing values were excluded from the analysis with degrees of freedom adjusted accordingly.

Results

Repeated in vitro T cell restimulation using IAV-infected P815 cells reveals a novel immunodominant T cell response

We previously reported $NP_{147-155}$ as the most immunodominant of 5 immunodominant IAV peptides recognized by $CD8^+$ T cells from IAV infected BALB/c mice (22). Interestingly, when we used IAV-infected P815 cells (PR8-P815) as antigen presenting cells (APC) to maintain a polyspecific T_{CD8+} line, the line gradually lost specificity for each of the 5 peptides, yet remained highly specific for IAV (Figure 1A). What did these T cells recognize?

To identify this potentially novel IAV epitope, we again raised a polyspecific T_{CD8+} line in the same manner. After the first round of stimulation, we determined recognition of 10 potential IAV gene products by infecting L929 cells ($H\text{-}2^k$) expressing individual $H\text{-}2^d$ allomorphs with recombinant vaccinia viruses (rVVs) expressing individual IAV mRNA (as a strictly cytoplasmic virus, rVV-encoded mRNAs are not spliced). HA-, NP-, and PA- specific responses were stimulated by L929-K^d cells, most likely representing the previously identified $HA_{518-526}$-, $NP_{147-155}$-, and PA-specific T cells respectively (18). A modest response to rVV-PB2 infected L929-D^d cells was observed (Figure 1B), as originally reported (29, 30) and are likely $PB2_{289-297}$-specific (21).

Critically, polyspecific T_{CD8+} cells also recognized L929-L^d cells infected with rVV-NS1, indicating the presence of novel NS1 epitope(s) (Figure 1B). The L^d-restricted recognition of rVV-encoded NS1 was confirmed by a repeatedly PR8-P815 re-stimulated T_{CD8+} line using $H\text{-}2L^d$ expressing HEK293T cells infected with PR8 or rVV-NS1 (Figure 1C).

Difficulties in identifying the NS1 L^d-restricted epitope

To identify the novel NS1 T cell epitope(s), we initially cloned six truncated NS1 gene fragments into the pcDNA™3.1/V5-His TOPO® vector to express NS1/1-38, NS1/1-64, NS1/1-98, NS1/1-131, NS1/1-164, NS1/1-198, and full-length NS1/1-230. Following transfection of HEK293T-L^d cells (31), none of the fragments displayed antigenicity (Supplementary Figure 1A).

As an alternative, we turned to $H\text{-}2L^d$ peptide binding motif prediction of potentially antigenic peptides. Accordingly, we selected peptides with a position 2 Serine (S) or Proline (P), and position 8, 9, 10, 11 Isoleucine (I), Phenylalanine (F), Valine (V), Leucine (L), Arginine (R) and Aspartic Acid (D). Of the 17 peptides tested, none demonstrated antigenicity using a repeatedly PR8-P815 re-stimulated T_{CD8+} line (Supplementary Figure 1B).

Many well-characterized antigenic peptides do not match predictive algorithms (32). Taking a broader approach, we synthesized a set of 13mer peptides with 10 overlapping AA to create a peptide panel spanning all possible 11mers and shorter linear epitopes, and likely all potential 12mer epitopes, as most T_{CD8+} recognize cognate peptides missing a single residue at either end when the peptide is provided at high concentration (µM level). Using a restimulated polyspecific T_{CD8+} line, none of the 76 13mer peptides were antigenic at even µM concentration (Supplementary Figure 1C).

Although MHC class I ligands are typically 9~10 AA and rarely longer than 12 AA, exceptions are known, e.g. 13mers from EBV BZLF1 antigen (33) or NY-ESO-1 (34), and a 14mer from M-CSF (35). To exclude the possibility of an unusually long epitope being recognized by our T_{CD8+} line, we screened a set of 18mer peptides with 12 AA overlapping. Once again, failure ensued (Supplementary Figure 1D).

These findings were consistent with the idea that the L^d-restricted T_{CD8+} recognize an unusual IAV peptide.

Identification of the novel epitope by alternative reading frame (ARF) peptides

Failing to find the epitope in the standard NS1 reading frame, we searched for predicted L^d-binding peptides in the +1 or +2 reading frames downstream of potential Met initiating residues. This revealed two peptides that we designate as NS1-ARF1 (MSANELQTK) and NS1-ARF2 (MPHSLIGFAEI). Testing synthetic peptides corresponding to the sequences, we found that

synthetic NS1ARF2 activated a major population of the PR8-P815 stimulated polyclonal T_{CD8+} (Figure 1D).

Within $ARF2_{1-11}$, coincidentally, is an 8-mer peptide that matches the H-$2L^d$ binding motif xPxxxxxF. We compared the abilities of synthetic $ARF2_{1-8}$ *vs.* $ARF2_{1-11}$ to generate T_{CD8+} lines from splenocytes derived from PR8 primed-BALB/c mice (Figure 1E, F). Interestingly, each peptide could elicit T_{CD8+} that greatly preferred their activating peptide compared to truncated versions, suggesting that each is naturally immunogenic in the context of PR8 infection *in vivo*.

To determine the extent to which cells naturally present $ARF2_{1-8}$ *vs.* $ARF2_{1-11}$ we HPLC fractionated acid soluble peptides from PR8 infected cells and assayed fractions for antigenicity against T_{CD8+} lines (25). Using an IAV polyspecific T_{CD8+} line, fractions eluting at 18.9 min, 21.2 min, 23.9 min, 24.3 min and 25.0 min, exhibited antigenic activity (Figure 2A-i). Using monospecific T_{CD8+} restimulated with synthetic peptides, it was clear that the two major peaks at 18.9 and 24.9 min are recognized by $NP_{147-155}$ and $ARF2_{1-8}$ respectively (Figure 2A-ii, iii). Corroborating this finding, this correlates perfectly with the elution of the cognate synthetic peptides. We fail to obtain evidence for natural presentation for $ARF2_{1-11}$ (Figure 2A-iv), which is expected to elute at 26.6 min based on the behavior of the synthetic version. This is mostly likely due to inefficient $ARF2_{1-11}$ epitope presentation by PR8-P815 cells.

Titration of peak antigenic fractions (Figure 2B) revealed that $ARF2_{1-8}$ is generated at least 3-fold more abundantly than $NP_{147-155}$, using the synthetic peptides as standards. We corroborated this finding with peptide generation kinetics, using brefeldin A to block further transport of class I – peptide complexes to the surface of PR8 infected P815 cells in conjunction with T_{CD8+} lines raised to individual peptides (26). As the two T_{CD8+} lines clearly had similar avidity shown by their cognate peptide titration (Figure 2C), $ARF2_{1-8}$ is presented approximately 2~3 fold faster than $NP_{147-155}$ (Figure 2D), which is consistent with the results of peptide elution and fraction titration results (Figure 2B) and likely explaining why multiple in vitro restimulated T_{CD8+} lines become $ARF2_{1-8}$-specific.

Highly functional $ARF2_{1-8}$-specific T_{CD8+} are a major component of BALB/c mice anti-IAV cellular immunity

Having defined NS1-$ARF2_{1-8}$ as the major peptide recognized by L^d restricted IAV-specific T_{CD8+}, we next used the synthetic peptide to measure the magnitude of the *in vivo* response 7 days post intraperitoneal infection. In mouse groups infected with IAV doses lower than $5x10^7$ pfu, NS1-$ARF2_{1-8}$ T_{CD8+} responses were slightly smaller than that of T_{CD8+} specific to the immunodominant $NP_{147-155}$. Remarkably, NS1-$ARF2_{1-8}$ response ascended to the α-position in the immunodominance hierarchy at higher virus doses (Figure 3A). In mice 10 d post intranasal infection with 100 pfu of IAV, in both spleen and lungs, NS1-$ARF2_{1-8}$ specific cells rank second in the immunodominance hierarchy to $NP_{147-155}$ (Figure 3B, C).

To be certain that recognition of NS1-$ARF2_{1-8}$ is a *bona fide* measure of its *in vivo* immunogenicity, we identified a 2009 pandemic H1N1 virus, A/Tasmania/09 (Pan09) (H1N1), possessing a F➔S substitution in the COOH-terminal anchor residue of NS1-$ARF2_{1-8}$ that obviates high affinity binding to L^d. Pan09 induced a vigorous response to all of the major epitopes except NS1-$ARF2_{1-8}$, confirming its status as a naturally processed epitope *in vivo* (Figure 3B, C).

