HHS Public Access

Author manuscript

Nat Med. Author manuscript; available in PMC 2011 May 01.

Published in final edited form as:

Nat Med. 2010 November ; 16(11): 1333–1340. doi:10.1038/nm.2248.

A novel Minimalist Cell-Free MHC Class II Antigen Processing System Identifies Immunodominant Epitopes

Isamu Z. Hartman[1,2,8], **AeRyon Kim**[1,2], **Robert J. Cotter**[4], **Kimberly Walter**[3], **Sarat K. Dalai**[3], **Tatiana Boronina**[5], **Wendell Griffith**[4], **Robert Schwenk**[7], **David E. Lanar**[7], **Urszula Krzych**[7], **Robert N. Cole**[5,6], and **Scheherazade Sadegh-Nasseri**[2,3,*]

[2] Graduate Program in Immunology

[3] Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[4] Department of Pharmacology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[5] Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[6] Mass Spectrometry and Proteomics Facility, Institute for Basic Biomedical Sciences, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[7] Division of Malaria Vaccine Development, Walter Reed Army Institute of Research, Silver Spring, MD 20910

Abstract

Immunodominance is defined as restricted responsiveness of T cells to a few selected epitopes from complex antigens. Strategies currently used for elucidating $CD4^+$ T cell epitopes are inadequate. To understand the mechanism of epitope selection for helper T cells, we established a cell-free antigen processing system composed of defined proteins: MHC class II, cathepsins, and HLA-DM. Our minimalist system successfully identified the physiologically selected immunodominant epitopes of model antigens, HA1 from influenza virus (A/Texas/1/77) and type II collagen. When applied for *de novo* epitope identification to a malaria antigen, or HA1 from H5N1 virus (Avian Flu), the system selected a single epitope from each protein that were confirmed to be immunodominant by their capacity to activate $CD4^+$ T cells in HLA-DR1 positive human volunteers or transgenic mice immunized with the corresponding proteins. Thus, we provide a powerful new tool for the identification of physiologically relevant helper T cell epitopes from antigens.

[*] Contact: ssadegh@jhmi.edu.

[1] equal contribution

[8] Current address: Department of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

Supplemental Data

Supplemental data include 14 Supplemental Figures, 4 Tables, and legends and can be found with this article online.

It is well known that the immune system focuses on and responds to very few representative immunodominant epitopes from invading pathogenic insults ranging from such as infectious agents and antigenic targets in autoimmune diseases, allergy, and cancer[1,2]. In each of these cases, the immune system either responds positively or fails to respond to the immunodominant epitopes in the context of MHC molecules. Recent advances in our understanding of the antigen presentation pathway have shown that the steps of antigen processing and selection critically influence the peptide repertoire presented to T-cells. Thus, methods for determining MHC-restricted immunodominant epitopes would be more accurate if all of the steps preceding antigen presentation were inclusively integrated together. Various approaches have been developed to identify MHC-restricted peptides that may elicit strong T-cell responses. One of these approaches uses large arrays of overlapping synthetic peptides that map the entire amino acid sequence of an antigenic protein to identify MHC-restricted peptides and, consequently, putative immunodominant epitopes. The large number of protein sequences available for mapping is now making peptide-mapping experiments highly expensive, labor intensive and often unreliable. Alternative approaches based on computer-driven algorithms have also been developed to predict putative MHC-restricted binders. These algorithms utilize matrix- and anchor-based motifs to screen amino acid sequences of antigenic proteins.

To examine if immunodominance is regulated during antigen processing, we designed a simple cell - free system composed of defined protein components: full-length antigen and five purified proteins (human MHC class II, HLA-DM, and cathepsins, B, H, and S). This system was tested successfully against model antigens; recombinant Influenza virus HA1 protein (rHA1) containing the well characterized DR1 restricted immunodominant peptide, $HA_{306\text{-}318}$, from A/Texas/1/77 strain[3] and type II collagen which is implicated in collagen induced arthritis in mice[4]. The minimalist system was evaluated for identifying HLA-DR1 restricted unknown immunodominant epitope(s) using HA1 from A/Vietnam/1203/2004 (H5N1) and recombinant malaria antigen, LSA-NRC, as the novel antigens. We demonstrate that the cell-free MHC class II antigen processing system identified a single epitope from either H5N1 or LSA-NRC, which was verified in HLA-DR1 positive humans or transgenic mice as immunodominant. Thus, our system provides a unique and novel technology for identifying the immunodominant epitopes from other unknown antigens.

Results

Components of a Cell-Free Antigen Processing System

Antigen processing is complex and involves multiple steps, many chaperones, and several accessory proteins. For MHC class II processing, antigens are taken up by antigen presenting cells from exogenous sources and shuttled through a series of endosomal compartments. These compartments contain a denaturing environment, accessory chaperones, and proteolytic enzymes that digest protein antigens and allow binding of some peptide fragments to the groove of MHC class II molecules[5]. To recreate the MHC class II antigen processing compartment[6], we selected a minimum number of essential components: a soluble form of the human MHC II molecule (HLA-DR1), soluble HLA-DM (DM), and cathepsins B, H, and S. We included DM in the system because of its role in peptide

editing[7],[8]. DM is known for catalyzing displacement of class II-associated invariant chain peptide (CLIP) and other peptides from the MHC groove as well as for inducing peptide binding[9],[10]. DM operates by generating a peptide-receptive MHC class II, which it accomplishes by exerting conformational changes in class II-peptide complexes that prevent formation of H-bonds between MHC class II and the peptide main chain[11]. By inducing MHC II to adopta peptide-receptive conformation that can quickly sample a large pool of peptides derived from exogenously acquired proteins, DM acts as a peptide editor that might exert influence on epitope selection.

Cathepsin S is the major endoprotease involved in class II antigen processing outside of the thymus and can generate smaller fragments from full - length proteins, so it was chosen as our only endoprotease[12],[13]. Exopeptidases, cathepsin B and cathepsin H, were chosen because they are constitutively expressed in all professional APCs and have carboxypeptidase (cathepsin B) and aminopeptidase (cathepsin H) activities important for trimming longer fragments bound to MHC II molecules. Cathepsins B and H have endoprotease activity as well[14]. To mimic endosomal compartment, free L-cysteine was included[15] to aid protein unfolding[16]. L-cysteine also helps the catalytic activity of thiol-dependent lysosomal enzymes. Since cathepsins and DM are active in acidic pH, we used citrate phosphate buffer at pH 5.

An SDS-PAGE assay showed little proteolysis of DR1 (either empty or peptide bound) to cathepsins B and H (Fig. 1a, b) under the conditions adopted for the digestion of rHA1 (Fig. 1a). Empty DR1 was susceptible to cathepsin S digestion but peptide bound DR1 complex was resistant (Fig. 1b). Soluble DM was somewhat sensitive to all three cathepsins. Because of the susceptibility of DM and empty DR1 to cathepsins, we pre-incubated antigens with DR1 and DM prior to the inclusion of the proteases. Once all components of the system were incubated together under endosomal/lysosomal-like conditions, peptide–DR1 complexes were isolated by immunoprecipitation, peptides eluted from DR1 and were analyzed on a Matrix-Assisted Laser Desorption Ionization (MALDI) mass spectrometer (Supplementary Fig. 1).

rHA1 of influenza with known immunodominant epitope

We first attempted a well-defined immunodominant epitope from rHA1 of influenza strain A/PR/8/34 to which the A/Texas/1/77-derived $HA_{306–318}$ (PKYVKQNTLKLAT) epitope[3] was genetically attached near its C-terminus (Supplementary Table 1). Comparison of spectra of peptides eluted from DR1 versus the background that did not contain rHA1 marked rHA1-derived peptide species, several of which could be detected in the m/z 1950-2530 range (Fig. 1 c–f). Sequences of 6 out of 10 such peptide species were determined through collision-induced dissociation (CID). Among those, 4 contained the immunodominant $HA_{306-318}$ epitope (in bold) from Influenza strain A/Texas/1/77 (at m/z 2153.09, 2217.27, 2281.27, 2524.45) (Fig. 1c, and Supplementary Fig. 3,4). Sequences of the other two peptides at m/z 2265.09 and 2339.09 (underlined) contained INSSLPYQNIHPVITIGECPK, derived from influenza strain A/PR/8/34 (Fig. 1c, and Supplementary Fig. 5, 6). For the last four peptides whose sequence identification through CID posed a challenge (in black), a search was conducted to match their masses with rHA1-

derived sequences using Findpept (http://www.expasy.org/tools/findpept.html), while taking certain artifactual modifications into consideration[17,18]. Peptides at m/z 1955.09, 2118.00, and 2503.18 were each determined to possibly contain peptides derived from influenza strain A/PR/8/34 (Supplementary Table 2). To rule out the possibility of preexisting peptide fragments containing $HA_{306-318}$ during the preparation of rHA1 proteins, we examined the repertoire of peptides bound to DR1 in the presence or absence of the added cathepsins. None of rHA1-derived peptides were detected without processing of protein antigens by the cathepsins (Supplementary Fig. 7). We prepared samples that included cathepsin S in addition to cathepsin B and H (Fig. 1e, f). This sample reaction produced a profile nearly identical to that of the reaction containing cathepsins B and H only (Fig. 1c), with the exception that including cathepsins S caused elimination of rHA1-derived peptides atm/z 1955.09 and $HA_{298-317}$ (A/PR/8/34) at m/z 2339.09 (shown by arrows).

Immunodominance of the identified peptides were determined in DR1 transgenic mice immunized with the rHA1 in CFA. These mice[4] express a fusion product of the peptide-binding groove of DR1 and the membrane proximal domain of I-E. These mice express DR1 as their only MHC II[19], as they were backcrossed to MHC class II deletion mutants for 13 generations to eliminate endogenous $I\text{-}A^f$. Draining lymph nodes were isolated and the cells were used in a recall T cell proliferation assay using titrating doses of the identified peptides, rHA1, or $CLIP_{89-105}$ as a negative control. We observed strong dose dependent responses to $HA_{306-318}$ (A/Texas/1/77) and to the rHA1 whereas but no response to $CLIP_{89-105}$. There was a significantly lower (2 out of five mice) (Fig. 1g) or no response (3 mice) to $HA_{298-317}$ (A/PR/8/34) (Fig. 1h). The sum of the proliferative responses to $HA_{306-318}$ and $HA_{298-317}$ peptides approached the magnitude of the response to the whole HA protein (Fig. 1g, dashed line). Thus, our minimalist system could identify the immunodominant epitope.

Type II collagen with known immunodominant epitope

We next proceeded with another well-defined immunodominant epitope, type II collagen (CII). CII, a major component of cartilage, is the main suspected autoantigen in Rheumatoid Arthritis (RA) in $DR1^{Positive}$ individuals. Through studies conducted on DR1 transgenic mice, $CII_{282-289}$ (FKGEQGPK), has been identified as its DR1-restricted immunodominant core epitope[4]. To recapitulate physiological conditions of digesting this antigen, we pre-digested CII with matrix metalloproteinase 9 (MMP9) because it has been shown that CII undergoes extracellular processing first, with the resulting fragments being further processed after uptake by professional APC[20]. MMP9-digested CII was included as a test antigen, and DR1-bound peptides were isolated and analyzed. The majority of CII-derived peptide species appeared in the m/z 3000-3500 range (Fig. 2a–c). The most prominent peak of this cluster was sequenced by tandem mass spectrometry and determined to be residues 273–305 of CII (QTGEPGIAGFKGEQGPKGEPGPAGVQGAPGPAG) with four hydroxylated residues ($(CII_{273-305})_{4OH}$; Supplementary Fig. 8 and Table 3), showing the core $CII_{282-289}$ epitope underlined. The other peptides in this cluster contained the same core epitope and they were consistent in mass with post-translational modification (PTM) variants of this peptide that are expected to be present (Fig. 2c)[20,21]. A baculovirus-encoded conotoxin-like peptide (CL_{13-23}, FAALNAQHVLA) (Supplementary Fig. 9) was present at m/z ~1153 Da.

This peptide likely binds to a portion of DR1 molecules during expression within baculovirus-transduced insect cells (Fig. 2a, b)[22]. As before, DR1 transgenic mice were immunized with CII protein/CFA and lymphoid cells were used in a recall proliferation assay using $CII_{280-294}$, the CII protein, or $CLIP_{89-105}$. We observed strong dose-dependent responses to $CII_{280-294}$ and to the CII protein, whereas no response to $CLIP_{89-105}$ was observed (Fig. 2d).

De novo identification of T Cell epitopes from rHA1 of H5N1

We next attempted *de novo* identification of immunodominant epitopes of HA1 protein of H5N1. Highly pathogenic influenza A H5N1 viruses caused disease outbreaks in poultry, wild birds, and humans in Asian countries with a fatality rate of approximately 60%. The experimental set up was identical to the chimeric rHA1 above. The Native or heat-denatured HA1 from the H5N1 Influenza strain A/Vietnam/1203/2004 (H5N1-rHA1) was used as antigen. Two unique H5N1-rHA1 derived peptide species were detected at m/z 1814.82 and 2201.00 compared to background spectrum regardless of whether the protein was native or heat-denatured (Fig. 3a–c). Microsequencing identified the sequences as $HA_{259-274}$ (SNGNFIAPEYAYKIVK) (Fig. 3d), and $HA_{259-278}$ (SNGNFIAPEYAYKIVKKGDS) (Supplementary Fig. 10) at m/z 1814.82 and m/z 2201.00, respectively, both sharing the core 16-residue sequence (underlined). Thus, a single core epitope was selected from intact H5N1-rHA1.

Immunodominance of the identified peptide was evaluated by its ability to induce T cell responses. DR1 transgenic mice were immunized with the H5N1-rHA1 in CFA. The lymphoid cells were used in recall proliferation and cytokine production assays using titrating doses of $HA_{259-274}$, $CLIP_{89-105}$, or H5N1-rHA1. We observed strong dose-dependent proliferation responses to $HA_{259-274}$ and to the H5N1-rHA1 but not to $CLIP_{89-105}$ (Fig. 4a). Proliferation in response to $HA_{259-274}$ was strikingly similar in magnitude to the response elicited by the whole H5N-rHA1. IL-2 production in response to the H5N1-rHA1 was greater than the $HA_{259-274}$ at 24 h but its production reached similar levels as H5N1-rHA1 at 48 h. Overall, a significantly higher amount of IFN-γ production was detected in comparison to IL-2 as determined by ELISA. IFN-γ production in response to stimulation by the H5N1-rHA1 and by the $HA_{259-274}$ was similar in magnitude both at 48 h and 72 h (Fig. 4b, c).

