

A reductionist cell-free major histocompatibility complex class II antigen processing system identifies immunodominant epitopes

Isamu Z Hartman^{1,7,8}, AeRyon Kim^{1,8}, Robert J Cotter², Kimberly Walter³, Sarat K Dalai³, Tatiana Boronina⁴, Wendell Griffith², David E Lanar⁵, Robert Schwenk⁵, Urszula Krzych⁵, Robert N Cole^{4,6} & Scheherazade Sadegh-Nasseri^{1,3}

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: human leukocyte antigen-DR1 (HLA-DR1), HLA-DM and cathepsins. Our reductionist system successfully identified the physiologically selected immunodominant epitopes of two model antigens: hemagglutinin-1 (HA1) from influenza virus (A/Texas/1/77) and type II collagen (CII). When applied for identification of new epitopes from a recombinant liver-stage antigen of malaria falciparum (LSA-NRC) or HA1 from H5N1 influenza virus ('avian flu'), the system selected single epitopes from each protein that were confirmed to be immunodominant by their capacity to activate CD4⁺ T cells from H5N1-immunized HLA-DR1-transgenic mice and LSA-NRC-vaccinated HLA-DR1-positive human volunteers. Thus, we provide a new tool for the identification of physiologically relevant helper T cell epitopes from antigens.

It is well known that the immune system focuses on and responds to very few representative immunodominant epitopes from invading pathogenic insults 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, in case of cancer, fails to respond to the immunodominant epitopes in the context of MHC molecules. Recent advances in understanding of the antigen presentation pathway have shown that the steps of antigen processing and selection crucially 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 reduced to a minimum. 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 use matrix- and anchor-based motifs to screen amino acid sequences of antigenic proteins.

To determine whether 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). We tested this system successfully against two model antigens: recombinant influenza virus HA1 protein (rHA1), containing the well-characterized DR1-restricted immunodominant peptide HA₃₀₆₋₃₁₈ from the A/Texas/1/77 strain³, and CII, which is implicated in collagen-induced arthritis in mice⁴. We evaluated the reductionist system for identifying HLA-DR1-restricted unknown immunodominant epitopes with HA1 from A/Vietnam/1203/2004 (H5N1) and recombinant malaria antigen, LSA-NRC, as the new antigens. We show that the cell-free MHC class II antigen-processing system identified a single epitope from either H5N1 or LSA-NRC, which we verified in HLA-DR1-positive humans or transgenic mice as immunodominant. Thus, our system provides a unique 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,

¹Graduate Program in Immunology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ²Department of Pharmacology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ³Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ⁴Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ⁵Division of Malaria Vaccine Development, Walter Reed Army Institute of Research, Silver Spring, Maryland, USA. ⁶Mass Spectrometry and Proteomics Facility, Institute for Basic Biomedical Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ⁷Present address: Department of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, Texas, USA. ⁸These authors contributed equally to this work. Correspondence should be addressed to S.S.-N. (ssadegh@jhmi.edu).

Received 16 October 2009; accepted 23 July 2010; published online 31 October 2010; doi:10.1038/nm.2248

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 chaperone, and proteolytic enzymes that digest protein antigens and allow binding of some peptide fragments to the groove of MHC class II molecules⁵. To recreate the MHC class II antigen processing compartment⁶, 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 hydrogen bonds between MHC class II and the peptide main chain¹¹. By inducing MHC II to adopt a 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}. We chose the exopeptidases cathepsin B and cathepsin H because they are constitutively expressed in all professional antigen-presenting cells 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 well¹⁴. To mimic the endosomal compartment, we included free L-cysteine¹⁵ to aid in protein unfolding¹⁶. L-cysteine also helps the catalytic activity of thiol-dependent lysosomal enzymes. Because 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) by 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 (Fig. 1b). Because of the susceptibility of DM and empty DR1 to cathepsins, we preincubated antigens with DR1 and DM before the inclusion of the proteases. Once all components of the system were incubated together under endosomal- and lysosomal-like conditions, we isolated peptide-DR1 complexes by immunoprecipitation and analyzed peptides eluted from DR1 on a matrix-assisted laser desorption/ionization (MALDI) mass spectrometer (Supplementary Fig. 1).