We next compared the effector function of T_{CD8+} specific for NS1-$ARF2_{1-8}$ *vs.* $NP_{147-155}$ after their *ex vivo* activation with cognate peptide. This revealed that T_{CD8+} populations expressed similar amounts of effector cytokines IFNγ and TNFα, and were equally cytolytic as judged by expression of the degranulation marker, CD107a (Figure 3D).

$ARF2_{1-8}$ is generated as an ARF peptide from NS-mRNA in a splicing independent manner

IAV generates NS2 via splicing of the NS mRNA. As the NS1-$ARF2_{1-8}$ epitope is encoded within the region that is normally spliced out to generate NS2-mRNA, we wanted to know whether the epitope is translated from NS1-mRNA as an ARF peptide or translated from the spliced off intronic mRNA. Since VV mRNAs are synthesized in the cytoplasm and do not traffic to the

nucleus, rVV-expressed NS1 is not spliced. Consequently, its recognition by NS1-$ARF2_{1-8}$ specific-T_{CD8+} (Figure 1D), demonstrates that the peptide can be made from the NS1 mRNA. This does not, however, exclude a role for splicing in generating the peptide in IAV infected cells, particularly given the evidence for nuclear translation as a source of peptides from IAV (36) and from intronic regions of plasmid encoded genes (37).

To examine the role of mRNA splicing in NS1-$ARF2_{1-8}$ generation, we cloned the PR8 NS gene into the retroviral vector PMIM as wild-type sequence (NS-full) or modified to abrogate splicing (NS-mut) (Figure 4A-ii). D2SV cells transduced with either NS-full or NS-mut equally activated an $ARF2_{1-8}$-specific T_{CD8+} line, demonstrating that mRNA splicing does not enhance ARF epitope generation (Figure 4B). We noted that in our NS1 constructs $ARF2_{1-8}$ epitope presentation is clearly weaker compared to IAV infected cells. The likely explanation is that NS1 expression levels are lower from our constructs compared to that from IAV infection (Supplementary Figure 2).

We further examined the contribution of NS1 splicing to peptide generation using NS1-GFP IAV (23), a recombinant IAV whose NS segment is modified to express NS1-GFP, followed by a FMDV 2A ribosome stop and go site (38) that releases NS2 (also called NEP) as it is synthesized. Two synonymous mutations are also present to prevent NS mRNA splicing. P815 cells infected with NS1-GFP IAV efficiently presented $ARF2_{1-8}$ epitope (Figure 4B).

Together, these results demonstrate that the $ARF2_{1-8}$ epitope is generated through the translation of the ARF2 of the NS1 mRNA, and that mRNA splicing plays, at most, a minor role in generating this peptide.

$ARF2_{1-8}$ epitope is also presented when the intronic sequence from NS is cloned into eGFP gene

Given the location of $ARF2_{1-8}$ epitope in the NS gene intron, we investigated whether its synthesis is influenced by regulatory sequences outside the intronic sequence by transplanting the NS intronic sequence into the eGFP gene (Figure 4A-iii). The construct was then used to transduce D2SV cells and transduced cells were enriched based on mCherry expression. Interestingly, the $ARF2_{1-8}$ epitope was generated more efficiently by eGFP-NS^{intron} transduced cells than those transduced with NS-full (Figure 4B), showing that $ARF2_{1-8}$ is generated in its local translational context. Again, when the splicing sites were eliminated in the eGFP-NS^{intron}, the generation of $ARF2_{1-8}$ epitope was not affected.

$ARF2_{1-8}$ epitope presentation from "autoantigen" eGFP-NS^{intron} is enhanced by IAV infection

Although the $ARF2_{1-8}$ epitope was generated from D2SV cells transfected with full-length NS gene or transduced with eGFP-NS^{intron}, in which either the NS or NS-intronic sequence is essentially converted to a self-gene by chromosomal integration, its presentation was clearly less efficient than in the context of IAV infection. We therefore questioned whether such ARF DRiP generation is enhanced with IAV infection.

To explore this, we generated the above-mentioned transfectant and transductant in the P815 cell line expressing either full-length NS or eGFP-NS^{intron} and then superinfected these cells with Pan09 IAV. As this IAV strain does not generate $ARF2_{1-8}$ from its own genome, all detected $ARF2_{1-8}$ is derived from the "autoantigens" expressed in these cell lines. As shown in Figure 4C, Pan09 IAV infection clearly enhanced $ARF2_{1-8}$ presentation, especially from cells that expressed eGFP-NS^{intron}. Again such enhanced presentation of endogenous DRiPs did not require mRNA-splicing as the epitope presentation was equally efficient in the splicing mutants. The infection by Pan09 IAV was as efficient as those by PR8 as shown by equal $NP_{147-155}$ epitope presentation (Figure 4D).

$ARF2_{1-8}$ epitope is a bona fide DRiP as no biological function was detectable for $ARF2_{1-14}$

As very short peptides can have significant biological functions (39), to formally exclude the possibility that $ARF2_{1-14}$ functions in IAV infection we generated $ARF2_{1-14}$ knockout IAV viruses using the reverse genetics approach (24). We then compared the mutant IAV to a control *wt* recombinant generated in parallel for *in vitro* infectivity and ability to trigger immunopathology and cellular immunity in two mouse strains. First, we infected the mouse lung epithelial adenoma cell line LA-4 (40) and measured the production of NP, M1, M2 and NS1 via flow cytometry with the appropriated mAbs. As shown in Figure 5A (i – iv), LA-4 cells infected with *wt* or the ARF2 mutant virus generated equal amounts of each of the proteins measured, indicating that the small ARF is not required for viral mRNA synthesis or translation. Further, when we infected P815 cells with *wt* vs mutant viruses and characterized the transcriptome and proteome, we detected only very minor differences. Only 12 out of 13167 genes and 2 out of 4419 proteins were differentially expressed. Moreover, the differentially detected genes and proteins did not overlap between replications, and no biological signaling pathways were over-represented (Supplementary Figure 3A and B).

Second, when we infected C57BL/6 and BALB/c mice intranasally with 100 pfu with *wt* or mutant viruses, mice showed identical weight-loss profiles, indicating that short ARF is not required for viral pathogenesis.

Finally, we assessed the cellular immune responses following intranasal infection of mice. In C57BL/6 mice, we found a nearly identical $CD8^+$ T cell response to 7 different peptides in spleen, with relatively minor variation in the BAL response to some peptides (Figure 5B, D, and F). Similarly, BALB/c mice, with the obvious exception of the $ARF2_{1-8}$ response, there were only minor differences in splenic or BAL responses to 6 other peptides (Figure 5C, E, and G). Higher infecting doses of 500pfu of $ARF2_{1-14}$ knockout virus did not demonstrate any alteration in lethality or the resulting immune response (data not shown).

Taken together, our findings do not support an important biological role for $ARF2_{1-14}$ in IAV infection of cultured cells or infected mice.

Discussion

Here we show that BALB/c mice mount a robust T_{CD8+} response to NS1-$ARF2_{1-8,}$ which under some conditions ascends to the top of the immunodominance hierarchy. We previously systematically searched for potential ARF epitopes presented by H-$2K^b$, D^b or K^d. Although 9 of the 35 synthetic peptides screened efficiently stimulated T_{CD8+} elicited with the autologous peptide, only a single peptide could stimulate or induce IAV-specific T_{CD8+} (41). Ironically, the corresponding 87 residue ARF is an ORF encoding PB1-F2 (8), which has proven to exert myriad functions contributing to viral fitness (42, 43). This highlights a critical issue in identifying ARF encoded peptides: how can we be certain that the gene product is not a short functional ORF? Indeed, it is not certain that the ARF peptides identified to date for HIV, HBV, and HCMV are truly non-functional (9, 44-47). In each of these published examples the ARF/ORF is greater than 30 residues, which is of sufficient length to form a membrane spanning oligomer, or at the very least, associate with a viral or cellular protein to modulate its function.