DR1 transgenic mice were immunized either with the H5N1-rHA1 in CFA or CFA alone. Freshly isolated lymphoid cells were stained with $HA_{259-274}$/DR1 or CLIP/DR1 tetramers, CD4+, and CD44+ (activation marker) (Fig. 4d and Supplementary Fig. 11). Cells positive for $HA_{259-274}$/DR1 tetramer comprised ∼0.01% of the CD4+CD44+ population in control group versus ∼0.1% among cells isolated from H5N1-rHA1/CFA immunized. After a 7 day *in vitro* expansion with H5N1-rHA1, the $HA_{259-274}$/DR1 tetramer-positive population comprised nearly 3% of the CD4+CD44+ population (Fig. 4e and Supplementary Fig. 12). In all, *in vitro* stimulation with the H5N1-rHA1 for seven days resulted in enrichment of $HA_{259-274}$ specific CD4+CD44+T cells. These combined results from proliferation assays, cytokine measurements and tetramer staining indicate that the minimalist system identified a major T cell epitope that activates CD4+ T cells specific for the H5N1-rHA1.

De novo detection of T Cell epitopes from a malaria antigen

Malaria is one of the main global causes of death from infectious disease, resulting in more than 300 million clinical cases and more than two million deaths annually. *Plasmodium falciparum* liver-stage antigen 1 (LSA-1) is exclusively expressed in malaria-infected hepatocytes at a pre-erythrocytic stage. LSA-1 contains a large central 17-amino-acid repeat region flanked by two highly conserved N- and C-terminal regions shown to contain B- and T-cell stimulating epitopes[23],[24]. LSA-NRC, a modified recombinant form of LSA-1, was designed as a vaccine against pre-erythrocyte stage malaria[25] (Supplementary Table 4).

LSA-NRC was incubated with DR1 and DM followed by cathepsins B, H, and S. Peptides were eluted from DR1 and analyzed by mass spectrometry. Comparing the spectra of experimental (Top) and background groups (Bottom), five unique peptide species were detected at m/z ∼1747,1877, 2278, 2295, and 3475 (Fig. 5a, b) and their sequences were determined by CID. The sequence of the peptides at m/z ∼1747 and 1877 were determined to be $LSA_{325\text{-}338}$ (YDNFQDEENIGIYK) (Supplementary Fig. 13) and $LSA_{325–339}$ (YDNFQDEENIGIYKE), respectively. The peptides at m/z ∼2278 and 2295 were identified as post-translation modification variants of a single peptide, $LSA_{434–453}$ (EDITKYFMKLGGSGSPHHHH) (Fig. 5c), with its methionine residue either unmodified or oxidized (16 Da difference). Although the intensity of peptide peak at m/z ∼3475 was weak and indistinguishable from the background, fragmentation data determined its sequence to be $LSA_{197–226}$ (LPSENERGYYIPHQSSLPQDNRGNSRDSKE) (Fig. 5b).

In vivo verification in DR1 transgenic mice confirmed $LSA_{436\text{-}449}$ as immunodominant by T cell proliferation and IL-2 and IFN-γ production (Fig. 5d, e and Supplementary Fig. 14). Here, we also used two peptides from the library of peptides originally synthesized as potentially active based on conventional 15mer overlapping peptide library design. The two peptides were $LSA_{323–337}$ (VQYDNFQDEENIGIY) and $LSA_{429–443}$ (VDELSEDITKYFMKL), which were specifically chosen for containing partial sequences from the epitopes selected by the minimalist system (underlined). Those two peptides did not recall any responses.

To confirm the potential clinical utility of this system, we used donor-derived T cells from volunteers who, as part of a malaria vaccine trial, had been immunized twice with LSA-NRC in a liposome-based formulation[26]. We tested samples from 8 individuals whose CD4+ T cells had responded vigorously to the LSA protein *in vitro*. PBMC were obtained pre-immunization and after second immunization with LSA-NRC. These PBMC were used for proliferation by stimulation *in vitro* with different concentrations of $LSA_{436\text{-}449}$, $CLIP_{89\text{-}105}$, and LSA-NRC. We found that while all eight post-immune samples responded to LSA protein, only two individuals (#8013 and #2060) responded vigorously to our epitope and none responded to $CLIP_{89–105}$ stimulation. HLA-DRB1 typing revealed that the two responding samples to the $LSA_{436–449}$ epitope came from $DR1^{positive}$ individuals (DR B1*0101) (Fig. 5f). The six non-responding samples were from DR1 negative individuals carrying DR4, 8,11,12, 13,14,15, or 17 alleles. These results indicate that $LSA_{436\text{-}449}$ is a DR1-restricted epitope, consistent with the fact that the minimalist system used to identify

this epitope utilized DR1. Importantly, these results demonstrate that the minimalist system is capable of predicting epitopes that human T cells recognize.

Discussion

We have established a minimalist antigen processing system composed of DR1, DM, and cathepsins B, H and S that identifies physiologically relevant immunodominant epitopes. We confirmed immunodominance in HLA-DR1-transgenic mice immunized and in human. Immunodominance of *de novo* epitopes of the H5N1-rHA1 or LSA-1 were confirmed by virtue of their capacity to induce T cell activation at levels similar to that induced by the whole proteins. Our results suggest that immunodominance is dictated by the combined interplay of cathepsins, HLA-DM and the MHC II during antigen processing.

The finding that so few components of MHC class II antigen processing can process full-length antigens and yield the physiologically selected immunodominant epitopes from protein antigens is striking. In addition to DM, the cell-free system relies on endo- and exopeptidases. We find that the inclusion of cathepsin S alone, in the absence of the C- and N-terminal exopeptidases used in the system was not adequate for successful detection of the dominant epitopes in rHA1 (data not shown). On the contrary, a combination of cathepsins B, H, and S, was always sufficient in all four antigens we studied here. Cathepsins B and H in the absence of cathepsin S, led to the capture of $HA_{306\text{-}318}$ and $HA_{298\text{-}317}$ from our model influenza antigen. There has been little appreciation for the roles of exopeptidases, cathepsin B and cathepsin H, in antigen processing. Cathepsin B deficient mice were shown not to have any significant defects in antigen processing[27]. Also, little information if any is available regarding a role for cathepsin H in antigen processing. As such, it was quite surprising that the combined use of cathepsins B and H led to the proper processing of protein antigens and the selection of immunodominant epitopes while the presence or absence of cathepsin S did not change the overall outcome of epitope selection for rHA1 of the A/Texas1/77 strain. Thus, a new important role for cathepsin H in conjunction with cathepsin B is defined here.

Our system relies on mass spectrometry for identifying epitopes, a technology that has been previously used for characterizing antibody epitopes[28] or for studying MHC-associated peptides[29]-[31]. Being so precise, mass spectrometry has the advantage of having the capacity of identifying post-translationally modified epitopes. Post-translational modification of antigenic peptides, specifically citrullination, has been shown to be important in activation of $CD4^+$ and $CD8^+$ T cells by MHC class II[32] and class I[33], respectively. Because our assay uses actual protein antigens, it intrinsically takes into account the effects that post-translational modifications may have on epitope selection.

In summary, we have established a cell-free antigen processing system that can successfully identify immunodominant epitopes from protein antigens that are recognized by CD4 T cells from immunized HLA-DR1 expressing humans and mice. Although in this report we have presented data using HLA-DR1 only, the system can be expanded to other MHC class II alleles and can be developed into a high throughput screening system. Our observations indicate that this system closely mimics antigen processing *in vivo*. If used for application

purposes, the system would help to reduce cost and increase effectiveness in determining immunodominant epitopes of antigens from infectious agents, autoimmune antigens, or cancer. The identified epitopes can be incorporated into MHC class II multimers[34] for detection of T cells reactive to the epitopes as correlates of immunity for diagnosis and/or monitoring disease progression in humans. Finally, this system can be used in design of peptide vaccines, or for evaluation of proper recombinant vaccine designs before human trial studies to ascertain that artifacts introduced to the recombinant vaccines do not interfere with the recognition of the wild type epitopes.

Methods

Peptides

$HA_{306-318}$ peptide (PKYVKQNTLKLAT), human short $CLIP_{89-105}$ peptide (KMRMATPLLMQALPM), and $CII_{280-294}$ (AGFKGEQGPKGEPGP) were synthesized by Global peptide (Fort Collins, CO). $HA_{298-317}$ peptide (INSSLPYQNIHPVTIGECPKY), $HA_{259-274}$ peptide (SNGNFIAPEYAYKIVK), and $LSA_{436-449}$ (ITKYFMKLGGSGSP) were synthesized by Elim Biopharmaceuticals (Hayward, CA) at >85% purity as analyzed by reversed phase HPLC, and their identities were confirmed by mass spectrometry. Overlapping peptides, $LSA_{323-337}$ (VQYDNFQDEENIGIY) and $LSA_{429-443}$ (VDELSEDITKYFMKL) were synthesized by GenScript.

Production of Recombinant Proteins

Soluble HLA-DR1*0101 and HLA-DM were produced in baculovirus-transduced insect cells as described previously[22,35]. Recombinant influenza hemagglutinin (rHA1) was produced in *E. coli* transformed with an expression vector for a dual 6xhistidine-tagged influenza hemagglutinin. The protein contains residues 12∼340 of the hemagglutinin of Influenza strain A/PR/8/34 HA gene with a MRGSHHHHHHTDPSSRSA tag on the N-terminus and a ACPKYVKQNTLKLATGMRKLHHHHHHN tag on the C-terminus (the underlined residues comprise the $HA_{306-318}$ epitope from Influenza strain A/Texas/1/77). Following affinity purification from bacterial lysate with Ni-NTA-charged agarose resin (Ni-NTA Superflow, Qiagen), the protein was refolded by stepwise dialysis into PBS/10% sucrose and then stored at −80 °C. H5N1 rHA1 from strain A/Vietnam/1203/2005 was purified from 293 cells (eEnzyme). The expression, purification, and biochemical and immunological characterization of E. coli produced, GMP manufactured LSA-NRC antigen has been described previously[25].

Experimental set up for the cell free processing system

DR1, antigen, and DM were incubated in citrate phosphate buffer (pH 5.0) at 37 °C for 3 h, after which cathepsin B (bovine spleen, Sigma) and cathepsin H (human liver, Calbiochem) or cathepsin B, cathepsin H, and cathepsin S (human, Calbiochem) were added with 6 mM L-Cysteine and 4 mM EDTA for an additional 2-3 h. After this, the pH was adjusted to 7.5, 10 mM iodoacetamide was added, and DR1 was immunoprecipitated with Sepharose conjugated with DR1 specific antibody. Bound peptides were eluted with TFA[36], filtered through a 10 kDa MWCO Microcon (Millipore), and lyophilized.

Mass Spectrometry

Lyophilized samples were re-suspended in 5-10 μl of 50% ethanol, 50% water, 0.2%TFA. 0.5 μl of re-suspended sample was spotted, dried, and covered with 0.5 μL matrix (40-50 mg/ml 2,5-dihydroxybenzoic acid or 2.5 mg/ml α-cyano-4-hydroxycinnamic acid in 50% ethanol, 50% water, 0.15%TFA). Samples were analyzed on the vMALDI-LTQ mass spectrometer (ThermoFisher, San Jose, California), using Tune Plus 2.2 Xcalibur 2.0 SR2 vMALDI LTQ 2. Full MS scans (m/z 1100-4000 Da) were acquired to select peptides of interest for identification by CID fragmentation (MS^2).

Full MS spectra were manually inspected to find the peptides of interest. MS^n data were collected on peptides of interest and searched with Bioworks 3.3.1 SP1 (ThermoFisher) against a custom-built database containing all protein components present in the sample, using the following criteria: no enzyme, mass tolerance of 2 Da for MS^1, mass tolerance of 1Da for MS^2 and MS^3, and with allowances for variable carboamidomethylated of cysteine and oxidation of methionine. For type II collagen, samples were analyzed on the Axima-CFR MALDI-TOF mass spectrometer (Kratos Analytical, Shimadzu) with data acquired in reflectron mode. Data were analyzed with LAUNCHPAD™ (Shimadzu/Kratos Analytical).

Proliferation and cytokine production assay

DR1 (DR B1*0101)-transgenic mice[19] were immunized with 50 μg proteins in CFA in the base of the tail. After 8-10 d, the draining lymph nodes were harvested and the cells were incubated with a range of peptide and protein concentrations for 3 d before adding [^{3}H] thymidine (Amersham). The cells were harvested and counted after a further incubation of 18-20 h, and the incorporated radioactivity was measured by Packard Matrix 96 beta counter. For each triplicate, lymphocyte proliferation was recorded as the mean counts per minute. For the cytokine assays, supernatants were collected after 24 h, 48 h, and 72 h incubation. IL-2 and IFN-γ concentrations were measured by enzyme-linked immunosorbent assay (R&D System). Optical density was measured with an ELISA reader (Dynex Technologies) with a test wavelength of 450 nm and a reference wavelength of 570 nm.