rHA1 of influenza with known immunodominant epitope

We first tried using rHA1 from influenza strain A/PR/8/34, to which we genetically attached the well-defined immunodominant epitope from A/Texas/1/77-derived HA₃₀₆₋₃₁₈ (PKYVKQNTLKLAT)³ near its C terminus (Supplementary Table 1 and Supplementary Fig. 2). Comparison of spectra of peptides eluted from DR1 versus the background that did not contain rHA1 showed rHA1-derived peptide species, several of which could be detected in the m/z 1,950–2,530 range (Fig. 1c–f). We determined the sequences of six out of 10 such peptide species through collision-induced dissociation (CID). Among those, four contained the immunodominant HA₃₀₆₋₃₁₈ epitope from influenza strain A/Texas/1/77 (at m/z 2,153.09, 2,217.27, 2,281.27 and 2,524.45) (Fig. 1c and Supplementary Figs. 3 and 4). The sequences of the other two peptides at m/z 2,265.09 and 2,339.09

contained INSSLPYQNIHPVITIGCEPK, derived from influenza strain A/PR/8/34 (Fig. 1c and Supplementary Figs. 5 and 6). For the last four peptides whose sequence identification through CID posed a challenge, we conducted a search to match their masses with rHA1-derived sequences with Findpept (<http://www.expasy.org/tools/findpept.html>) while taking certain artifactual modifications, such as peptide oxidation or alkylation, into consideration^{17,18}. We found that the peaks at m/z 1,955.09, 2,118.00 and 2,503.18 each possibly contained peptides derived from influenza strain A/PR/8/34 (Supplementary Table 2). To rule out the possibility of preexisting peptide fragments containing HA₃₀₆₋₃₁₈ contaminating the mixture during the preparation of rHA1 proteins, we examined the repertoire of peptides bound to DR1 in the presence or absence of the added cathepsins. We detected none of the rHA1-derived peptides without processing of protein antigens by the cathepsins (Supplementary Fig. 7). We prepared samples that included cathepsin S in addition to cathepsins 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 at m/z 1,955.09 and HA₂₉₈₋₃₁₇ (A/PR/8/34) at m/z 2,339.09 (Fig. 1c,e).

We determined the immunodominance of the identified peptides in DR1-transgenic mice immunized with the rHA1 in complete Freund's adjuvant (CFA). These mice⁴ express a fusion product of the peptide-binding groove of DR1 and the membrane-proximal domain of mouse MHC class II I-E. These mice express DR1 as their only MHC II¹⁹, as they were backcrossed to MHC class II deletion mutants for 13 generations to eliminate endogenous mouse MHC class II I-A^f. We isolated draining lymph and used the cells in a recall T cell proliferation assay of increasing doses of the identified peptides, rHA1 or CLIP₈₉₋₁₀₅ as a negative control. We observed strong dose-dependent responses to HA₃₀₆₋₃₁₈ (A/Texas/1/77) and to the rHA1 but no response to CLIP₈₉₋₁₀₅ (Fig. 1g). There was a significantly lower response (two out of five mice) (Fig. 1g) or no response (three mice) to HA₂₉₈₋₃₁₇ (A/PR/8/34) (Fig. 1h). The sum of the proliferative responses to HA₃₀₆₋₃₁₈ and HA₂₉₈₋₃₁₇ peptides approached the magnitude of the response to the whole HA protein (Fig. 1g). Thus, our reductionist system was able to identify the immunodominant epitope.