NS1-ARF2 is but 14 residues. While not impossible, it is unlikely that such a short peptide could be functional. A drosophila 11-mer seems to have a function (13) and plant 18- and 20-mer microproteins were recently described to possess biological activity (48, 49). The shortest viral gene products with known biological activity are greater than 40 residues. Out of >5000 NS1 sequences available online, only one (A/chicken/Jiangsu/JT34/2011(H9N2)) lacks the start codon of the $ARF2_{1-8}$ epitope. All others have sequences highly similar (or identical) to PR8 or Pan09. Interpreting this sequence conservation is muddied by the high conservation of the corresponding

overlapping NS1 sequence, which encodes residues 30~43 of the RNA-binding domain critical for sequestering IAV dsRNA during its replication (50, 51). To formally exclude the possibility that NS1-ARF2$_{1-14}$ might be an intended viral product with biological function, we generated NS1-ARF2$_{1-14}$ deletion mutant and its control wildtype IAV virus by reverse genetics. In our parallel *in vitro* and *in vivo* and systems biology experiments, we were not able to detect any function or influence of NS1-ARF2$_{1-14}$ during IAV infection. We therefore conclude that it is likely that NS1-ARF2$_{1-14}$ is a *bona fide* viral DRiP.

While 14 residues may be too short to generate a functional gene product, we have not completely ruled out that NS1-ARF2 is synthesized as something more than a 14-mer. Our evidence conclusively demonstrates NS1-ARF2$_{1-8}$ being generated from NS1 mRNA and not a splice variant. This follows from NS1-ARF2$_{1-8}$ specific T_{CD8+} recognition of rVV-infected cells expressing NS1 mRNA, or cells expressing unspliceable forms of NS1 mRNA expressed by IAV, or from a transgene. IAV generate at least one functional frame shifted protein (PA-X) (15), but there is very little room upstream of NS1-ARF2 for frame-shifting, with only a Val residue before encountering a stop codon, and no known frame shift abetting sequence. Although it seems unlikely, it is possible that the NS1-ARF2 is extended by read through of the stop codon or by shifting back into the NS1 frame between residue 8 and the stop codon after residue 14.

Our findings clearly demonstrate the potential importance of unusual translation events in generating biologically important viral peptides, extending the recent exciting evidence for the potential involvement of nuclear translation in generating viral (36, 52) and cell peptides, in the latter case from intronic regions (37). More importantly, we further demonstrated that IAV infection enhances cellular ARF DRiP generation, a finding very similar to that recently reported by Prasad et al (53), in which the generation of a model T cell epitope translated via alternative start codon was enhanced by viral infection. These findings may help us to further understand the relationship between virus infection and autoimmunity. Perhaps virus infection could trigger autoimmunity by simply increasing DRiP antigen presentation?

Pragmatically, it is important to reconsider strategies for identifying immunogenic and antigenic peptides. In our own specific circumstances, we would have found NS1-ARF2$_{1-8}$ had we synthesized all potential ARF peptides as at their predicted full length, utilizing serum proteases to liberate antigenically active fragments. Surely, many potentially biologically relevant peptides lurk in unexpected locations in the genomes of medically important viruses. We note that NS1-ARF2$_{1-8}$ is predicted to bind HLA-B*5101, HLA-B*3505 and HLA-B*5502 with sufficient affinity to be immunogenic, an obvious area of future research.

ARF2$_{1-8}$ elicits a robust response *in vivo*. Given predicted low levels of synthesis of the ARF combined with predicted rapid degradation, it is extremely unlikely that there is a sufficient amount of protein for cross-priming (54). This provides a clear example of the importance of direct priming in eliciting T_{CD8+} to a natural viral gene product. The position of NS1-ARF2$_{1-8}$ in the immunodominance hierarchy varies depending on the route and dose of infection. A similar phenomenon occurs with the T_{CD8+} response to VV in B6 mice (55). Multiple factors contribute to immunodominance (18, 22, 56), whose complexities are the natural product of the highly complex interaction between cells, viruses and all other constituent parts (19, 57). We note that NS1-ARF2$_{1-8}$ is a reasonably abundant peptide (copy number of ~ 2500 complexes per cell), ~3 fold more abundant than NP$_{147-155}$, which is present at 800 copies per cell (22). Yet there is no clear relationship between peptide abundance and immunodominance, with very abundant peptides sometime occupying low rungs on the immunodominance ladder (22).

In summary, we show that what is very likely to be mistranslation of a non-functional ARF encoded by IAV leads to the rapid presentation of an epitope at or near the top of the anti-viral T_{CD8+} immunodominance hierarchy. Along with the findings of Yang *et al (58)* demonstrating CUG

initiation of a +1 RF model peptide after the M1 mRNA stop codon, this provides clear biological evidence for the importance of DRiPs in peptide generation, extending the voluminous evidence from studies of peptide generation in cultured cells.

Figure legends

Figure 1

PR8 re-stimulation reveals a novel immunodominant H-2L^{d}-restricted response to NS1.

A) BALB/c mice were i.p. infected with 10^7 pfu PR8. 30 days later, splenic cells containing memory T cells specific to IAV were harvested and $3x10^6$ splenic cells infected with PR8 at 10 MOI for 5 hrs and used to restimulate $2.7x10^7$ splenic cells. T cell cultures were re-stimulated every 14 days with PR8-P815 infected at 10 MOI. 10 days following re-stimulation, culture specificity was assessed by ICS for IFN-γ using synthetic peptides or PR8-P815. Black bars denote T cell antigen specificity following the first stimulation; white bars denote specificity following the 4th stimulation. B) Assessment of T cell antigen specificity in a PR8-P815 restimulated T cell culture using single H-2^{d} allele transfected HEK293T cells infected with PR8 or rVVs encoding single IAV proteins. C) $NP_{147-155}$ and a 9-round PR8-P815 re-stimulated polyspecific T_{CD8+} line were assessed for specificity using P815 and HEK293T-L^{d} cells infected with PR8 or rVVs encoding single IAV proteins. All experiments are representative of at least two independent experiments. D) Assessment for the presentation of the two predicted NS1 ARF peptides and IAV infected P815 cells by IFN-γ ICS using a PR8-P815 restimulated T_{CD8+}. Splenic cells were harvested from PR8 infected memory BALB/c mice and two separate T cell cultures raised by re-stimulating splenic cells with (E) $ARF2_{1-8}$ and (F) $ARF2_{1-11}$ synthetic peptides. Antigen-specific T cell activation by ARF2 peptides or ARF2 peptides with an amino acid truncation or extension at either end was assessed by IFN-γ ICS using log-fold dilutions in serum-free condition. All experiments are representative of at least two independent experiments.

Figure 2

$ARF2_{1-8}$ is presented more abundantly than $NP_{147-155}$.

A) P815 cells were infected with PR8 at 10 MOI for 5 hrs, lysed and peptides extracted by acid and fractionated by RP-HPLC. Fractions were collected at 0.1 min intervals and pulsed onto low-temp induced P815 cells and used to assess for IFN-γ production in an ICS by i) PR8-P815-polyspecific, ii) $NP_{147-155}$-specific, iii) $ARF2_{1-8}$-specific and iv) $ARF2_{1-11}$-specific T_{CD8+} to confirm epitope presentation following a natural IAV infection. B) Peak fractions corresponding to synthetic peptide elution times for $NP_{147-155}$ and $ARF2_{1-8}$ were titrated with $NP_{147-155}$-specific and $ARF2_{1-8}$-specific T_{CD8+} respectively by IFN-γ ICS. C) Synthetic peptide titration curves for $NP_{147-155}$ and $ARF2_{1-8}$ specific T_{CD8+}. D) P815 cells were infected with PR8 at 10 MOI for 1 hr. Cells were washed and added to $NP_{147-155}$-specific (triangles) or $ARF2_{1-8}$-specific (square) T_{CD8+} lines, with BFA added at specific time pointed after infection and T cell lines assessed by IFN-γ ICS for peptide presentation kinetics. Max IFN-γ^{+} response determined by synthetic peptide shown in C). All experiments are representative of at least two independent experiments.

Figure3

Polyfunctional $ARF2_{1-8}$ specific T cell responses occur in IAV infected mice.

A) BALB/c mice were i.p. infected with 5 x 10^6 to 2 x 10^8 pfu of PR8 IAV. 7 days later, intraperitoneal wash cells were isolated and enumerated for T cell antigen specificity using synthetic peptides by IFN-γ ICS. BALB/c mice were i.n. infected with 100 pfu of PR8 or 10^7 pfu of Pan09 (both H1N1) IAV. 10 days later, B) bronchoalveolar lavage (BAL) and C) splenic cells were isolated and enumerated for T cell antigen specificity using synthetic peptides by IFN-γ ICS. D) IFN-γ^{+ve} i) $NP_{147-155}$- and ii) $ARF2_{1-8}$-specific T_{CD8+} were assessed for TNF-α production and iii) cell degranulation as measured by surface CD107a expression. N=4 mice/ gp. Error bars indicate SEM. All experiments are representative of at least two independent experiments.