Human PBMC proliferation assay

Full description of the screening and enrolment of the volunteers, preparation of LSA-NRC, the adjuvant used for immunization etc are described in Cummings et al[26]. Peripheral blood mononuclear cells (PBMC) from both pre-immunization and LSA-NRC post immunization[26] were collected and tested for proliferation. PBMC from 8 different donors were incubated with a different concentration of peptides, or protein for 3 d before adding [^{3}H] thymidine. The cells were harvested and counted after a further incubation of 18-20 h, and the incorporated radioactivity was measured. The stimulation index was calculated by dividing antigen-induced proliferation by background proliferation (media). HLA-typing of donors was done in Immunogenetics laboratory at JHU.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

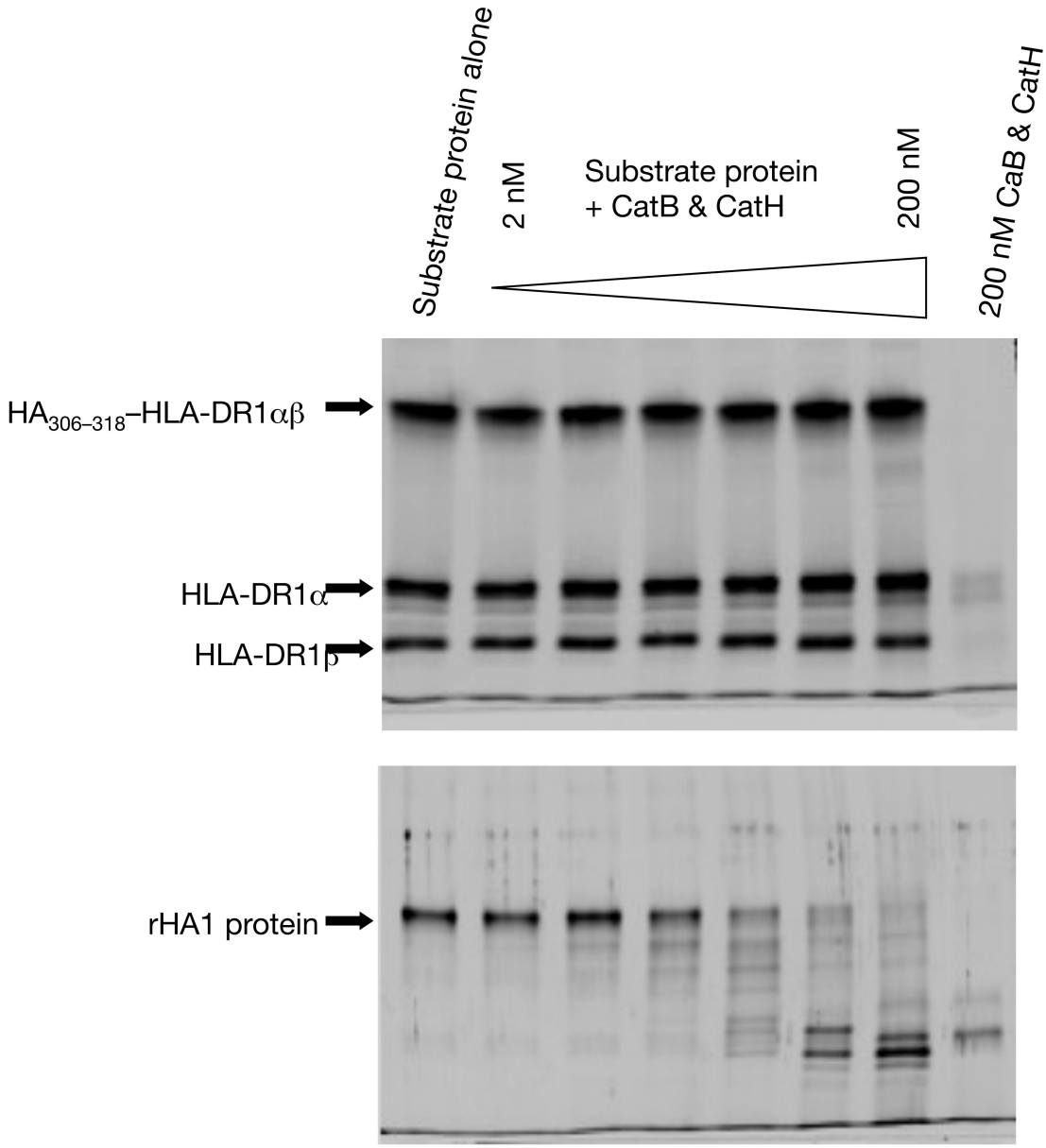
Acknowledgments

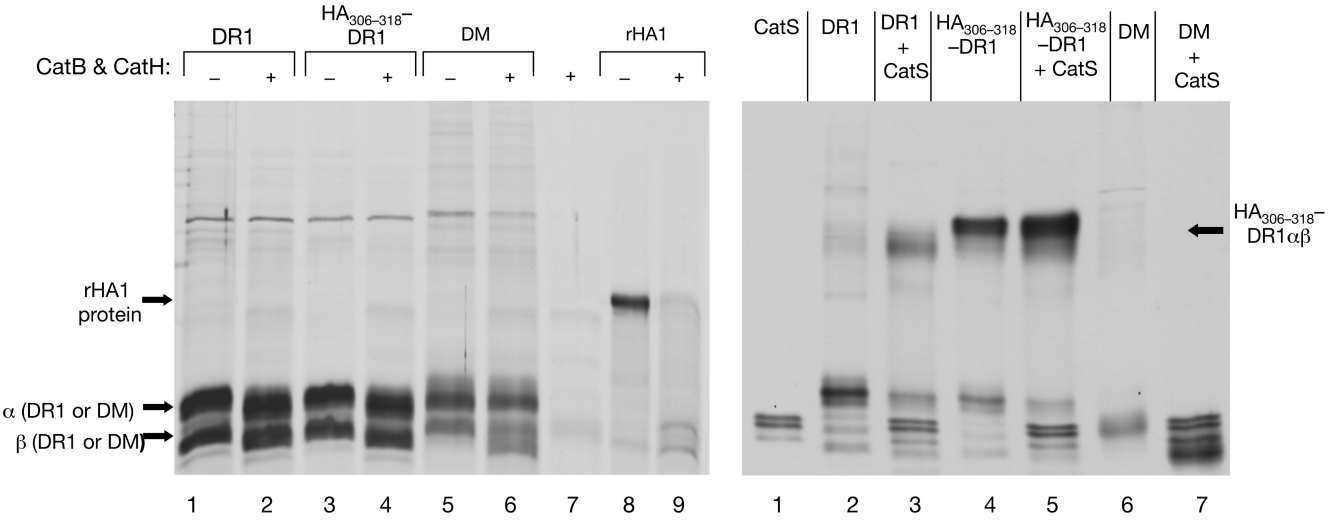
Authors wish to thank S. Landry, A. Hamad, J. Pomerantz and K. Narayan for reading of the manuscript and insightful discussions, S. Khoruzhenko for protein production, S. Kalb-Ramirez and D. Wang for the initial mass spectrometry, F. Korangy and D. Pardoll (Johns Hopkins University) for E. coli transformed with an expression vector influenza hemagglutinin, D. Zaller (MERCK) for the original DR1 Tg mice, L. Egner (WRAIR) for testing the human samples, and US National Institute of Health Tetramer Facility for providing DR1/CLIP monomers. This work was supported by R01 grants AI063764, and GM053549, a grant from Johns Hopkins Malaria Research Institute to SS-N, an NSF predoctoral award to IZH, and a JHU Rheumatology T32 Fellowship to AK. IZH and AK designed and executed experiments analyzed data and wrote the paper, RJC, WG, TB, RNC provided mass spectrometry data and analyses, KW cloned DM, SKD contributed to *in vivo* testing, RS, DEL, UK did LSA-1 experiments in humans, and SS-N supervised the project, obtained funding and wrote the paper.

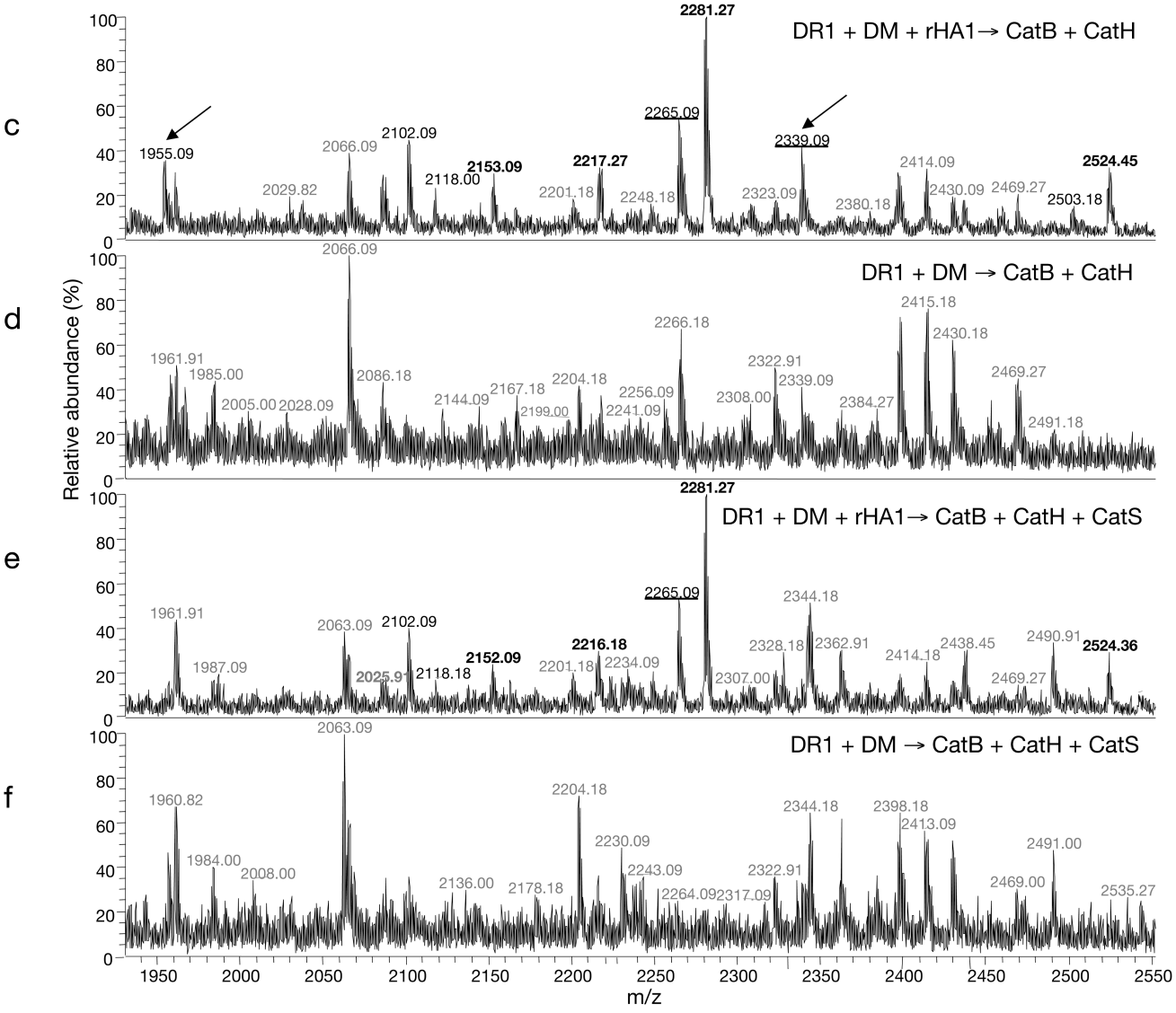
References

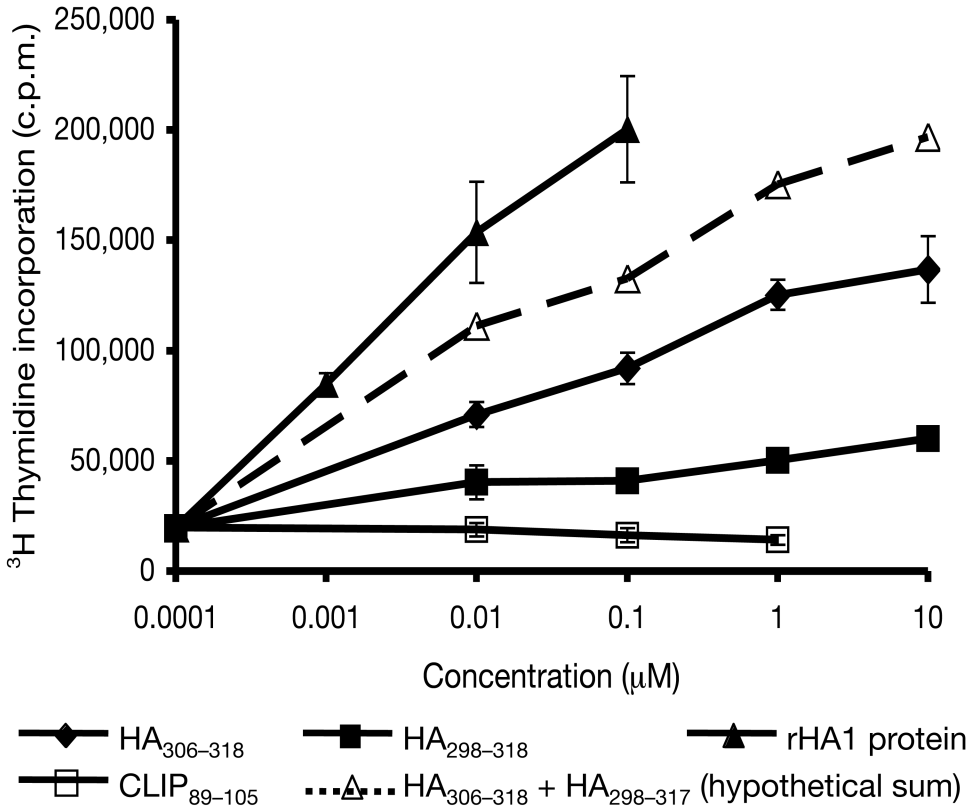
1. Sercarz EE, et al. Dominance and crypticity of T cell antigenic determinants. Annu Rev Immunol. 1993; 11:729–766. [PubMed: 7682817]
2. Sercarz EE. Immune focusing vs diversification and their connection to immune regulation. Immunol Rev. 1998; 164:5–10. [PubMed: 9795758]
3. Lamb JR, Eckels DD, Lake P, Woody JN, Green N. Human T-cell clones recognize chemically synthesized peptides of influenza haemagglutinin. Nature. 1982; 300:66–69. [PubMed: 6982419]
4. Rosloniec EF, et al. An HLA-DR1 transgene confers susceptibility to collagen-induced arthritis elicited with human type II collagen. J Exp Med. 1997; 185:1113–1122. [PubMed: 9091584]
5. Jensen PE. Recent advances in antigen processing and presentation. Nat Immunol. 2007; 8:1041–1048. [PubMed: 17878914]
6. Amigorena S, Drake JR, Webster P, Mellman I. Transient accumulation of new class II MHC molecules in a novel endocytic compartment in B lymphocytes [see comments]. Nature. 1994; 369:113–120. [PubMed: 8177316]
7. Kropshofer H, et al. Editing of the HLA-DR-peptide repertoire by HLA-DM. Embo J. 1996; 15:6144–6154. [PubMed: 8947036]
8. Chou CL, Sadegh-Nasseri S. HLA-DM recognizes the flexible conformation of major histocompatibility complex class II. J Exp Med. 2000; 192:1697–1706. [PubMed: 11120767]
9. Riberdy JM, Newcomb JR, Surman MJ, Barbosa JA, Cresswell P. HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. Nature. 1992; 360:474–477. [PubMed: 1448172]
10. Sloan VS, et al. Mediation by HLA-DM of dissociation of peptides from HLA-DR. Nature. 1995; 375:802–806. [PubMed: 7596415]
11. Sadegh-Nasseri S, Chen M, Narayan K, Bouvier M. The convergent roles of tapasin and HLA-DM in antigen presentation. Trends Immunol. 2008; 29:141–147. [PubMed: 18261958]
12. Pluger EB, et al. Specific role for cathepsin S in the generation of antigenic peptides in vivo. Eur J Immunol. 2002; 32:467–476. [PubMed: 11813165]
13. Riese RJ, et al. Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. Immunity. 1996; 4:357–366. [PubMed: 8612130]
14. Chapman HA. Endosomal proteases in antigen presentation. Curr Opin Immunol. 2006; 18:78–84. [PubMed: 16338127]
15. Pisoni RL, Acker TL, Lisowski KM, Lemons RM, Thoene JG. A cysteine-specific lysosomal transport system provides a major route for the delivery of thiol to human fibroblast lysosomes: possible role in supporting lysosomal proteolysis. J Cell Biol. 1990; 110:327–335. [PubMed: 2404990]
16. Maric M, et al. Defective antigen processing in GILT-free mice. Science. 2001; 294:1361–1365. [PubMed: 11701933]
17. Boja ES, Fales HM. Overalkylation of a protein digest with iodoacetamide. Anal Chem. 2001; 73:3576–3582. [PubMed: 11510821]
18. Schey KL, Finley EL. Identification of peptide oxidation by tandem mass spectrometry. Acc Chem Res. 2000; 33:299–306. [PubMed: 10813874]