CII with known immunodominant epitope

We next proceeded with another well-defined immunodominant epitope, CII. CII, a major component of cartilage, is the main suspected autoantigen in rheumatoid arthritis in DR1⁺ individuals. Through studies conducted on DR1-transgenic mice, CII₂₈₂₋₂₈₉ (FKGEQGPK) has been identified as its DR1-restricted immunodominant core epitope⁴. To recapitulate the physiological conditions in which this antigen is digested, we predigested CII with matrix metalloproteinase-9 (MMP-9) because it has been shown that CII undergoes extracellular processing first, with the resulting fragments being further processed after uptake by professional antigen-presenting cells²⁰. We included MMP-9-digested CII, a test antigen, and we isolated and analyzed DR1-bound peptides. The majority of CII-derived peptide species appeared in the m/z 3,000–3,500 range (Fig. 2a–c). We sequenced the most prominent peak of this cluster by tandem mass spectrometry and determined it to be residues 273–305 of CII (QTGEPGIAGFKGEQGPKGEPGPAGVQGAPGPAG, underlining indicates the core CII₂₈₂₋₂₈₉ epitope), with four hydroxylated residues ((CII₂₇₃₋₃₀₅)₄OH; Supplementary Fig. 8 and Supplementary Table 3). The other peptides in

this cluster contained the same core epitope, and they were consistent in mass with the post-translational modification variants of this peptide that would be expected to be present (Fig. 2c)^{20,21}. A baculovirus-encoded conotoxin-like peptide (CL₁₃₋₂₃, FAALNAQHVLA) (Supplementary Fig. 9) was present at *m/z* ~1,153 Da. This peptide probably bound a portion of the DR1 molecules when they were initially synthesized within baculovirus-transduced insect cells (Fig. 2a,b)²². As before, we immunized DR1-transgenic mice with CII protein in CFA, and we used lymphoid cells in a recall proliferation assay with CII₂₈₀₋₂₉₄, the

CII protein or CLIP₈₉₋₁₀₅. We observed strong dose-dependent responses to CII₂₈₀₋₂₉₄ and the CII protein, whereas we saw no response to CLIP₈₉₋₁₀₅ (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. The experimental set up was identical to that with the chimeric rHA1 above. We used the native or heat-denatured HA1 from the H5N1 influenza strain A/Vietnam/1203/2004 (H5N1-rHA1) as antigen.

mpg © 2010 Nature America, Inc. All rights reserved.