Figure 4

Splicing is not required for $ARF2_{1-8}$ epitope generation from NS and eGFP-NS^{intron}

A) i) Schematic diagram of NS mRNA containing the position of $ARF2_{1-14}$ and NS1/2 mRNA splicing, and the key features and requirements for intron splicing. Schematic diagram of ii) NS-

full, NS-mut, iii) eGFP-NS^{intron}-WT and eGFP-NS^{intron}-mut constructs. B) Wild type D2SV cells and D2SV-NS transductants were infected with PR8 (10 MOI) or rVV-NS1 (10 MOI) for 4 hrs and then incubated with either $ARF2_{1-8}$-specific or $NP_{147-155}$-specific T_{CD8+} to assess for IFN-γ production by ICS. Far-right: Wild type P815 cells were infected with PR8 or PR8-GFP (No NS splicing) at 10MOI for 4 hrs and then incubated with either $ARF2_{1-8}$-specific or $NP_{147-155}$-specific T_{CD8+} to assess for IFN-γ production by ICS. Wild type P815 cells and P815-NS transductants were infected with 10 MOI PR8 or Pan09 IAV for 4 hrs and then incubated with either C) $ARF2_{1-8}$-specific or D) $NP_{147-155}$-specific T_{CD8+} to assess for IFN-γ production by ICS. Max IFN-γ^+ response for B) determined by synthetic peptide: $ARF2_{1-8}$: 60%, $NP_{147-155}$: 75%. All experiments are representative of at least two independent experiments.

Figure 5

$ARF2_{1-8}$ is a non-functional DRiP.

A) LA-4 cells were infected with 10 MOI wtPR8 or $ARF2_{1-8}$ KO PR8 IAV and assessed for i) NP, ii) M1, iii) M2 and iv) NS1 protein production by flow cytometry. B) C57B6 and C) BALB/c mice were infected intransally with 100 pfu of wtPR8 or $ARF2_{1-8}$ KO PR8 IAV and daily body weight recorded. Following 10 days, D, E) BAL wash and F, G) splenic cells were isolated and enumerated for T cell antigen specificity using synthetic peptides by IFN-γ ICS. N=4 mice/ gp. Error bars indicate SEM. All experiments are representative of at least two independent experiments.