19. Narayan K, et al. Staphylococcal enterotoxin A induces small clusters of HLA-DR1 on B cells. PLoS One. 2009; 4:e6188. [PubMed: 19587800]
20. Van den Steen PE, et al. Cleavage of denatured natural collagen type II by neutrophil gelatinase B reveals enzyme specificity, post-translational modifications in the substrate, and the formation of remnant epitopes in rheumatoid arthritis. Faseb J. 2002; 16:379–389. [PubMed: 11874987]
21. Van den Steen PE, et al. Generation of glycosylated remnant epitopes from human collagen type II by gelatinase B. Biochemistry. 2004; 43:10809–10816. [PubMed: 15311942]
22. Stern LJ, Wiley DC. The human class II MHC protein HLA-DR1 assembles as empty alpha beta heterodimers in the absence of antigenic peptide. Cell. 1992; 68:465–477. [PubMed: 1371238]
23. Krzych U, et al. T lymphocytes from volunteers immunized with irradiated Plasmodium falciparum sporozoites recognize liver and blood stage malaria antigens. J Immunol. 1995; 155:4072–4077. [PubMed: 7561118]
24. Fidock DA, et al. Plasmodium falciparum liver stage antigen-1 is well conserved and contains potent B and T cell determinants. J Immunol. 1994; 153:190–204. [PubMed: 7515922]
25. Hillier CJ, et al. Process development and analysis of liver-stage antigen 1, a preerythrocyte-stage protein-based vaccine for Plasmodium falciparum. Infect Immun. 2005; 73:2109–2115. [PubMed: 15784552]
26. Cummings JF, et al. Recombinant Liver Stage Antigen-1 (LSA-1) formulated with AS01 or AS02 is safe, elicits high titer antibody and induces IFN-gamma/IL-2 CD4+ T cells but does not protect against experimental Plasmodium falciparum infection. Vaccine. 2009
27. Deussing J, et al. Cathepsins B and D are dispensable for major histocompatibility complex class II-mediated antigen presentation. Proc Natl Acad Sci USA. 1998; 95:4516–4521. [PubMed: 9539769]
28. Parker CE, Papac DI, Trojak SK, Tomer KB. Epitope mapping by mass spectrometry: determination of an epitope on HIV-1IIIB p26 recognized by a monoclonal antibody. J Immunol. 1996; 157:198–206. [PubMed: 8683115]
29. Chicz RM, et al. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. Nature. 1992; 358:764–768. [PubMed: 1380674]
30. Hunt DF, et al. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science. 1992; 255:1261–1263. [PubMed: 1546328]
31. Strug I, et al. Vaccinia peptides eluted from HLA-DR1 isolated from virus-infected cells are recognized by CD4+ T cells from a vaccinated donor. J Proteome Res. 2008; 7:2703–2711. [PubMed: 18507432]
32. Ireland J, Herzog J, Unanue ER. Cutting edge: unique T cells that recognize citrullinated peptides are a feature of protein immunization. J Immunol. 2006; 177:1421–1425. [PubMed: 16849447]
33. Beltrami A, et al. Citrullination-dependent Differential Presentation of a Self-peptide by HLA-B27 Subtypes. J Biol Chem. 2008; 283:27189–27199. [PubMed: 18650441]
34. Gebe JA, Kwok WW. Tracking antigen specific CD4+ T-cells with soluble MHC molecules. Methods Mol Med. 2007; 136:39–50. [PubMed: 17983139]
35. Narayan K, et al. HLA-DM targets the hydrogen bond between the histidine at position beta81 and peptide to dissociate HLA-DR-peptide complexes. Nat Immunol. 2007; 8:92–100. [PubMed: 17143275]
36. Suri A, et al. In APCs, the autologous peptides selected by the diabetogenic I-Ag7 molecule are unique and determined by the amino acid changes in the P9 pocket. J Immunol. 2002; 168:1235–1243. [PubMed: 11801660]
37. Natarajan SK, Stern LJ, Sadegh-Nasseri S. Sodium dodecyl sulfate stability of HLA-DR1 complexes correlates with burial of hydrophobic residues in pocket 1. J Immunol. 1999; 162:3463–3470. [PubMed: 10092802]









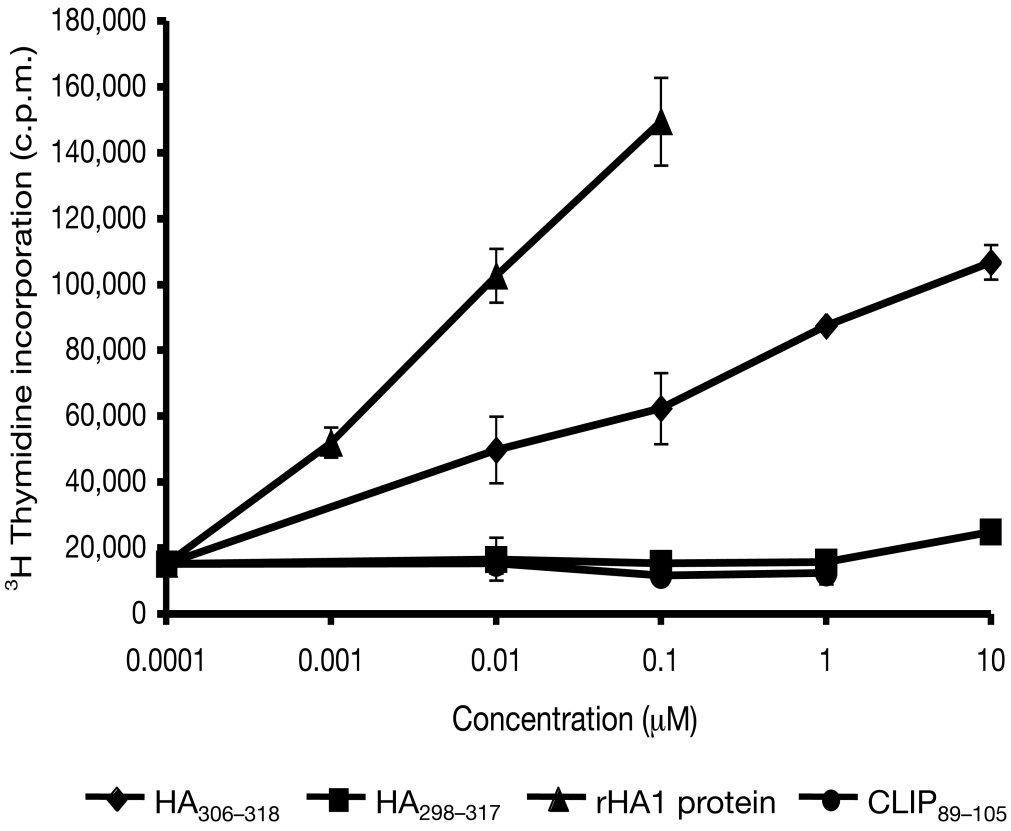
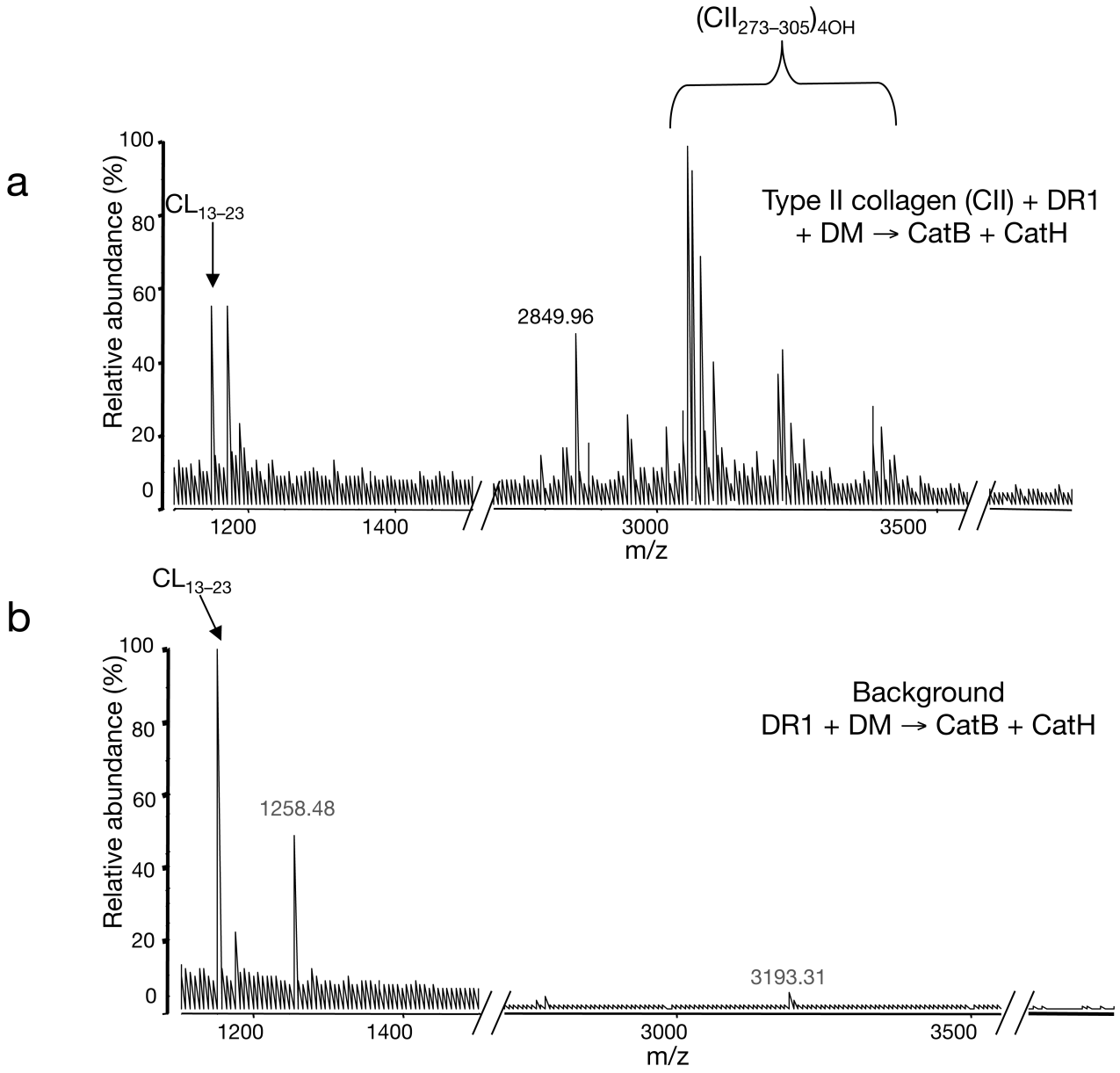
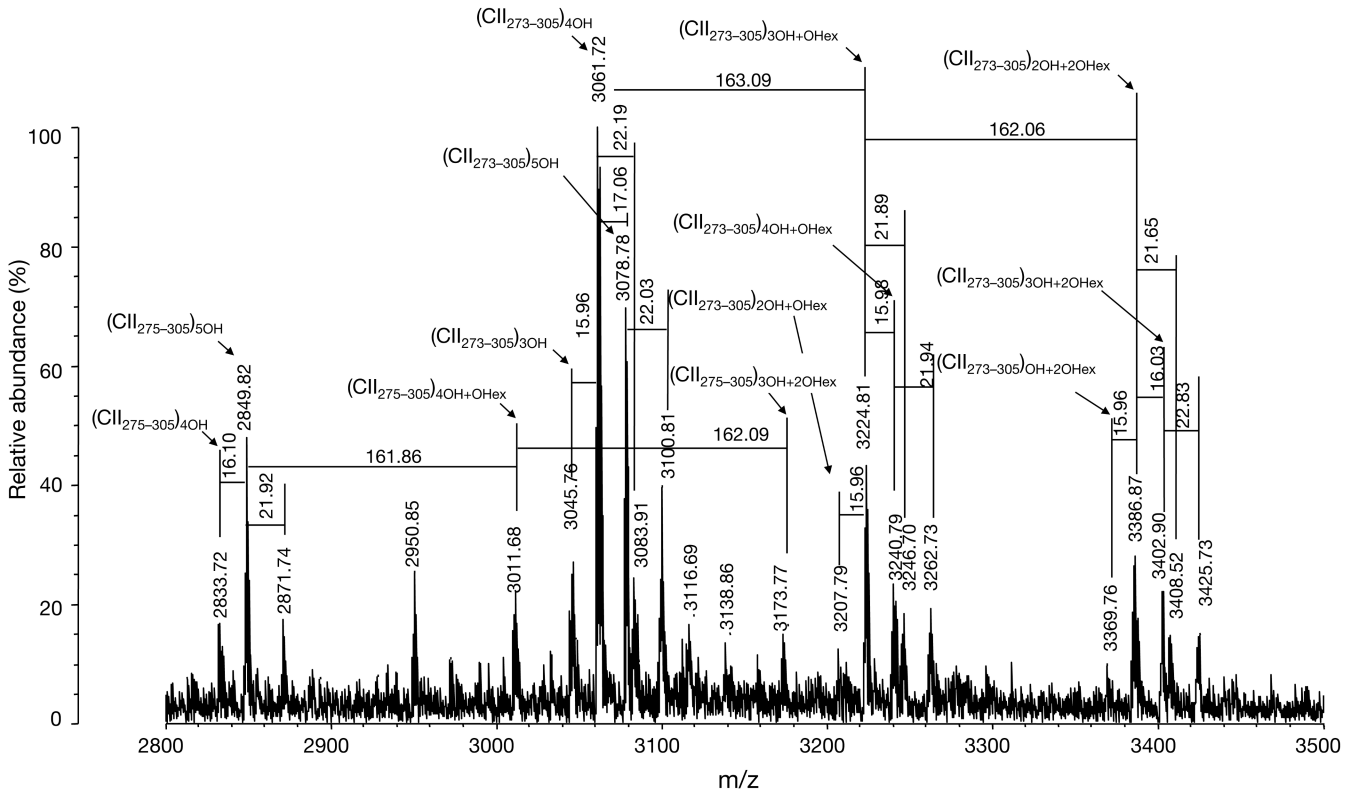