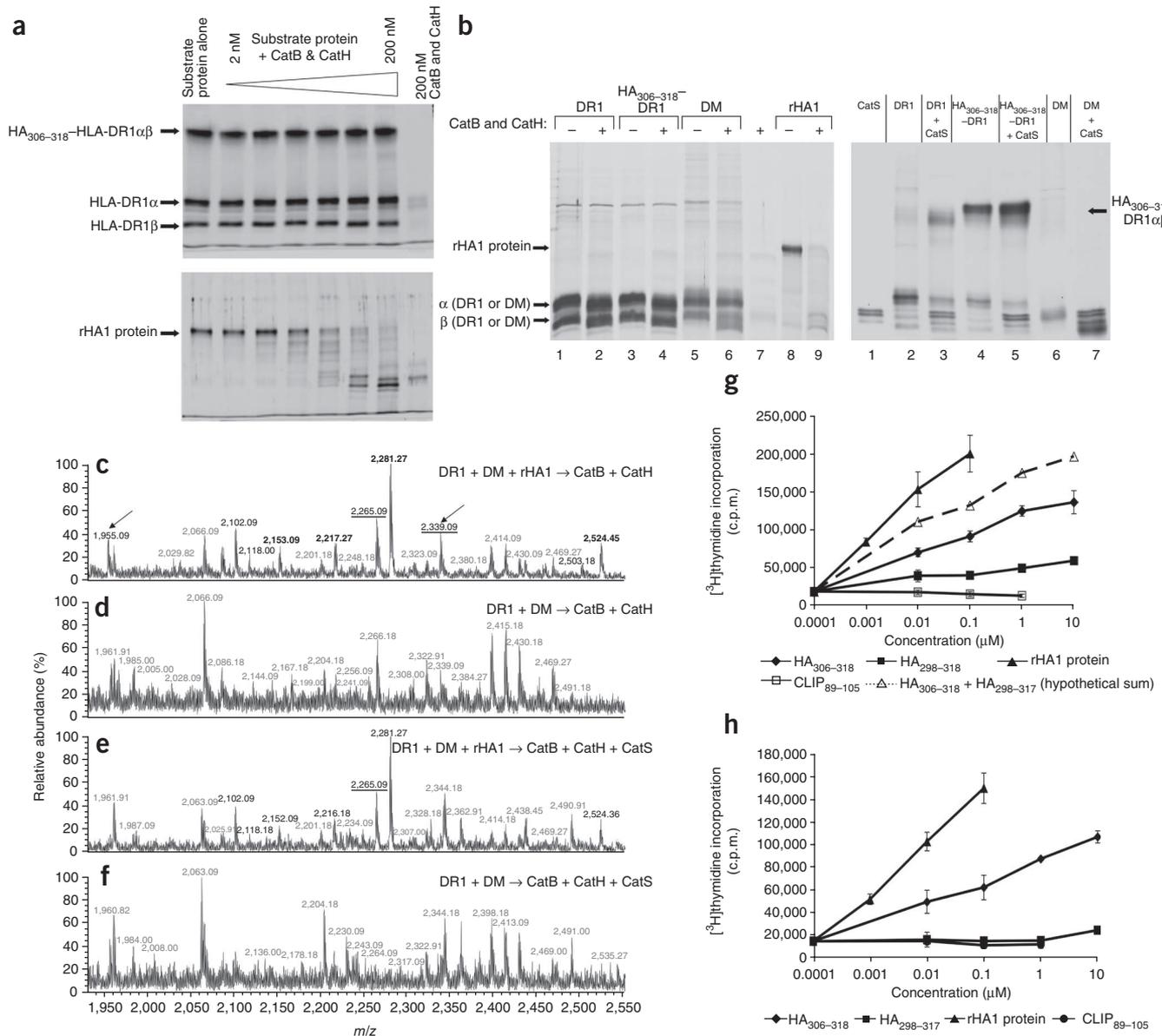


Figure 1 Sensitivity to cathepsins and identification of rHA1-derived peptides by mass spectrometry. (a) Sensitivity to cathepsin B (CatB) and cathepsin H (CatH). HA₃₀₆₋₃₁₈-DR1 (top) and rHA1 (bottom) were treated with CatB and CatH at various concentrations of each enzyme. Samples were resolved by gentle SDS-PAGE where samples were not boiled to preserve peptide-DR1αβ complex³⁷. (b) Sensitivity of DM and DR1 to cathepsins. Left, conventional SDS-PAGE of empty DR1 (lanes 1 and 2), pre-formed HA₃₀₆₋₃₁₈-DR1 (lanes 3 and 4) and DM (lanes 5 and 6) incubated in the presence or absence of 200 nM CatB and CatH. rHA1 (lanes 8 and 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), HA₃₀₆₋₃₁₈-DR1 (lane 5) and DM (lane 7) treated with 100 nM CatS. (c-f) Mass spectra of rHA1-derived peptides eluted from DR1 for *m/z* 1,950–2,550 Da. d is background for c, and f is background for e. Mass species in bold represent rHA1 fragments eluted from DR1 containing HA₃₀₆₋₃₁₈ epitope (Supplementary Figs. 3 and 4), underlining represents other rHA1-derived peptides, and in gray are the background peaks. Experiments were repeated more than three times. Pointed arrows indicate peaks that disappeared in e. (g,h) *In vitro* proliferation of rHA1-immunized cells, as measured by [3H]thymidine incorporation (two individual mice out of five tested).