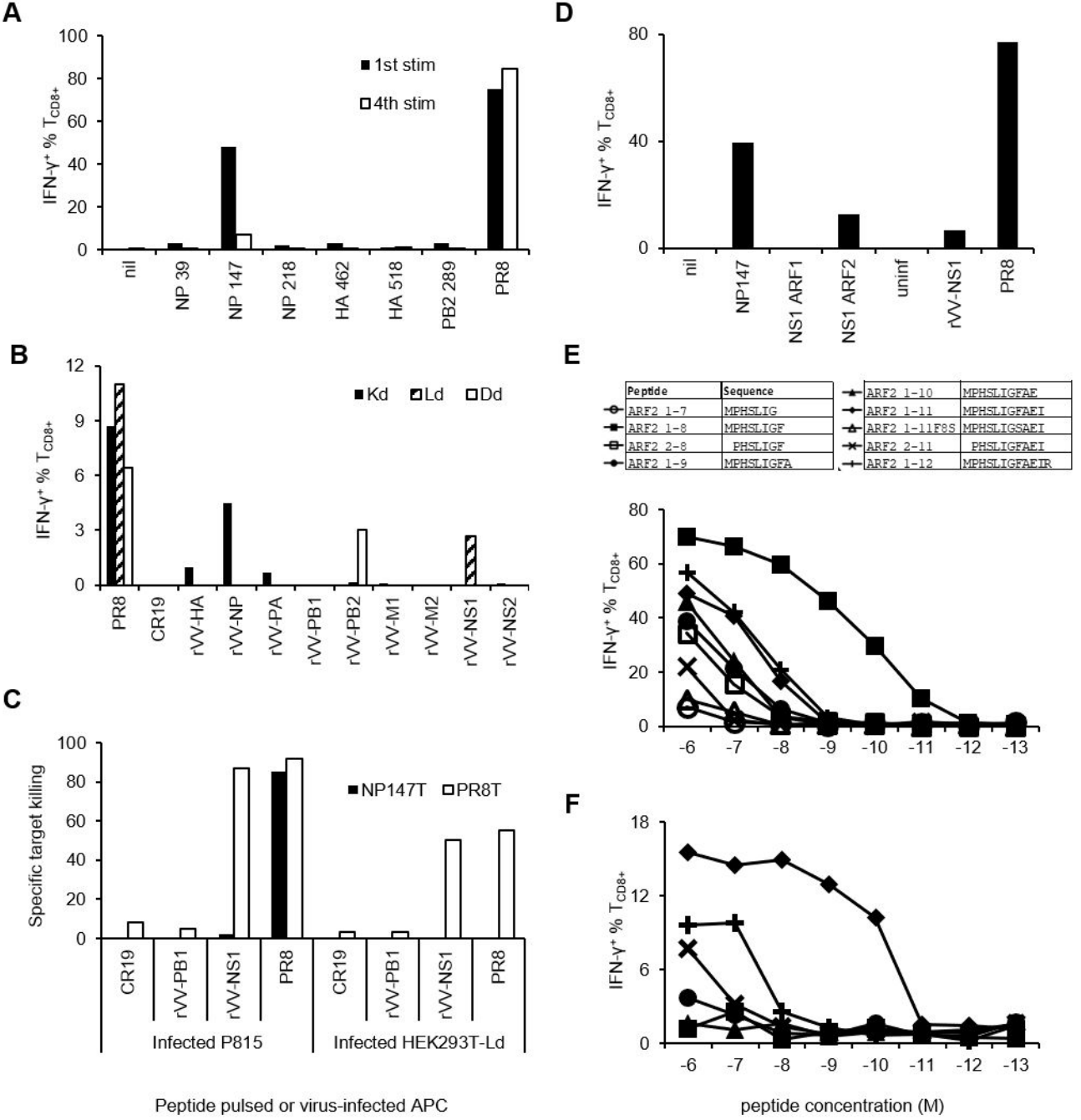
References

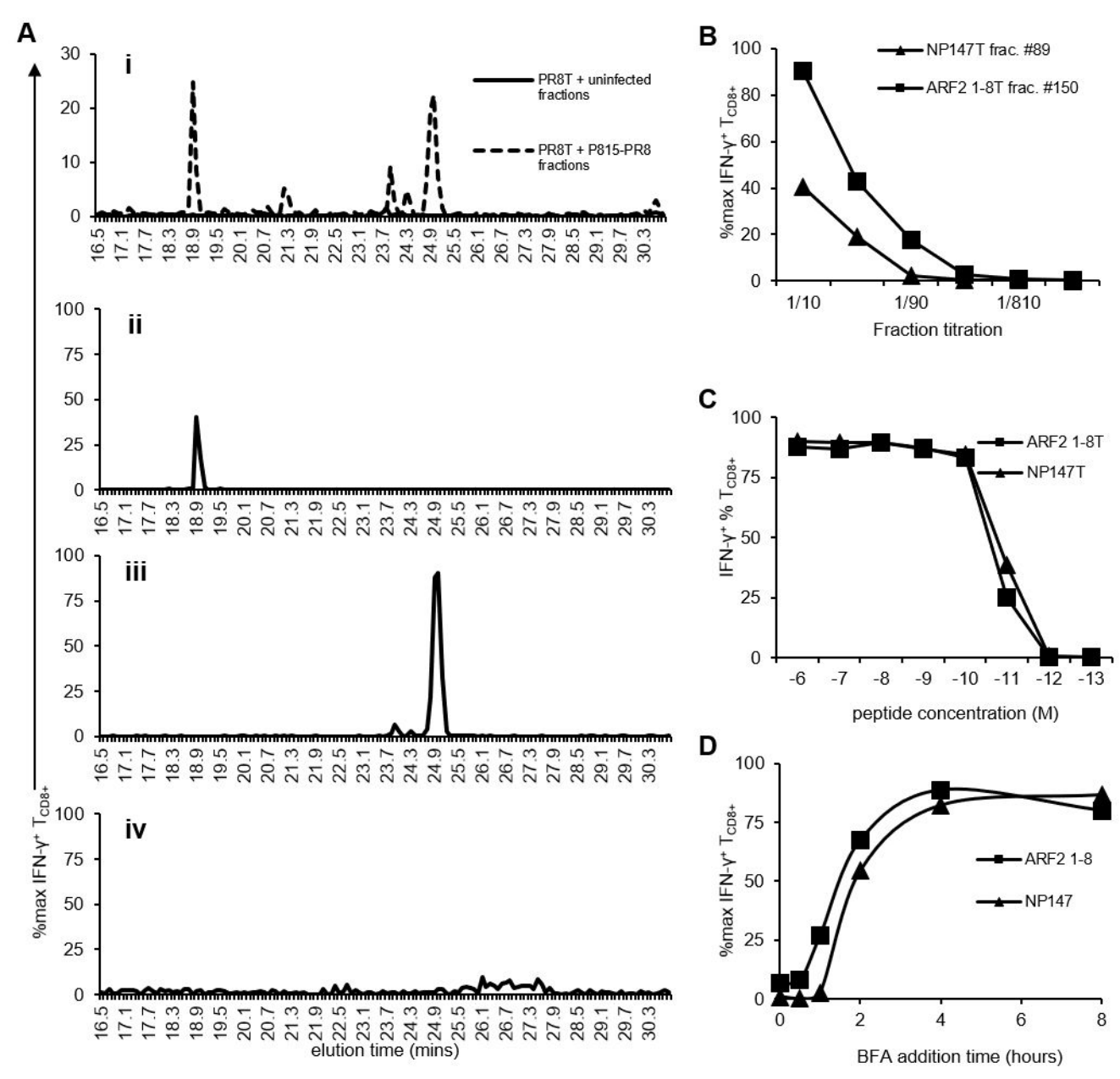
1. Esquivel, F., J. Yewdell, and J. Bennink. 1992. RMA/S cells present endogenously synthesized cytosolic proteins to class I-restricted cytotoxic T lymphocytes. *The Journal of experimental medicine* 175: 163-168.
2. Croft, N. P., A. W. Purcell, and D. C. Tscharke. 2015. Quantifying epitope presentation using mass spectrometry. *Molecular immunology*.
3. Anton, L. C., and J. W. Yewdell. 2014. Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors. *Journal of leukocyte biology* 95: 551-562.
4. Zook, M. B., M. T. Howard, G. Sinnathamby, J. F. Atkins, and L. C. Eisenlohr. 2006. Epitopes derived by incidental translational frameshifting give rise to a protective CTL response. *J Immunol* 176: 6928-6934.
5. Berglund, P., D. Finzi, J. R. Bennink, and J. W. Yewdell. 2007. Viral alteration of cellular translational machinery increases defective ribosomal products. *Journal of virology* 81: 7220-7229.
6. Bullock, T. N., A. E. Patterson, L. L. Franlin, E. Notidis, and L. C. Eisenlohr. 1997. 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.
7. Bullock, T. N. J., and L. C. Eisenlohr. 1996. Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. *Journal of Experimental Medicine* 184: 1319-1329.
8. Chen, W., P. A. Calvo, D. Malide, J. Gibbs, U. Schubert, I. Bacik, S. Basta, R. O'Neill, J. Schickli, P. Palese, P. Henklein, J. R. Bennink, and J. W. Yewdell. 2001. A novel influenza A virus mitochondrial protein that induces cell death. *Nature medicine* 7: 1306-1312.
9. Cardinaud, S., A. Moris, M. Fevrier, P. S. Rohrlich, L. Weiss, P. Langlade-Demoyen, F. A. Lemonnier, O. Schwartz, and A. Habel. 2004. Identification of cryptic MHC I-restricted epitopes encoded by HIV-1 alternative reading frames. *The Journal of experimental medicine* 199: 1053-1063.
10. Maness, N. J., A. D. Walsh, S. M. Piaskowski, J. Furlott, H. L. Kolar, A. T. Bean, N. A. Wilson, and D. I. Watkins. 2010. CD8+ T cell recognition of cryptic epitopes is a ubiquitous feature of AIDS virus infection. *Journal of virology* 84: 11569-11574.
11. Maness, N. J., L. E. Valentine, G. E. May, J. Reed, S. M. Piaskowski, T. Soma, J. Furlott, E. G. Rakasz, T. C. Friedrich, D. A. Price, E. Gostick, A. L. Hughes, J. Sidney, A. Sette, N. A. Wilson, and D. I. Watkins. 2007. AIDS virus specific CD8+ T lymphocytes against an immunodominant cryptic epitope select for viral escape. *The Journal of experimental medicine* 204: 2505-2512.
12. Smirnova, E., A. E. Firth, W. A. Miller, D. Scheidecker, V. Brault, C. Reinbold, A. M. Rakotondrafara, B. Y. Chung, and V. Ziegler-Graff. 2015. Discovery of a Small Non-AUG-Initiated ORF in Poleroviruses and Luteoviruses That Is Required for Long-Distance Movement. *PLoS pathogens* 11: e1004868.
13. Galindo, M. I., J. I. Pueyo, S. Fouix, S. A. Bishop, and J. P. Couso. 2007. Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS Biol* 5: e106.
14. Wise, H. M., A. Foeglein, J. Sun, R. M. Dalton, S. Patel, W. Howard, E. C. Anderson, W. S. Barclay, and P. Digard. 2009. A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *Journal of virology* 83: 8021-8031.
15. Jagger, B. W., H. M. Wise, J. C. Kash, K. A. Walters, N. M. Wills, Y. L. Xiao, R. L. Dunfee, L. M. Schwartzman, A. Ozinsky, G. L. Bell, R. M. Dalton, A. Lo, S. Efstathiou, J. F. Atkins, A. E. Firth, J. K. Taubenberger, and P. Digard. 2012. An overlapping protein-