Figure 1. Sensitivity to cathepsins and identification of rHA1-derived peptides by Mass Spectrometry

(**a**) Sensitivity to CatB and CatH. $HA_{306\text{-}318}$–DR1 (top) and rHA1 (bottom) were treated with CatB and CatH at various concentrations of each enzyme. Samples were resolved by gentle SDS-PAGE where samples were not boiled to preserve peptide–DR1αβ complex[37]. (**b**) Sensitivity of DM and DR1 to Cathepsins. Left: Conventional SDS-PAGE of empty DR1 (lanes 1-2), pre-formed $HA_{306\text{-}318}$–DR1 (lanes 3-4), and DM (lanes 5-6) incubated in the presence or absence of 200 nM CatB and CatH, rHA1 (lanes 8-9) served as a positive control for digestion by CatB and CatH. Right: Gentle SDS-PAGE (12% Acrylamide, silver-stained) of empty DR1 (lane 3), $HA_{306\text{-}318}$–DR1 (lane 5), and DM (lane 7) treated with 100 nM CatS. (**c**-**f**) Mass spectra of rHA1 derived peptides eluted from DR1 are shown between m/z 1950-2550 Da. **b** is background for **a**, and **d** is background for **c**. Mass species in bold represent rHA1 fragments eluted from DR1 containing $HA_{306\text{-}318}$ epitope (Supplementary Fig. 3–4), underlined represent other rHA1-derived peptides, in grey are the background peaks. Experiments were repeated more than three times. (**g**, **h**) *In vitro* proliferation of rHA1-immunized cells as measured by [^{3}H] thymidine incorporation (two individual mice out of five tested).





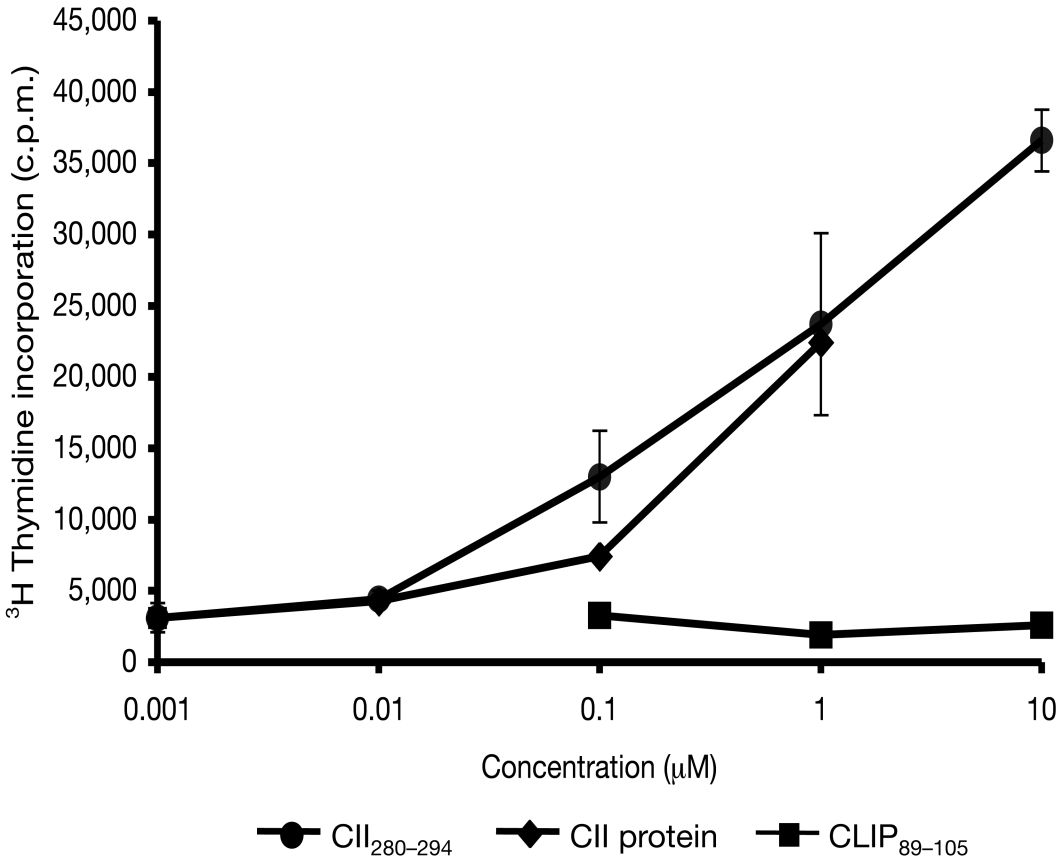
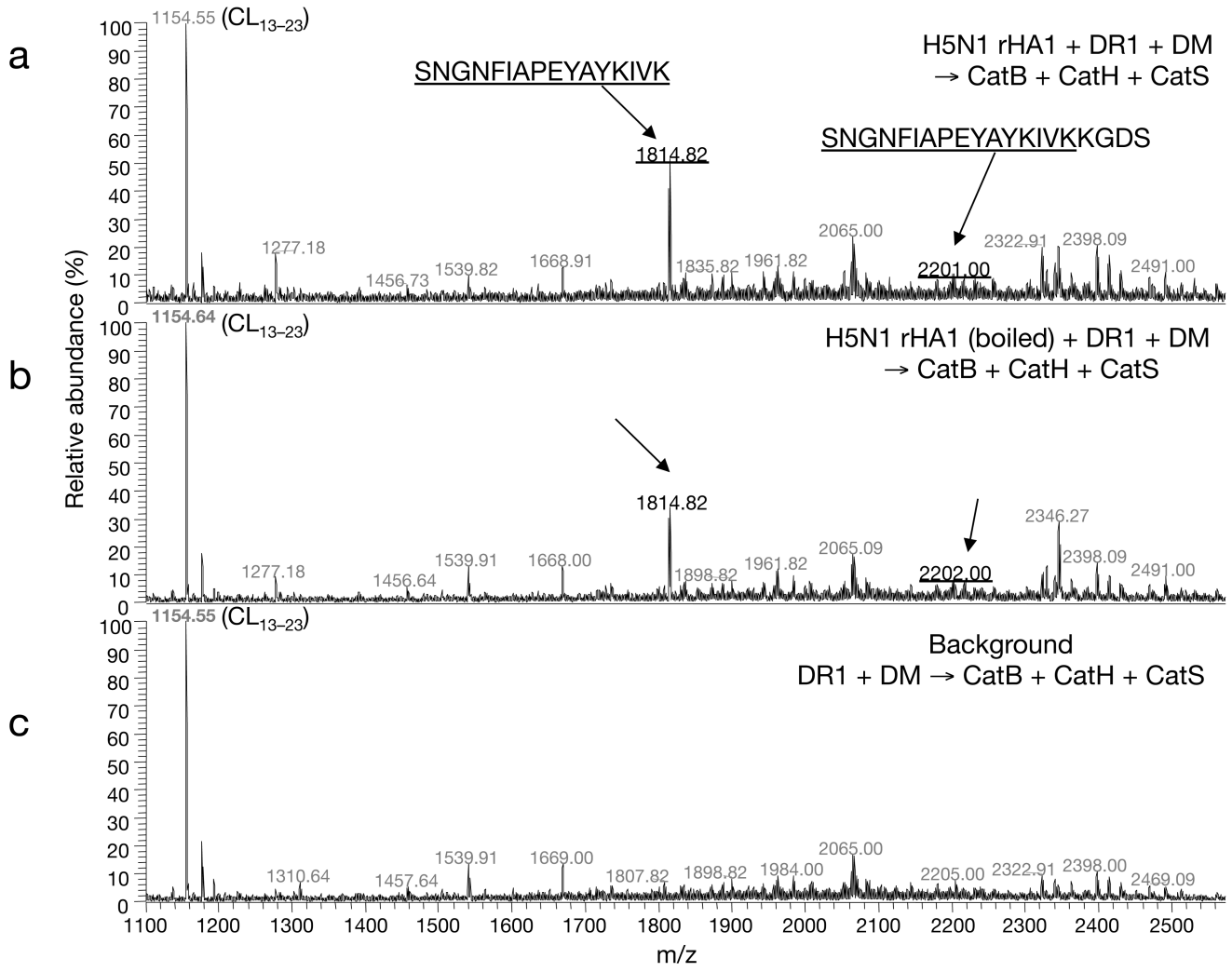


Figure 2. Identification of type II collagen-derived peptides eluted from DR1

(**a**, **b**) Mass spectra of CII derived-peptides eluted from DR1. (**a**) DR1 was incubated with the following components: Matrix Metalloproteinase 9 (MMP9)-fragmented bovine type II collagen (CII)[20], DM, CatB, and CatH (**b**) the negative control reactions carried out without including MMP9-fragmented CII. (CII $_{273\text{-}305})_{4OH}$ represents the sequence of collagen peptide containing four hydroxylation of Pro or Lys (QTGEP$_{OH}$GIAGFKGEQGPK$_{OH}$GEP$_{OH}$GPAGVQGAP$_{OH}$GPAG). The experiment was repeated more than three times. (**c**) Expanded spectrum of (**a**) between m/z 2800 and 3500 Da. Peptide modification: Hexose: +162 Da, Hydroxylation: +16 Da, Na^+ adduct: +22Da (**d**) Proliferation of cells immunized with native CII protein in CFA incubated with titrating doses of CII$_{280\text{-}294}$ (AGFKGEQGPKGEPGP), CLIP$_{89\text{-}105}$, and heat denatured CII protein *in vitro* as measured by [^{3}H] thymidine incorporation (one out of three representative individual mice tested).