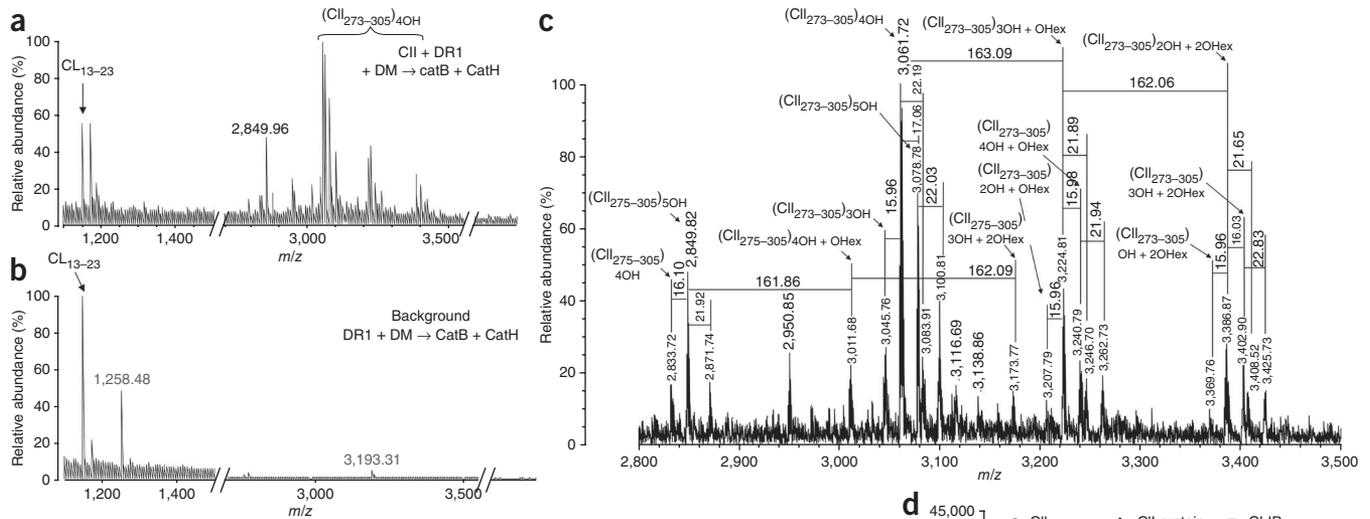


Figure 2 Identification of type II collagen-derived peptides eluted from DR1. (a,b) Mass spectra of CII-derived peptides eluted from DR1. (a) DR1 was incubated with the following components: MMP-9-fragmented bovine CII²⁰, DM, CatB and CatH. (b) The negative control reactions carried out without including MMP-9-fragmented CII. (CII₂₇₃₋₃₀₅)_{4OH} represents the sequence of collagen peptide containing four hydroxylations of proline and lysine (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* 2,800 and 3,500 Da. Peptide modifications: hexose, 162 Da; hydroxylation, 16 Da; Na⁺ adduction, 22Da. (d) Proliferation of cells immunized with native CII protein in CFA incubated with increasing doses of CII₂₈₀₋₂₉₄ (AGFKGEQGPKGEPGP), CLIP₈₉₋₁₀₅ and heat-denatured CII protein *in vitro*, as measured by [³H]thymidine incorporation (one out of three representative individual mice tested).

We detected two unique H5N1-rHA1-derived peptide species at *m/z* 1,814.82 and 2,201.00 compared to the background spectrum, regardless of whether the protein was native or heat denatured (Fig. 3a-c). Microsequencing identified the sequences as HA₂₅₉₋₂₇₄ (SNGNFIAP^uEYAYKIVK) (Fig. 3d) and HA₂₅₉₋₂₇₈ (SNGNFIAP^uEYAYKIVKKGDS) (Supplementary Fig. 10) at *m/z*

1,814.82 and *m/z* 2,201.00, respectively, both sharing the core 16-residue sequence (underlined). Thus, the system selected a single core epitope from intact H5N1-rHA1.