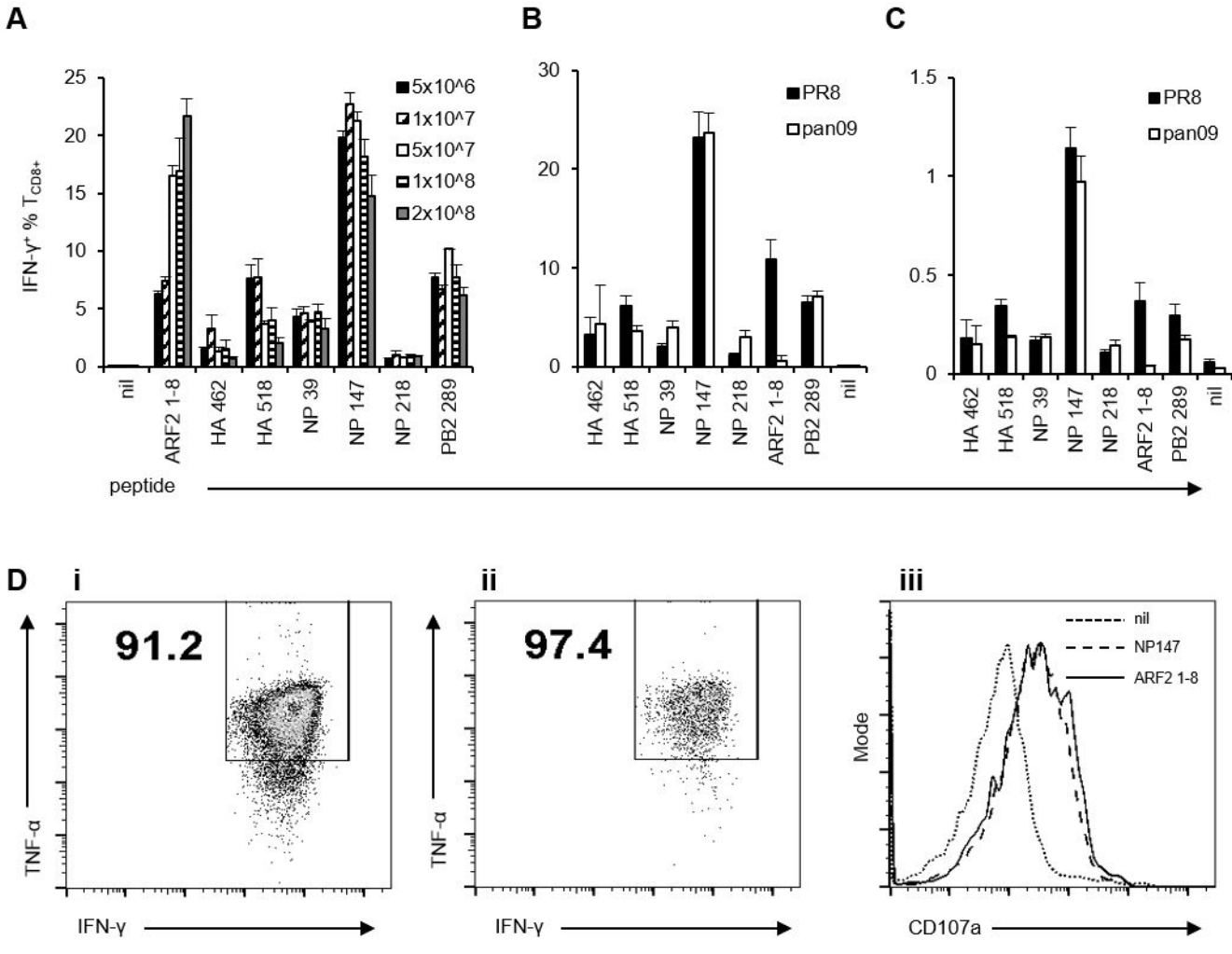
coding region in influenza A virus segment 3 modulates the host response. *Science (New York, N.Y* 337: 199-204.
16. Muramoto, Y., T. Noda, E. Kawakami, R. Akkina, and Y. Kawaoka. 2013. Identification of novel influenza A virus proteins translated from PA mRNA. *Journal of virology* 87: 2455-2462.
17. Pala, P., H. C. Bodmer, R. M. Pemberton, J. C. Cerottini, J. L. Maryanski, and B. A. Askonas. 1988. Competition between unrelated peptides recognized by H-2-Kd restricted T cells. *J Immunol* 141: 2289-2294.
18. Deng, Y., J. W. Yewdell, L. C. Eisenlohr, and J. R. Bennink. 1997. MHC affinity, peptide liberation, T cell repertoire, and immunodominance all contribute to the paucity of MHC class I-restricted peptides recognized by antiviral CTL. *J Immunol* 158: 1507-1515.
19. Zanker, D., J. Waithman, J. W. Yewdell, and W. Chen. 2013. Mixed Proteasomes Function To Increase Viral Peptide Diversity and Broaden Antiviral CD8+ T Cell Responses. *J Immunol* 191: 52-59.
20. Chen, W., K. Pang, K. A. Masterman, G. Kennedy, S. Basta, N. Dimopoulos, F. Hornung, M. Smyth, J. R. Bennink, and J. W. Yewdell. 2004. Reversal in the immunodominance hierarchy in secondary CD8+ T cell responses to influenza A virus: roles for cross-presentation and lysis-independent immunodomination. *J Immunol* 173: 5021-5027.
21. Chen, W., J. R. Bennink, P. A. Morton, and J. W. Yewdell. 2002. Mice deficient in perforin, CD4+ T cells, or CD28-mediated signaling maintain the typical immunodominance hierarchies of CD8+ T-cell responses to influenza virus. *Journal of virology* 76: 10332-10337.
22. Chen, W., L. C. Anton, J. R. Bennink, and J. W. Yewdell. 2000. Dissecting the multifactorial causes of immunodominance in class I-restricted T cell responses to viruses. *Immunity* 12: 83-93.
23. Manicassamy, B., S. Manicassamy, A. Belicha-Villanueva, G. Pisanelli, B. Pulendran, and A. Garcia-Sastre. 2010. Analysis of in vivo dynamics of influenza virus infection in mice using a GFP reporter virus. *Proceedings of the National Academy of Sciences of the United States of America* 107: 11531-11536.
24. Hoffmann, E., G. Neumann, Y. Kawaoka, G. Hobom, and R. G. Webster. 2000. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proceedings of the National Academy of Sciences of the United States of America* 97: 6108-6113.
25. Chen, W., J. W. Yewdell, R. L. Levine, and J. R. Bennink. 1999. Modification of cysteine residues in vitro and in vivo affects the immunogenicity and antigenicity of major histocompatibility complex class I-restricted viral determinants. *The Journal of experimental medicine* 189: 1757-1764.
26. Pang, K. C., M. T. Sanders, J. J. Monaco, P. C. Doherty, S. J. Turner, and W. Chen. 2006. Immunoproteasome subunit deficiencies impact differentially on two immunodominant influenza virus-specific CD8+ T cell responses. *J Immunol* 177: 7680-7688.
27. Kim, D., B. Langmead, and S. L. Salzberg. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12: 357-360.
28. Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15: 550.
29. Bennink, J. R., and J. W. Yewdell. 1988. Murine cytotoxic T lymphocyte recognition of individual influenza virus proteins: High frequency of nonresponder MHC class I alleles. *J. Exp. Med.* 168: 1935-1939.
30. Reay, P. A., I. M. Jones, and G. G. Brownlee. 1988. A region of the influenza A/NT/60/68 PB2 protein containing an antigenic determinant recognized by murine H-2Dd restricted cytotoxic T lymphocytes. *Virology* 167: 261-268.
31. Tscharke, D. C., W. P. Woo, I. G. Sakala, J. Sidney, A. Sette, D. J. Moss, J. R. Bennink, G. Karupiah, and J. W. Yewdell. 2006. Poxvirus CD8+ T-cell determinants and cross-reactivity in BALB/c mice. *Journal of virology* 80: 6318-6323.