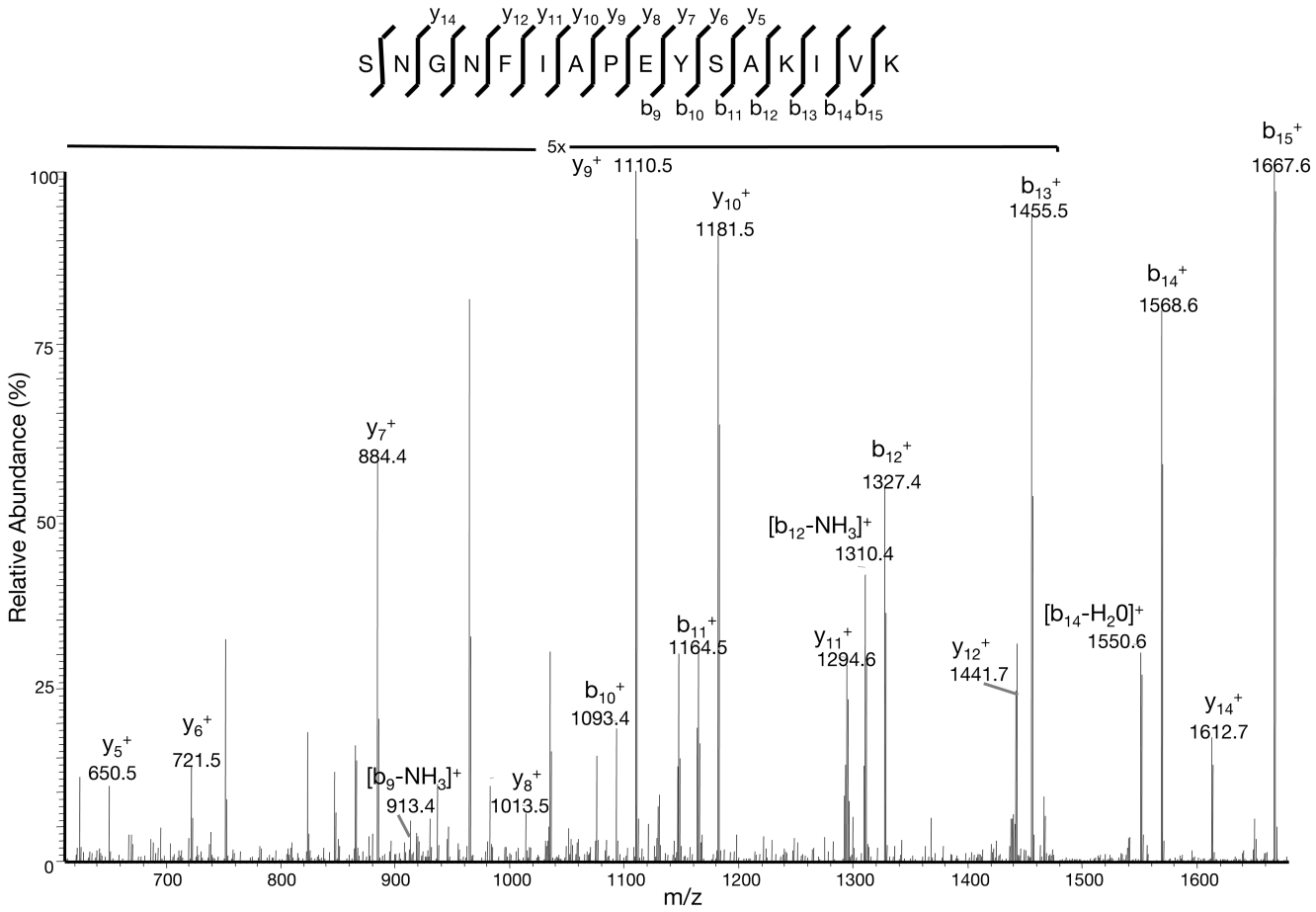
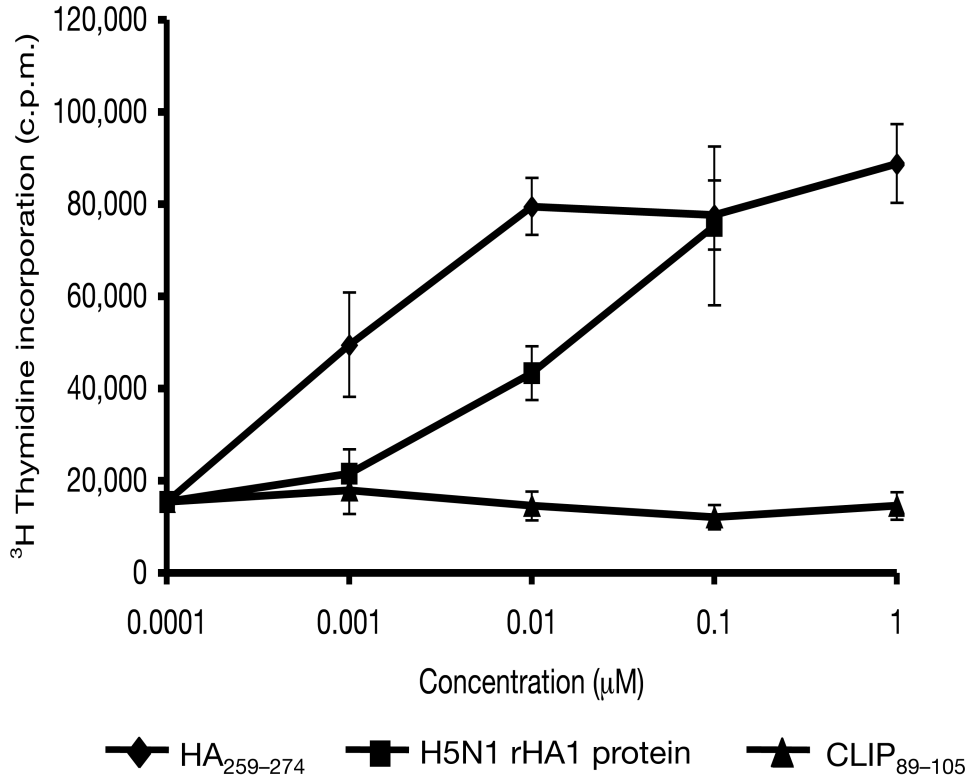
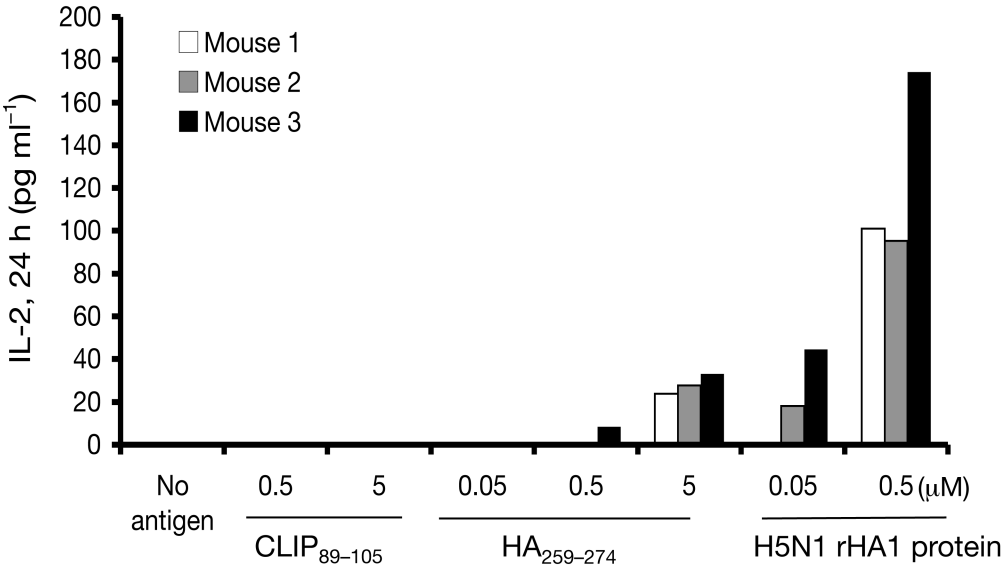
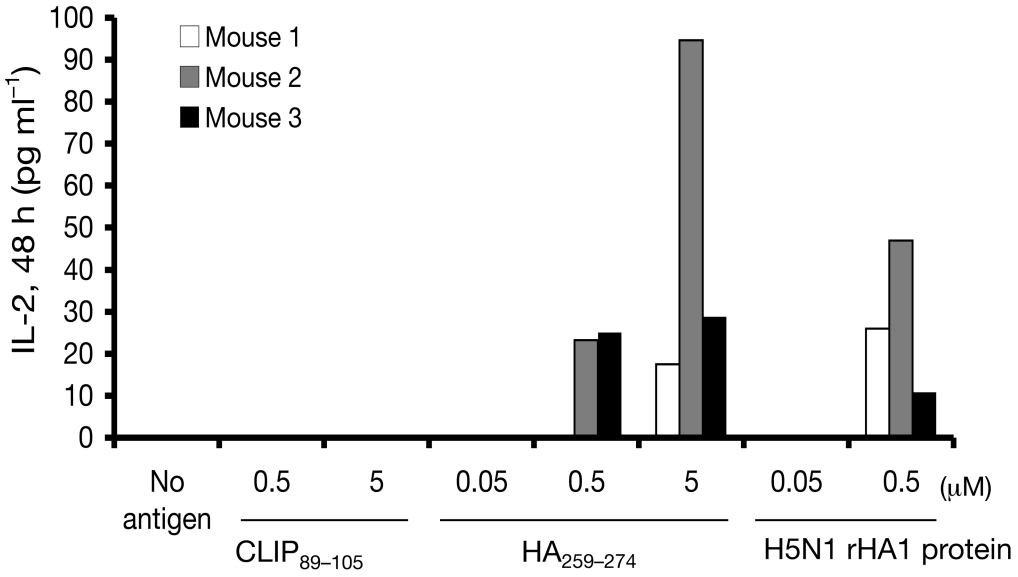


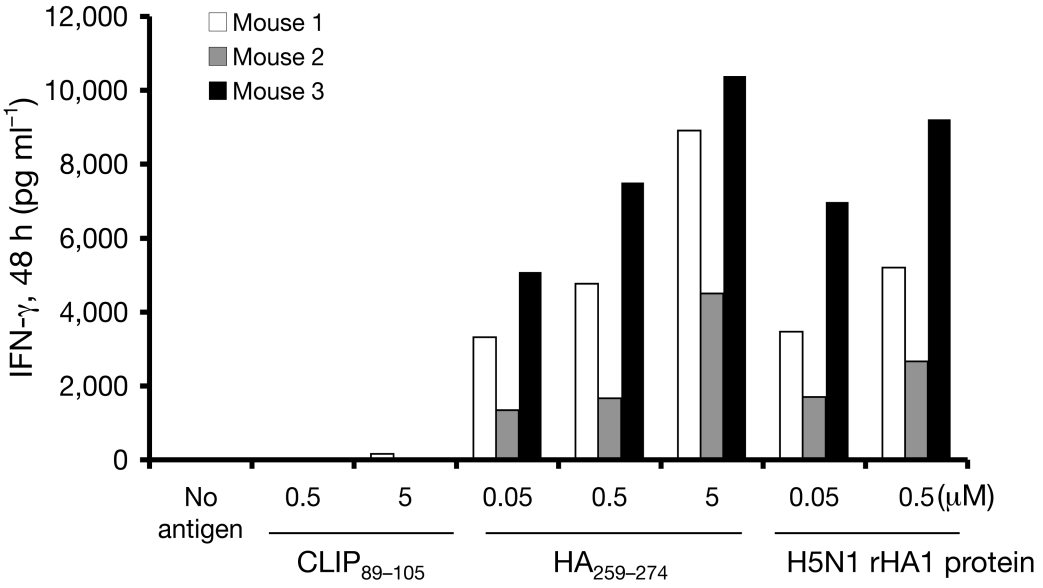
Figure 3. Identification of DR1 restricted epitope of H5N1 rHA1 by the cell-free antigen processing system

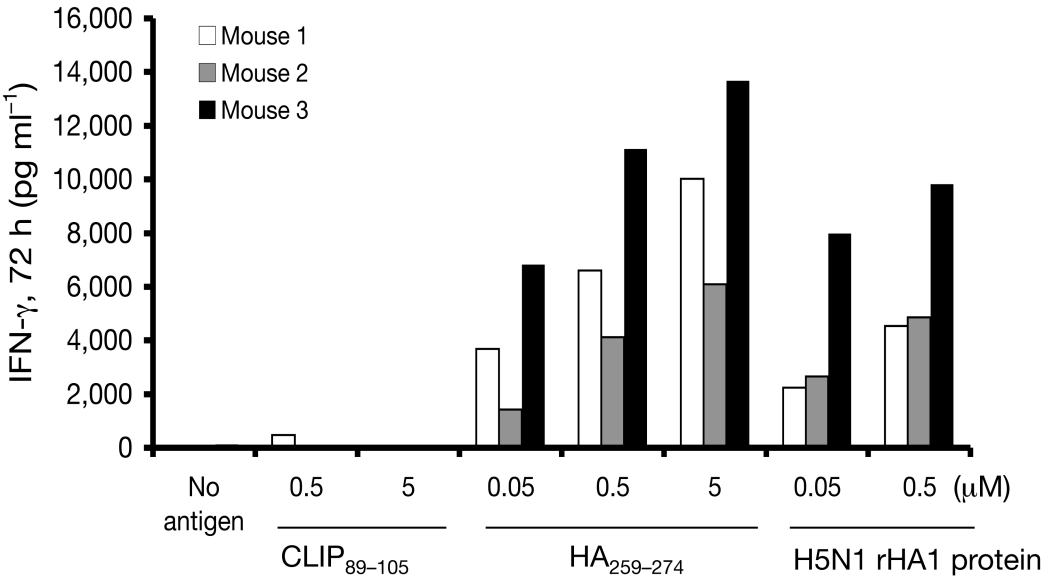
(**a–c**) Mass spectra of H5N1 rHA1 (A/Vietnam/1203/2004 H5N1 strain, Genebank No. AY651334) derived peptides eluted from DR1. DR1 and DM were incubated with (**a**) Native H5N1 rHA1, (**b**) Heat denatured H5N1 rHA1, or (**c**) no protein antigens, followed by addition of CatB, CatH and CatS, and then immunoprecipitation and peptide elution. Mass species that are underlined represent H5N1-derived $HA_{259-274}$, and $HA_{259-278}$ peptides eluted from DR1. Background mass species are labeled in grey. (**d**) MS/MS spectra of $HA_{259-274}$ at m/z 1814.82 Da. These spectra represent one of four repeated experiments.