We evaluated the immunodominance of the identified peptide by its ability to induce T cell responses. We immunized DR1-transgenic mice with the H5N1-rHA1 in CFA. We used the lymphoid cells in recall

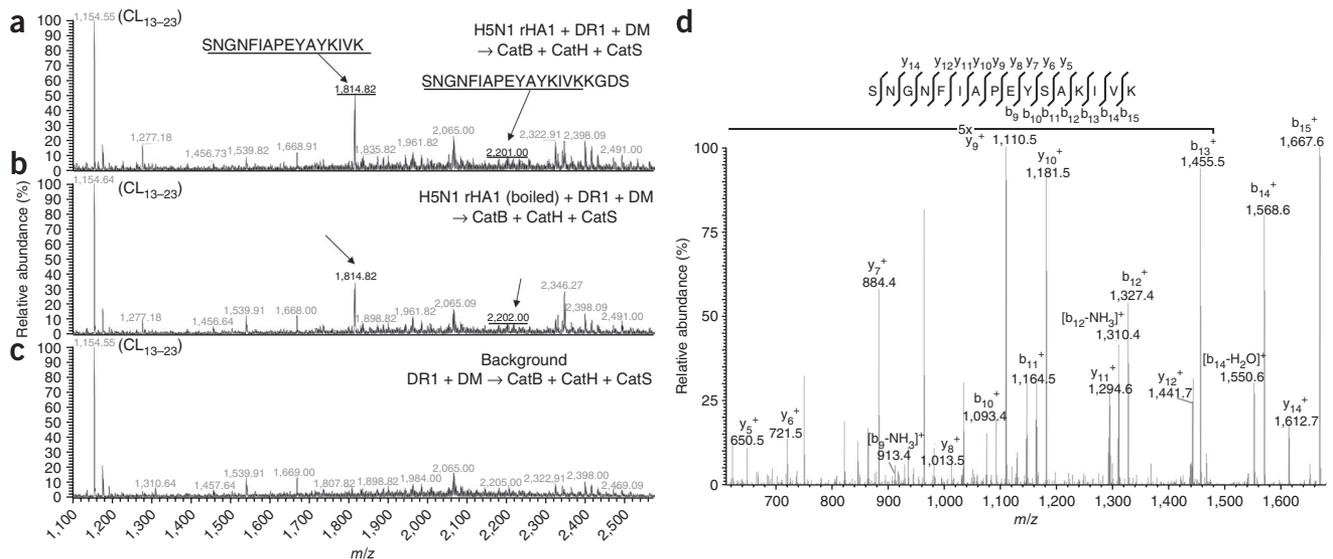


Figure 3 Identification of DR1 restricted epitope of H5N1 rHA1 by the cell-free antigen processing system. (a-c) Mass spectra of H5N1 rHA1 (A/Vietnam/1203/2004 H5N1 strain, Genbank AY651334)-derived peptides eluted from DR1. DR1 and DM were incubated with native H5N1 rHA1 (a), heat-denatured H5N1 rHA1 (b) or no protein antigens (c), followed by addition of CatB, CatH and CatS, and then immunoprecipitation and peptide elution. Mass species that are underlined represent H5N1-derived HA₂₅₉₋₂₇₄ and HA₂₅₉₋₂₇₈ peptides eluted from DR1. Background mass species are labeled in gray. (d) Tandem mass spectra of HA₂₅₉₋₂₇₄ at *m/z* 1,814.82 Da. These spectra represent one of four repeated experiments.