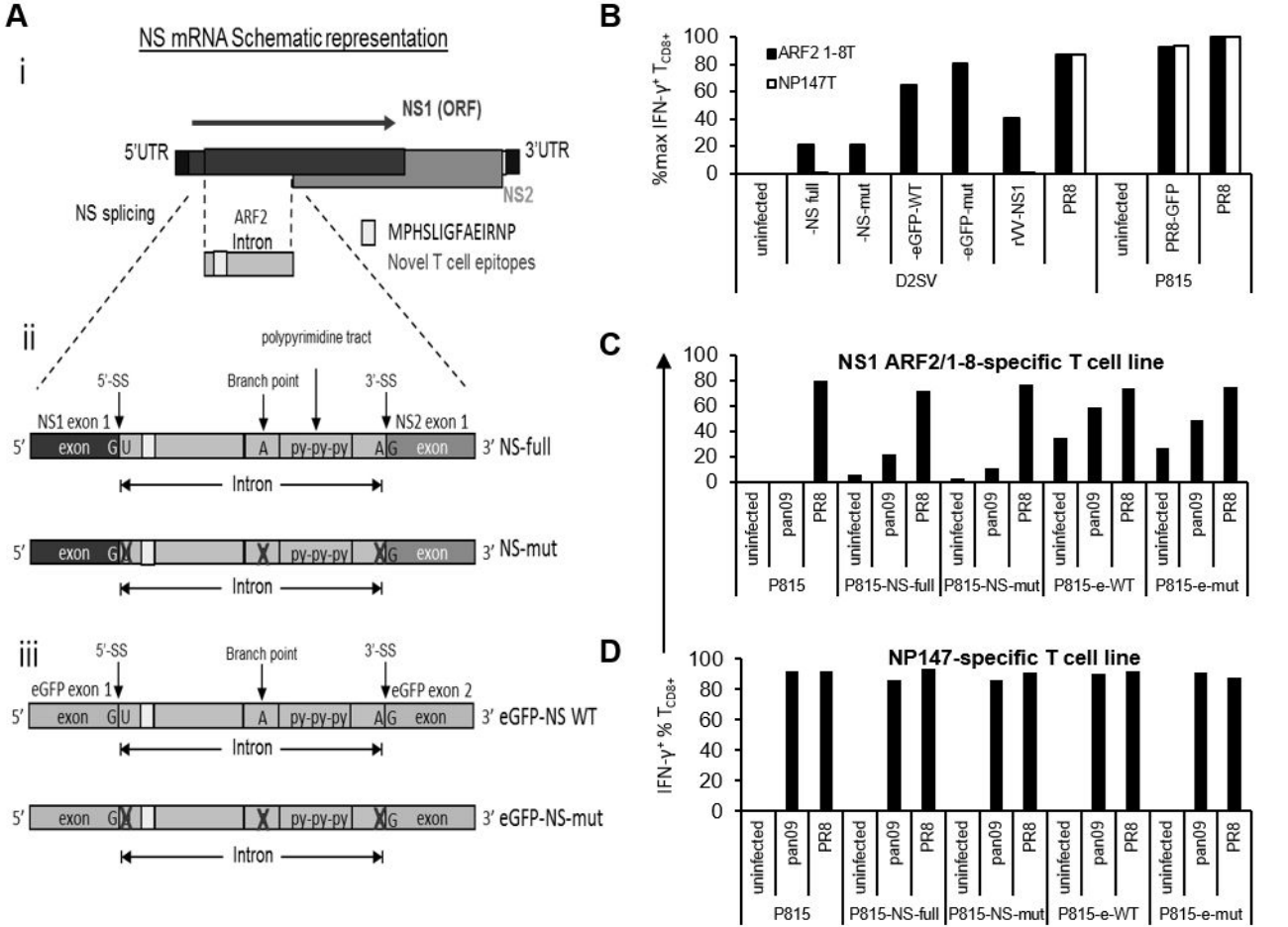
32. Kim, Y., J. Sidney, S. Buus, A. Sette, M. Nielsen, and B. Peters. 2014. Dataset size and composition impact the reliability of performance benchmarks for peptide-MHC binding predictions. *BMC Bioinformatics* 15: 241.
33. Green, K. J., J. J. Miles, J. Tellam, W. J. van Zuylen, G. Connolly, and S. R. Burrows. 2004. Potent T cell response to a class I-binding 13-mer viral epitope and the influence of HLA micropolymorphism in controlling epitope length. *European journal of immunology* 34: 2510-2519.
34. Ebert, L. M., Y. C. Liu, C. S. Clements, N. C. Robson, H. M. Jackson, J. L. Markby, N. Dimopoulos, B. S. Tan, I. F. Luescher, I. D. Davis, J. Rossjohn, J. Cebon, A. W. Purcell, and W. Chen. 2009. A long, naturally presented immunodominant epitope from NY-ESO-1 tumor antigen: implications for cancer vaccine design. *Cancer research* 69: 1046-1054.
35. Probst-Kepper, M., V. Stroobant, R. Kridel, B. Gaugler, C. Landry, F. Brasseur, J. P. Cosyns, B. Weynand, T. Boon, and B. J. Van Den Eynde. 2001. An alternative open reading frame of the human macrophage colony- stimulating factor gene is independently translated and codes for an antigenic peptide of 14 amino acids recognized by tumor-infiltrating CD8 T lymphocytes. *The Journal of experimental medicine* 193: 1189-1198.
36. Dolan, B. P., J. J. Knowlton, A. David, J. R. Bennink, and J. W. Yewdell. 2010. RNA polymerase II inhibitors dissociate antigenic peptide generation from normal viral protein synthesis: a role for nuclear translation in defective ribosomal product synthesis? *J Immunol* 185: 6728-6733.
37. Apcher, S., G. Millot, C. Daskalogianni, A. Scherl, B. Manoury, and R. Fahraeus. 2013. Translation of pre-spliced RNAs in the nuclear compartment generates peptides for the MHC class I pathway. *Proceedings of the National Academy of Sciences of the United States of America* 110: 17951-17956.
38. Doronina, V. A., C. Wu, P. de Felipe, M. S. Sachs, M. D. Ryan, and J. D. Brown. 2008. Site-specific release of nascent chains from ribosomes at a sense codon. *Molecular and cellular biology* 28: 4227-4239.
39. Storz, G., Y. I. Wolf, and K. S. Ramamurthi. 2014. Small proteins can no longer be ignored. *Annual review of biochemistry* 83: 753-777.
40. Stoner, G. D., Y. Kikkawa, A. J. Kniazeff, K. Miyai, and R. M. Wagner. 1975. Clonal isolation of epithelial cells from mouse lung adenoma. *Cancer research* 35: 2177-2185.
41. Chen, W., J. R. Bennink, and J. W. Yewdell. 2003. Systematic search fails to detect immunogenic MHC class-I-restricted determinants encoded by influenza A virus noncoding sequences. *Virology* 305: 50-54.
42. Kosik, I., J. Holly, and G. Russ. 2013. PB1-F2 expedition from the whole protein through the domain to aa residue function. *Acta Virol* 57: 138-148.
43. Chakrabarti, A. K., and G. Pasricha. 2013. An insight into the PB1F2 protein and its multifunctional role in enhancing the pathogenicity of the influenza A viruses. *Virology* 440: 97-104.
44. Bain, C., P. Parroche, J. P. Lavergne, B. Duverger, C. Vieux, V. Dubois, F. Komurian-Pradel, C. Trepo, L. Gebuhrer, G. Paranhos-Baccala, F. Penin, and G. Inchauspe. 2004. Memory T-cell-mediated immune responses specific to an alternative core protein in hepatitis C virus infection. *Journal of virology* 78: 10460-10469.
45. Cardinaud, S., G. Consiglieri, R. Bouziat, A. Urrutia, S. Graff-Dubois, S. Fourati, I. Malet, J. Guergnon, A. Guihot, C. Katlama, B. Autran, P. van Endert, F. A. Lemonnier, V. Appay, O. Schwartz, P. M. Kloetzel, and A. Moris. 2011. CTL escape mediated by proteasomal destruction of an HIV-1 cryptic epitope. *PLoS pathogens* 7: e1002049.
46. Ingolia, N. T., G. A. Brar, N. Stern-Ginossar, M. S. Harris, G. J. Talhouarne, S. E. Jackson, M. R. Wills, and J. S. Weissman. 2014. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell reports* 8: 1365-1379.