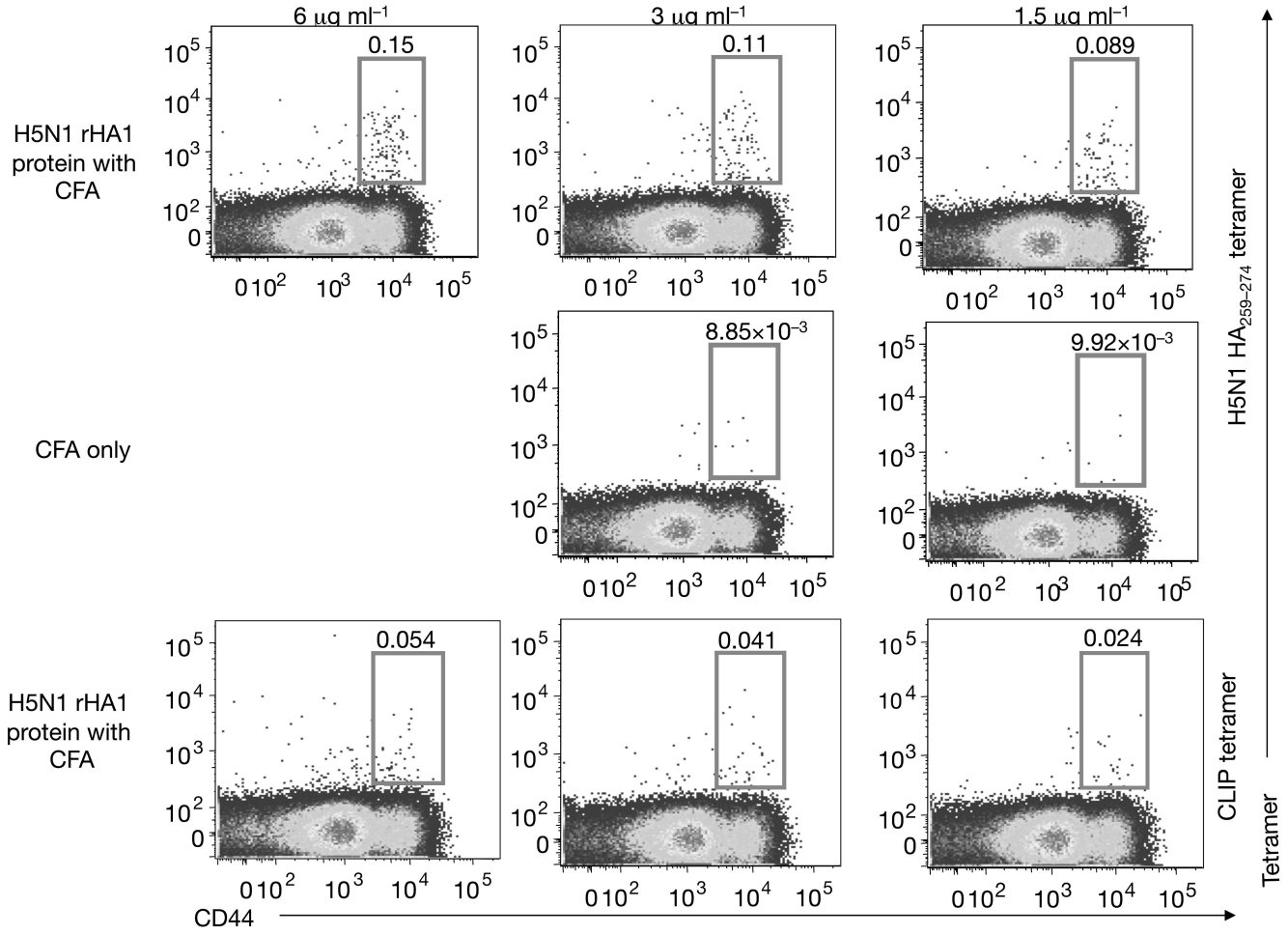












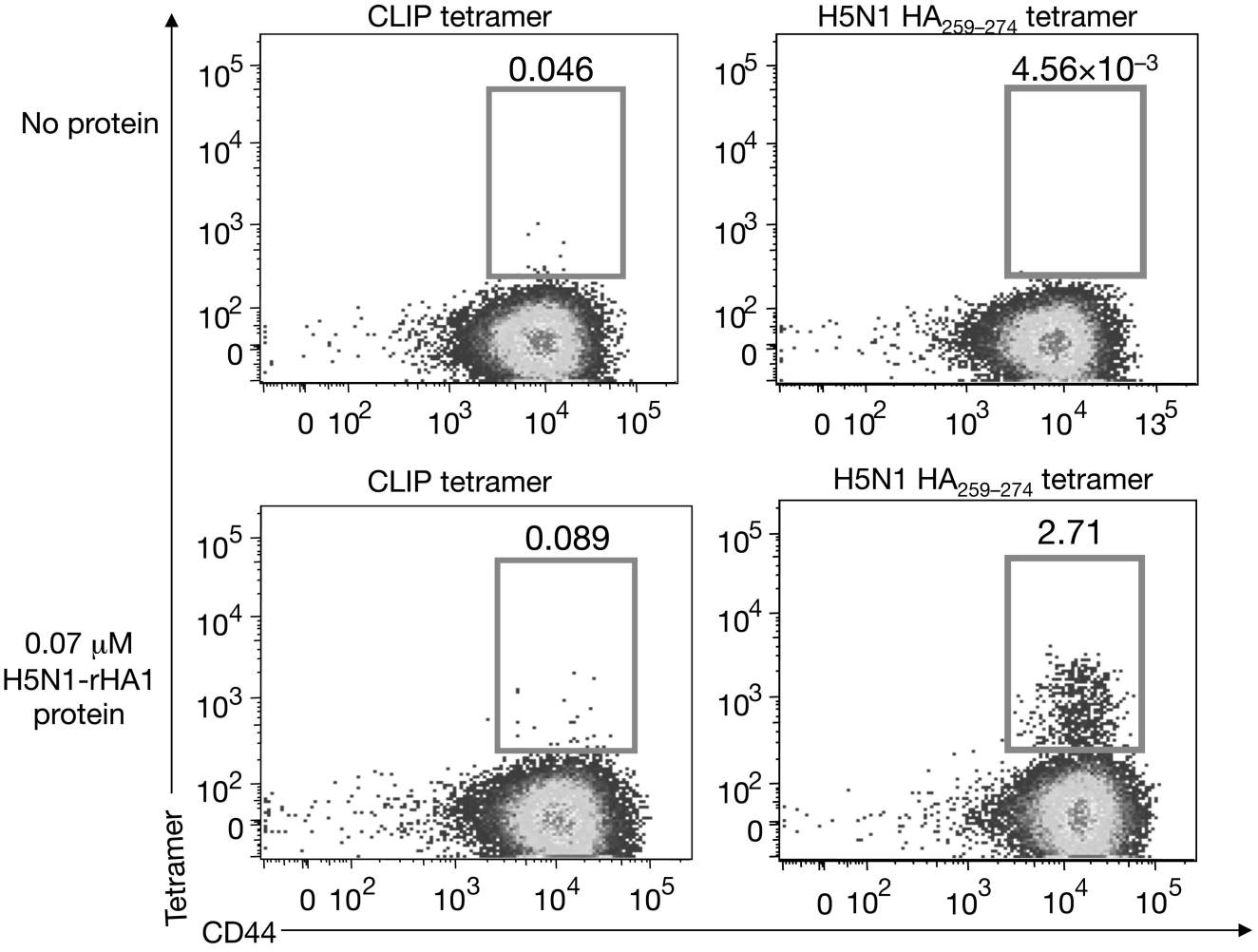
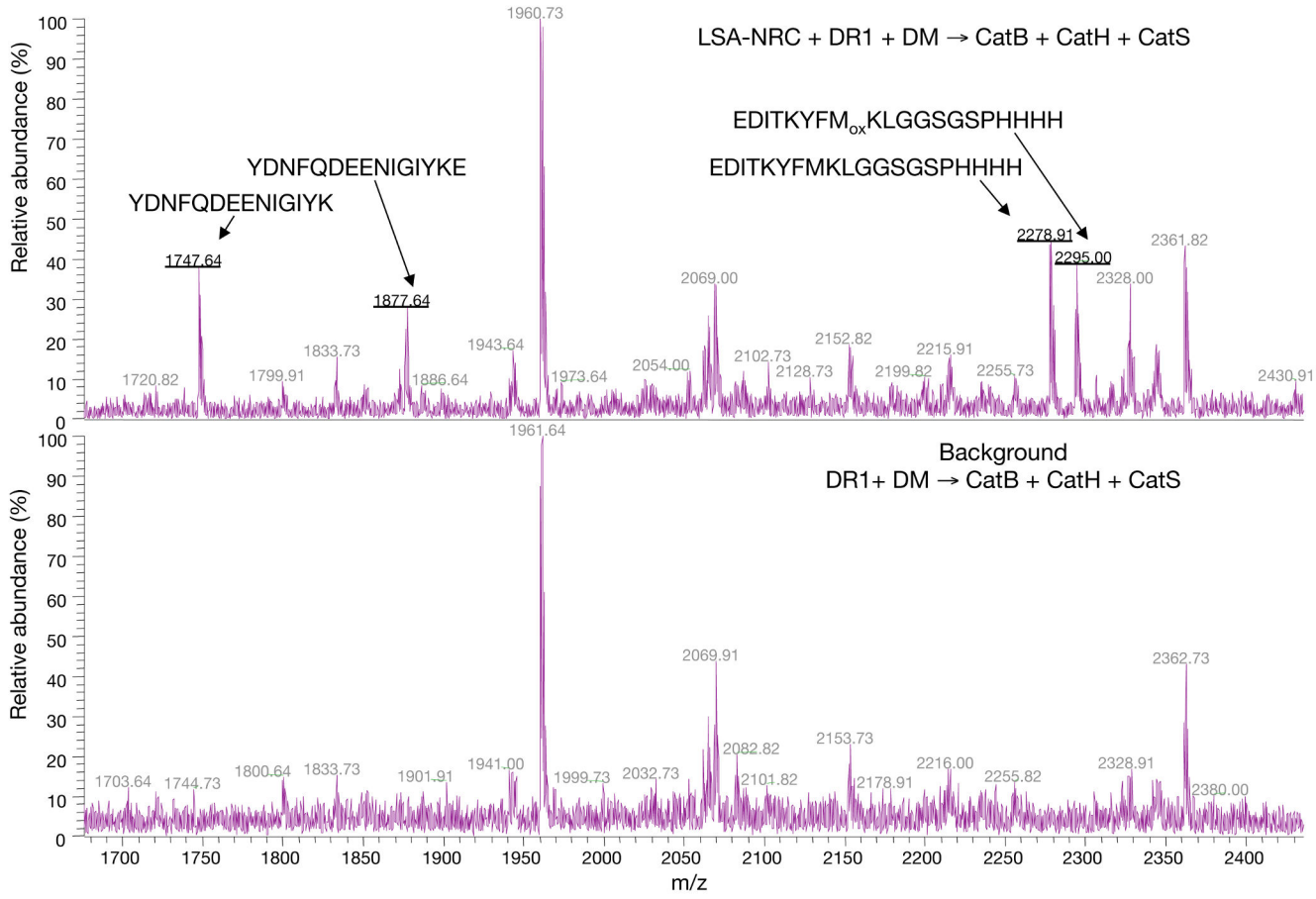
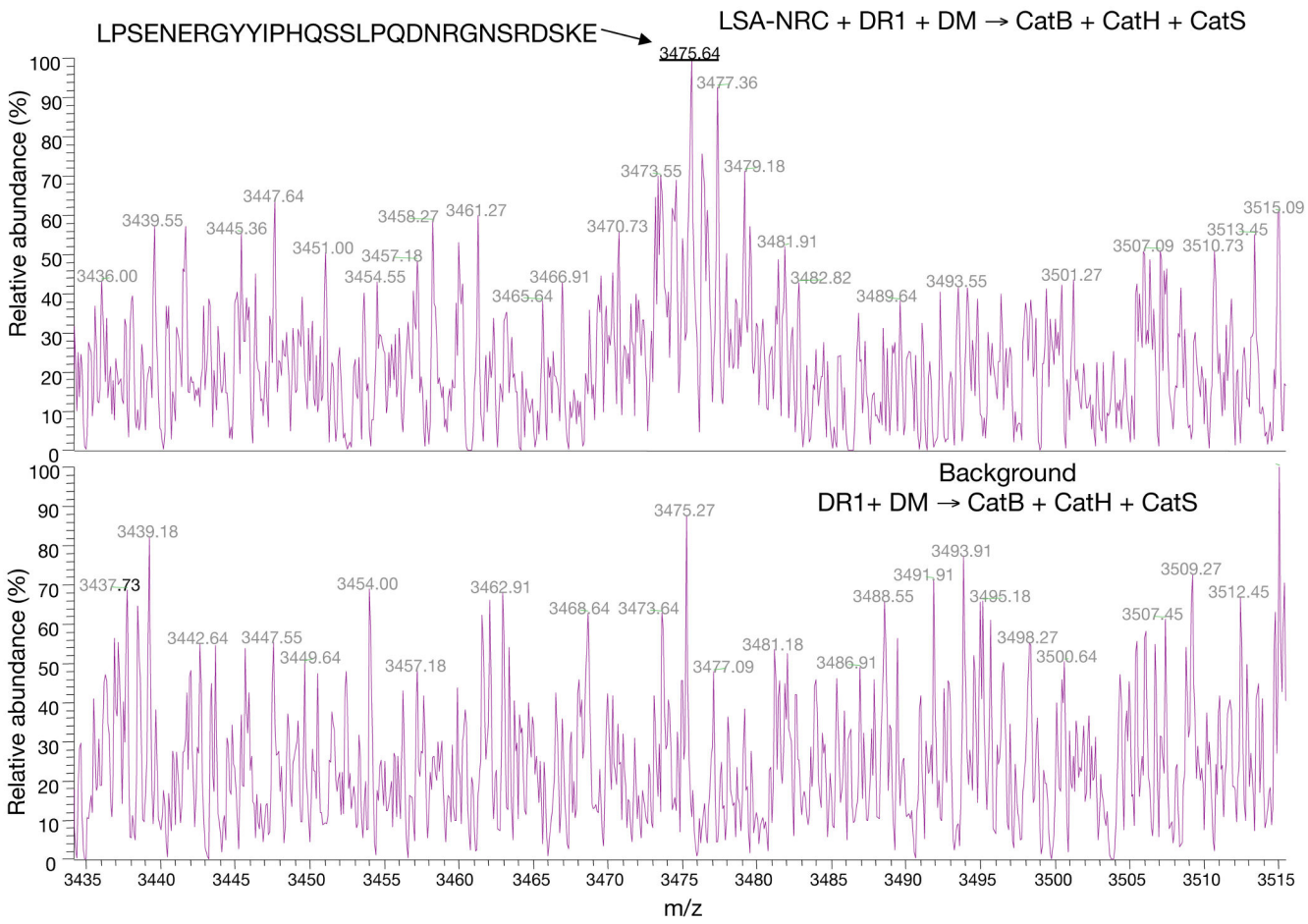