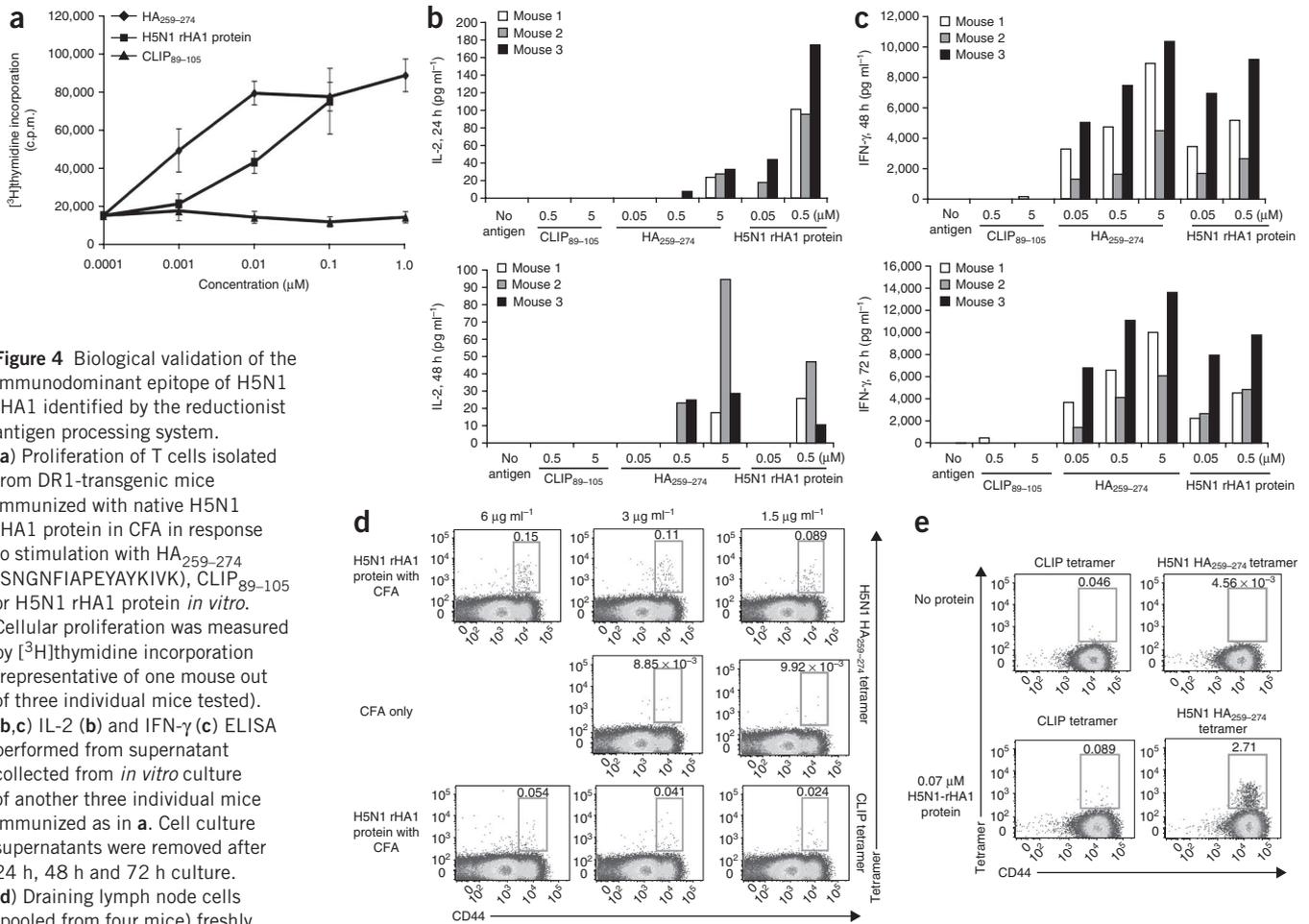


Figure 4 Biological validation of the immunodominant epitope of H5N1 rHA1 identified by the reductionist antigen processing system. **(a)** Proliferation of T cells isolated from DR1-transgenic mice immunized with native H5N1 rHA1 protein in CFA in response to stimulation with HA₂₅₉₋₂₇₄ (SNGNFIAPYAYKIVK), CLIP₈₉₋₁₀₅ or H5N1 rHA1 protein *in vitro*. Cellular proliferation was measured by [³H]thymidine incorporation (representative of one mouse out of three individual mice tested). **(b,c)** IL-2 **(b)** and IFN- γ **(c)** 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 for T cells specific for HA₂₅₉₋₂₇₄-DR1 or CLIP-DR1. Cells from three mock-immunized mice were also stained for the presence of T cell receptors specific for HA₂₅₉₋₂₇₄-DR1 and served as controls. Cells were stained for 2 h at 37 °C with varying concentrations of tetramers as shown followed by staining with monoclonal antibodies for CD4, CD8, CD44, F4/80 and B220. **(e)** Cells from **d** were expanded with 0.07 μM H5N1 rHA1 for an additional 7 d *in vitro*. Protein-stimulated cells were stained with the tetramers (1.5 $\mu\text{g ml}^{-1}$) and the antibodies described in **d**. Numbers represent the percentage of tetramer-positive events among parent CD4⁺ (CD4⁺CD44⁺CD8⁻F4/80⁻B220⁻7AAD⁻) cells.