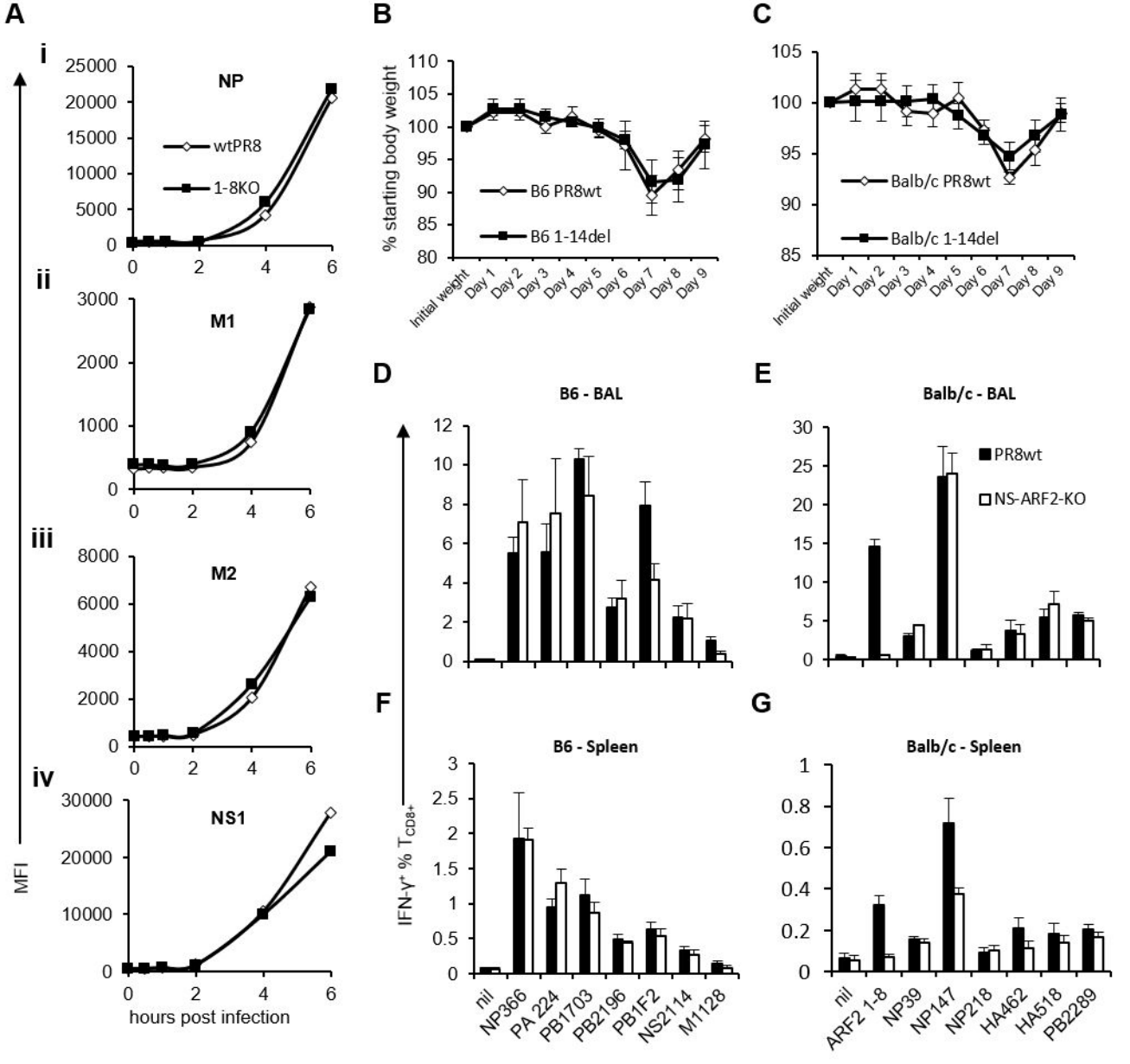
47. Stern-Ginossar, N., B. Weisburd, A. Michalski, V. T. Le, M. Y. Hein, S. X. Huang, M. Ma, B. Shen, S. B. Qian, H. Hengel, M. Mann, N. T. Ingolia, and J. S. Weissman. 2012. Decoding human cytomegalovirus. *Science* 338: 1088-1093.
48. McLean, D. T., M. T. McCrudden, G. J. Linden, C. R. Irwin, J. M. Conlon, and F. T. Lundy. 2014. Antimicrobial and immunomodulatory properties of PGLa-AM1, CPF-AM1, and magainin-AM1: potent activity against oral pathogens. *Regulatory peptides* 194-195: 63-68.
49. Lauressergues, D., J. M. Couzigou, H. S. Clemente, Y. Martinez, C. Dunand, G. Becard, and J. P. Combier. 2015. Primary transcripts of microRNAs encode regulatory peptides. *Nature* 520: 90-93.
50. Qian, X. Y., F. Alonso-Caplen, and R. M. Krug. 1994. Two functional domains of the influenza virus NS1 protein are required for regulation of nuclear export of mRNA. *Journal of virology* 68: 2433-2441.
51. Dauber, B., J. Schneider, and T. Wolff. 2006. Double-stranded RNA binding of influenza B virus nonstructural NS1 protein inhibits protein kinase R but is not essential to antagonize production of alpha/beta interferon. *Journal of virology* 80: 11667-11677.
52. David, A., B. P. Dolan, H. D. Hickman, J. J. Knowlton, G. Clavarino, P. Pierre, J. R. Bennink, and J. W. Yewdell. 2012. Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *The Journal of cell biology* 197: 45-57.
53. Prasad, S., S. R. Starck, and N. Shastri. 2016. Presentation of Cryptic Peptides by MHC Class I Is Enhanced by Inflammatory Stimuli. *J Immunol* 197: 2981-2991.
54. Norbury, C. C., S. Basta, K. B. Donohue, D. C. Tscharke, M. F. Princiotta, P. Berglund, J. Gibbs, J. R. Bennink, and J. W. Yewdell. 2004. CD8+ T cell cross-priming via transfer of proteasome substrates. *Science (New York, N.Y* 304: 1318-1321.
55. Tscharke, D. C., G. Karupiah, J. Zhou, T. Palmore, K. R. Irvine, S. M. Haeryfar, S. Williams, J. Sidney, A. Sette, J. R. Bennink, and J. W. Yewdell. 2005. Identification of poxvirus CD8+ T cell determinants to enable rational design and characterization of smallpox vaccines. *The Journal of experimental medicine* 201: 95-104.
56. Yewdell, J. W. 2006. Confronting complexity: real-world immunodominance in antiviral CD8+ T cell responses. *Immunity* 25: 533-543.
57. Lin, L. C., I. E. Flesch, and D. C. Tscharke. 2013. Immunodomination during peripheral vaccinia virus infection. *PLoS pathogens* 9: e1003329.
58. Yang, N., J. S. Gibbs, H. D. Hickman, G. V. Reynoso, A. K. Ghosh, J. R. Bennink, and J. W. Yewdell. 2016. 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.


A
IFN-γ+ % TCD8+
1st stim
4th stim
nil
NP 39
NP 147
NP 218
HA 462
HA 518
PB2 289
PR8
B
IFN-γ+ % TCD8+
Kd
Ld
Dd
PR8
CR19
rVV-HA
rVV-NP
rVV-PA
rVV-PB1
rVV-PB2
rVV-M1
rVV-M2
rVV-NS1
rVV-NS2
C
Specific target killing
NP147T
PR8T
CR19
rVV-PB1
rVV-NS1
PR8
Infected P815
Infected HEK293T-Ld
Peptide pulsed or virus-infected APC
D
IFN-γ+ % TCD8+
nil
NP147
NS1 ARF1
NS1 ARF2
uninf
rVV-NS1
PR8
E
Peptide
Sequence
ARF2 1-7
MPHSLIG
ARF2 1-8
MPHSLIGF
ARF2 2-8
PHSLIGF
ARF2 1-9
MPHSLIGFA
ARF2 1-10
MPHSLIGFAE
ARF2 1-11
MPHSLIGFAEI
ARF2 1-11F8S
MPHSLIGSAEI
ARF2 2-11
PHSLIGFAEI
ARF2 1-12
MPHSLIGFAEIR
IFN-γ+ % TCD8+
F
IFN-γ+ % TCD8+
peptide concentration (M)


A
i
PR8T + uninfected fractions
PR8T + P815-PR8 fractions
ii
iii
iv
%max IFN-γ+ TCD8+
elution time (mins)
B
NP147T frac. #89
ARF2 1-8T frac. #150
%max IFN-γ+ TCD8+
1/10
1/90
1/810
Fraction titration
C
ARF2 1-8T
NP147T
IFN-γ+ % TCD8+
peptide concentration (M)
D
ARF2 1-8
NP147
%max IFN-γ+ TCD8+
BFA addition time (hours)


A
B
C
D
i
ii
iii
IFN-γ+ % TCD8+
5x10^6
1x10^7
5x10^7
1x10^8
2x10^8
PR8
pan09
nil
ARF2 1-8
HA 462
HA 518
NP 39
NP 147
NP 218
PB2 289
peptide
91.2
97.4
TNF-α
IFN-γ
NP147
Mode
CD107a


A
NS mRNA Schematic representation
i
NS1 (ORF)
5'UTR
3'UTR
NS2
NS splicing
ARF2
Intron
MPHSLIGFAEIRNP
Novel T cell epitopes
ii
polypyrimidine tract
5'-SS
Branch point
3'-SS
NS1 exon 1
NS2 exon 1
5' exon GU A py-py-py AG exon 3' NS-full
Intron
5' exon GX X py-py-py XG exon 3' NS-mut
Intron
iii
5'-SS
Branch point
3'-SS
eGFP exon 1
eGFP exon 2
5' exon GU A py-py-py AG exon 3' eGFP-NS WT
Intron
5' exon GX X py-py-py XG exon 3' eGFP-NS-mut
Intron
B
%max IFN-γ+ TCD8+
ARF2 1-8T
NP147T
uninfected
-NS full
-NS-mut
-eGFP-WT
-eGFP-mut
rVV-NS1
PR8
uninfected
PR8-GFP
PR8
D2SV
P815
C
NS1 ARF2/1-8-specific T cell line
uninfected
pan09
PR8
P815
P815-NS-full
P815-NS-mut
P815-e-WT
P815-e-mut
D
IFN-γ+ % TCD8+
NP147-specific T cell line
uninfected
pan09
PR8
P815
P815-NS-full
P815-NS-mut
P815-e-WT
P815-e-mut


A
i
NP
wtPR8
1-8KO
ii
M1
iii
M2
iv
NS1
MFI
hours post infection
B
% starting body weight
B6 PR8wt
B6 1-14del
Initial weight
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Day 7
Day 8
Day 9
C
Balb/c PR8wt
Balb/c 1-14del
D
B6 - BAL
E
Balb/c - BAL
PR8wt
NS-ARF2-KO
F
B6 - Spleen
IFN-γ+ % T CD8+
nil
NP366
PA 224
PB1703
PB2196
PB1F2
NS2114
M1128
G
Balb/c - Spleen
nil
ARF2 1-8
NP39
NP147
NP218
HA462
HA518
PB2289