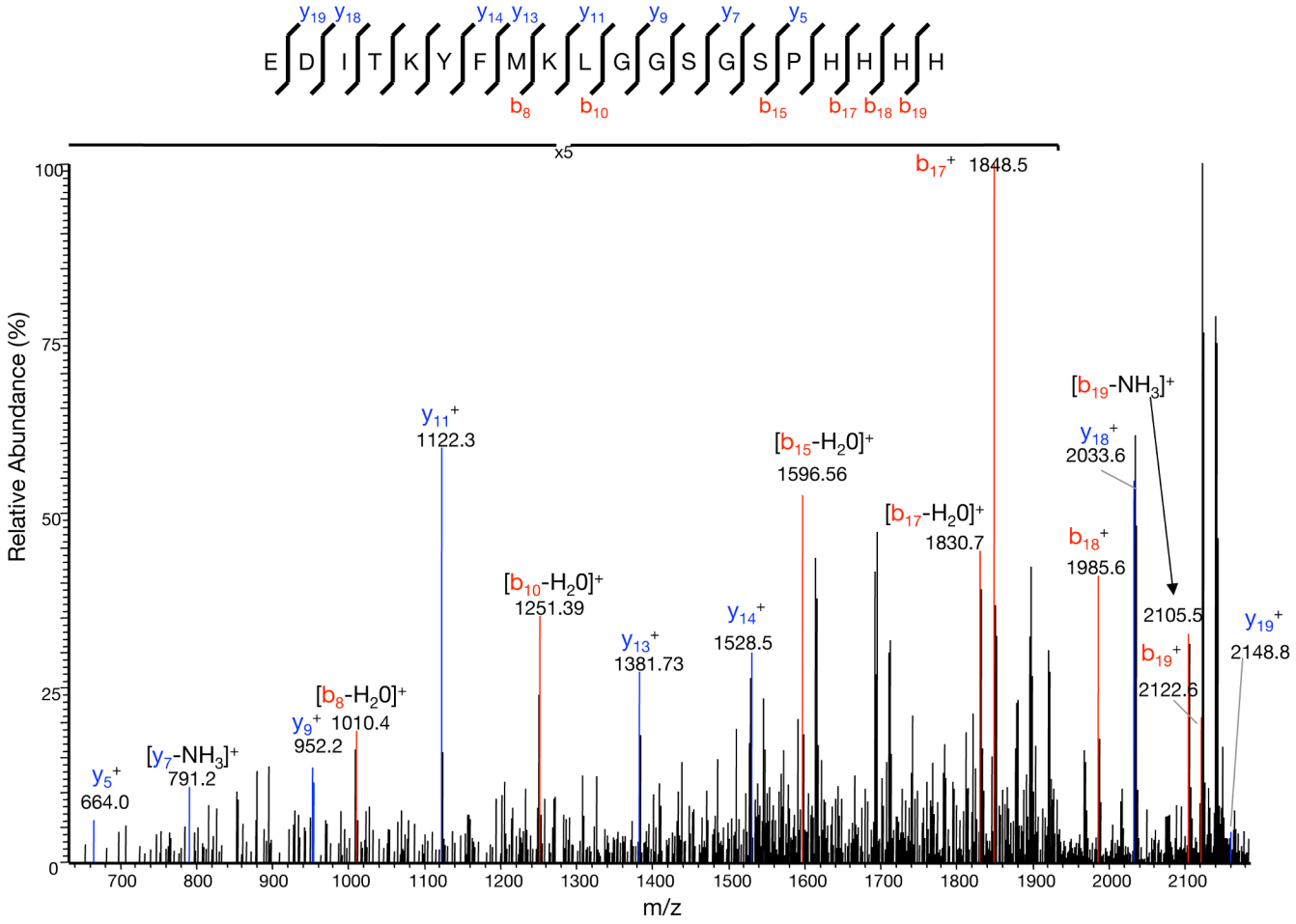
Figure 4. Biological validation of the immunodominant epitope of H5N1 rHA1 identified by the minimalist antigen processing system

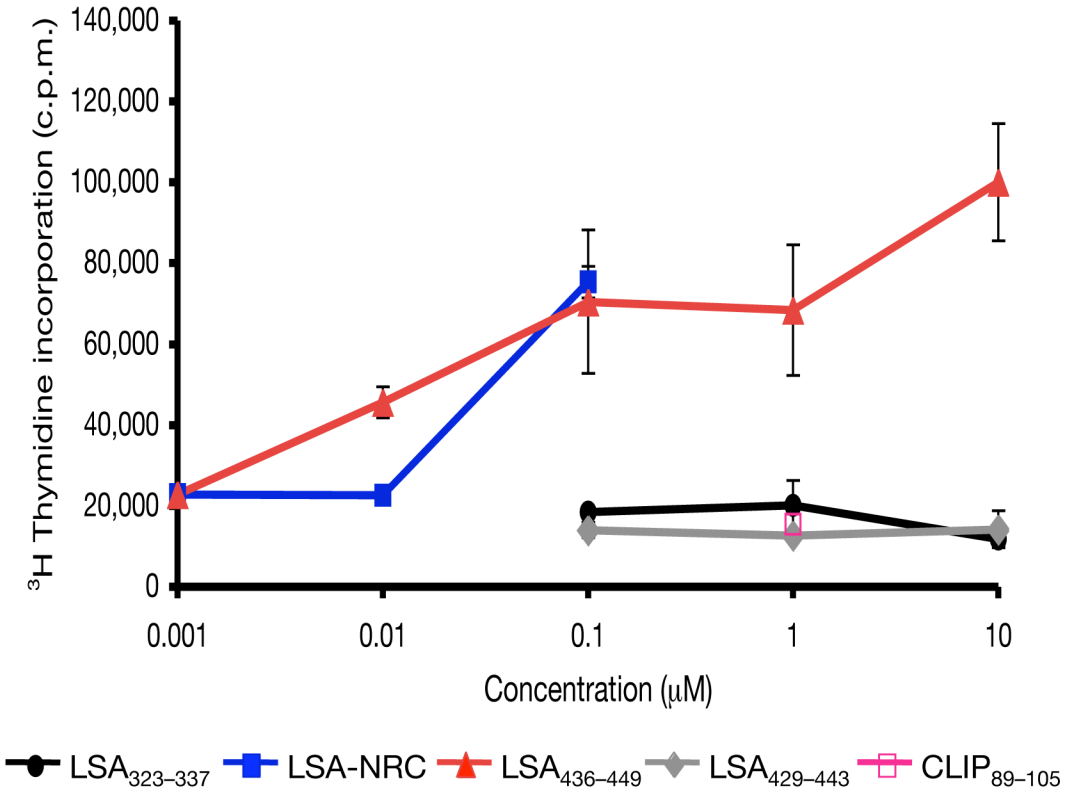
(**a**) T cell proliferation of DR1 Tg mice immunized with native H5N1 rHA1 protein in CFA stimulated with $HA_{259–274}$ (SNGNFIAPEYAYKIVK), $CLIP_{89-105}$, or H5N1 rHA1 protein *in vitro*. Cellular proliferation was measured by [^{3}H] thymidine incorporation (representative of one mouse out of three individual mice tested). (**b**–**c**) IL-2 and IFN-γ ELISA performed from supernatant collected from *in vitro* culture of another three individual mice immunized as in **a**. Cell culture supernatants were removed after 24 h, 48 h, and 72 h culture. (**d**) Draining lymph node cells (pooled from four mice) freshly isolated from H5N1 rHA1 protein immunized mice on day 8 and directly stained with either $HA_{259-274}$/DR1 or CLIP/DR1 tetramers. Cells from another three mice immunized with PBS/CFA as control were stained with $HA_{259-274}$/DR1 tetramers. Cells were stained for 2 h at 37 °C with varying concentrations of tetramers (6 μg ml^{-1}, 3 μg ml^{-1}, or 1.5 μg ml^{-1}) followed by staining with monoclonal antibodies for CD4-FITC, CD8-eFlore605, CD44-Alexa700, F4/80-APC, and B220-APC for additional 20 min at 4 °C. (**e**) Cells from (**d**) were expanded with 0.07 μM H5N1 rHA1 for additional 7 d *in vitro*. Protein stimulated cells were stained with the tetramers (1.5 μg ml^{-1}) and antibodies as described in (**d**). Numbers represent the percentage

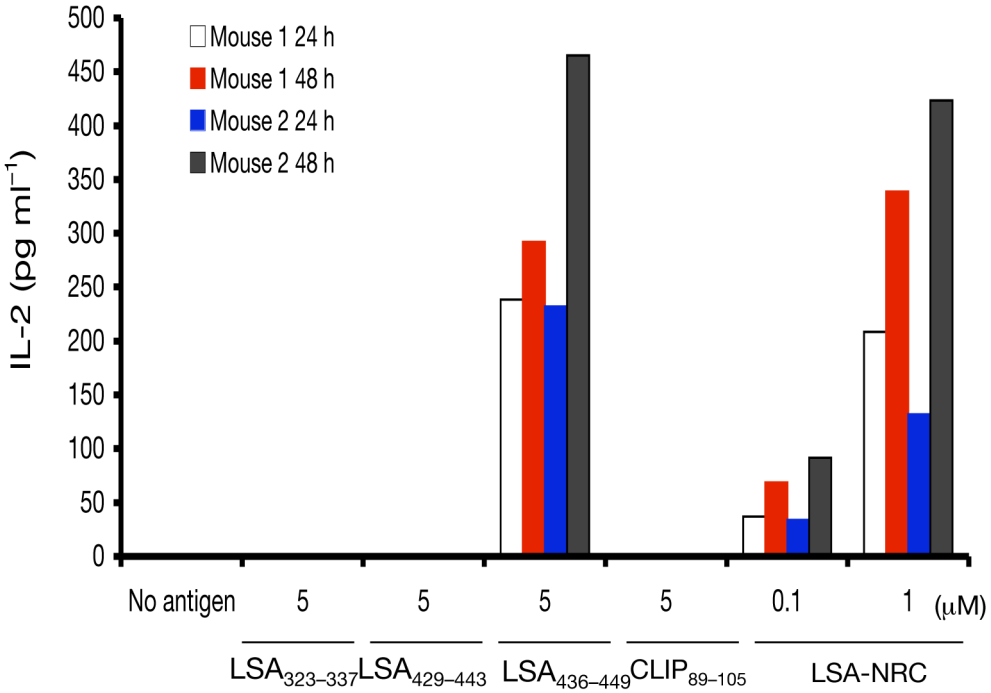
of tetramer-positive events among parent $CD4^+$ ($CD4^+CD44^+CD8^-F4/80^-$ $B220^-7AAD^-$) cells.











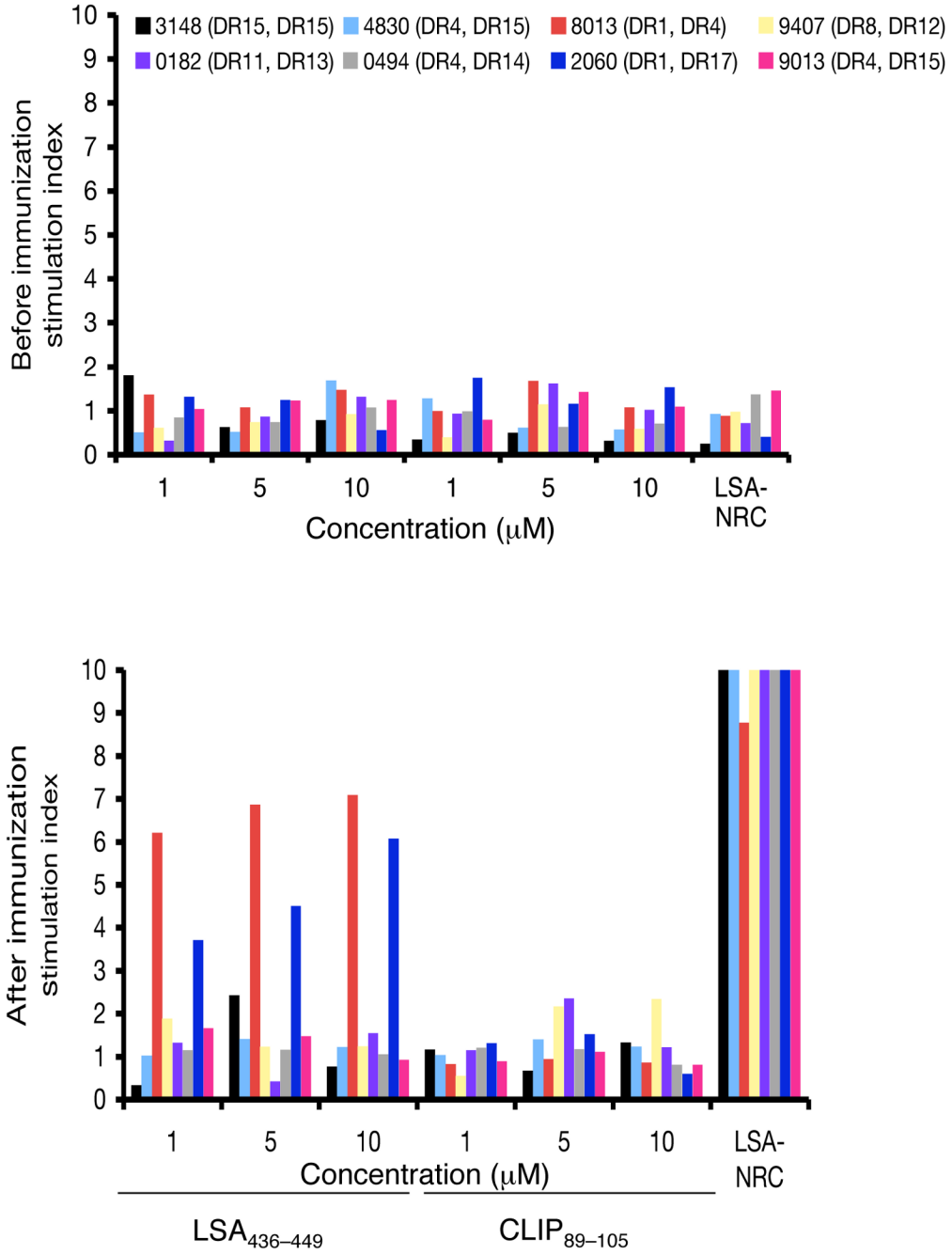


Figure 5. Identification of DR1 restricted epitope of LSA-NRC and biological validation

(**a**, **b**) Mass spectra of LSA-NRC derived peptides eluted from DR1. LSA-NRC (top) or no protein antigens (bottom) were incubated with DR1 and DM followed by addition of CatB, CatH and CatS. LSA-NRC-derived peptides captured by DR1 are underlined. (**a**) Expanded spectrum between m/z 1700-2430 Da and (**b**) Spectrum between m/z 3430-3515 Da. The experiments were repeated more than five times. (**c**) MS/MS spectrum of m/z ~2278 Da that was identified. (**d**, **e**) LSA-NRC immunized cells were incubated with $LSA_{323-337}$ (VQYDNFQDEENIGIY), $LSA_{429-443}$ (VDELSEDITKYFMKL), $LSA_{436-449}$

(ITKYFMKLGGSGSP), $CLIP_{89–105}$, and LSA-NRC *in vitro*. (**d**) T cell proliferation was measured by [^{3}H] thymidine incorportation (one out of three individual mice tested). (**e**) Cell culture supernatant from another two mice under similar experimental set up as in (**d**) were removed 24 h, or 48 h later, and the amount of IL-2 produced was determined by ELISA. (**f**) PBMC from eight volunteers were obtained pre- or post immunization with LSA-NRC Liposome-based vaccine and were stimulated with the intact LSA-NRC protein (0.2 μM) as a positive control, different concentrations of $LSA_{436-449}$, and $CLIP_{89-105}$ for 3 d and were assayed by [^{3}H] thymidine incorporation. Samples of each individual are shown by different colors matching their HLA-DR haplotypes shown on top of the figure.