proliferation and cytokine production assays with increasing doses of HA₂₅₉₋₂₇₄, CLIP₈₉₋₁₀₅ or H5N1-rHA1. We observed strong dose-dependent proliferation responses to HA₂₅₉₋₂₇₄ and H5N1-rHA1 but not CLIP₈₉₋₁₀₅ (Fig. 4a). Proliferation in response to HA₂₅₉₋₂₇₄ was strikingly similar in magnitude to the response elicited by the whole H5N1-rHA1 (Fig. 4a). Interleukin-2 (IL-2) production in response to H5N1-rHA1 was greater than that in response to HA₂₅₉₋₂₇₄ at 24 h, but its production reached similar levels to H5N1-rHA1 at 48 h (Fig. 4b). Overall, we detected a significantly higher amount of interferon- γ (IFN- γ) production in comparison to IL-2, as determined by ELISA (Fig. 4b,c). IFN- γ production in response to stimulation by the H5N1-rHA1 and by the HA₂₅₉₋₂₇₄ was similar in magnitude both at 48 h and at 72 h (Fig. 4c).

We immunized DR1-transgenic mice with either the H5N1-rHA1 in CFA or CFA alone. We stained freshly isolated lymphoid cells with HA₂₅₉₋₂₇₄-DR1 or CLIP-DR1 tetramers and antibodies to CD4 and CD44 (activation marker) (Fig. 4d and Supplementary Fig. 11). Cells positive for HA₂₅₉₋₂₇₄-DR1 tetramer comprised ~0.01% of the CD4⁺CD44⁺ population in the control group versus ~0.1% among cells isolated from mice immunized with H5N1-rHA1-CFA. After a 7-d *in vitro* expansion with H5N1-rHA1, the HA₂₅₉₋₂₇₄-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 7 d resulted in enrichment of HA₂₅₉₋₂₇₄-specific CD4⁺CD44⁺ T cells. These combined results from proliferation assays, cytokine measurements and tetramer staining indicate that the reductionist system identified a major T cell epitope that activates CD4⁺ T cells specific for H5N1-rHA1.

Detection of T cell epitopes from a malaria antigen

Plasmodium falciparum liver-stage antigen 1 (LSA-1) is exclusively expressed in malaria-infected hepatocytes at a preerythrocytic 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 preerythrocyte stage malaria²⁵ (Supplementary Table 4).

We incubated LSA-NRC with DR1 and DM followed by cathepsins B, H and S. We eluted the peptides from DR1 and analyzed them by mass spectrometry. Comparing the spectra of experimental and background groups, we detected five unique peptide species at *m/z* ~1,747, 1,877, 2,278, 2,295 and 3,475 (Fig. 5a,b) and determined their sequences by CID. The sequence of the peptides at *m/z* ~1,747 and 1,877 were LSA₃₂₅₋₃₃₈ (YDNFQDEENIGIYK) (Supplementary Fig. 13) and LSA₃₂₅₋₃₃₉ (YDNFQDEENIGIYKE),

respectively. The peptides at $m/z \sim 2,278$ and 2,295 were post-translation modification variants of a single peptide, LSA_{434–453} (EDITKYFMKLGSGSPHHH) (Fig. 5c), with its methionine residue either unmodified or oxidized (16-Da difference). Although

the intensity of peptide peak at $m/z \sim 3,475$ was weak and indistinguishable from the background, fragmentation data determined its sequence to be LSA_{197–226} (LPSENERGYYPHQSSLPQDNRGNS RDSKE) (Fig. 5b).

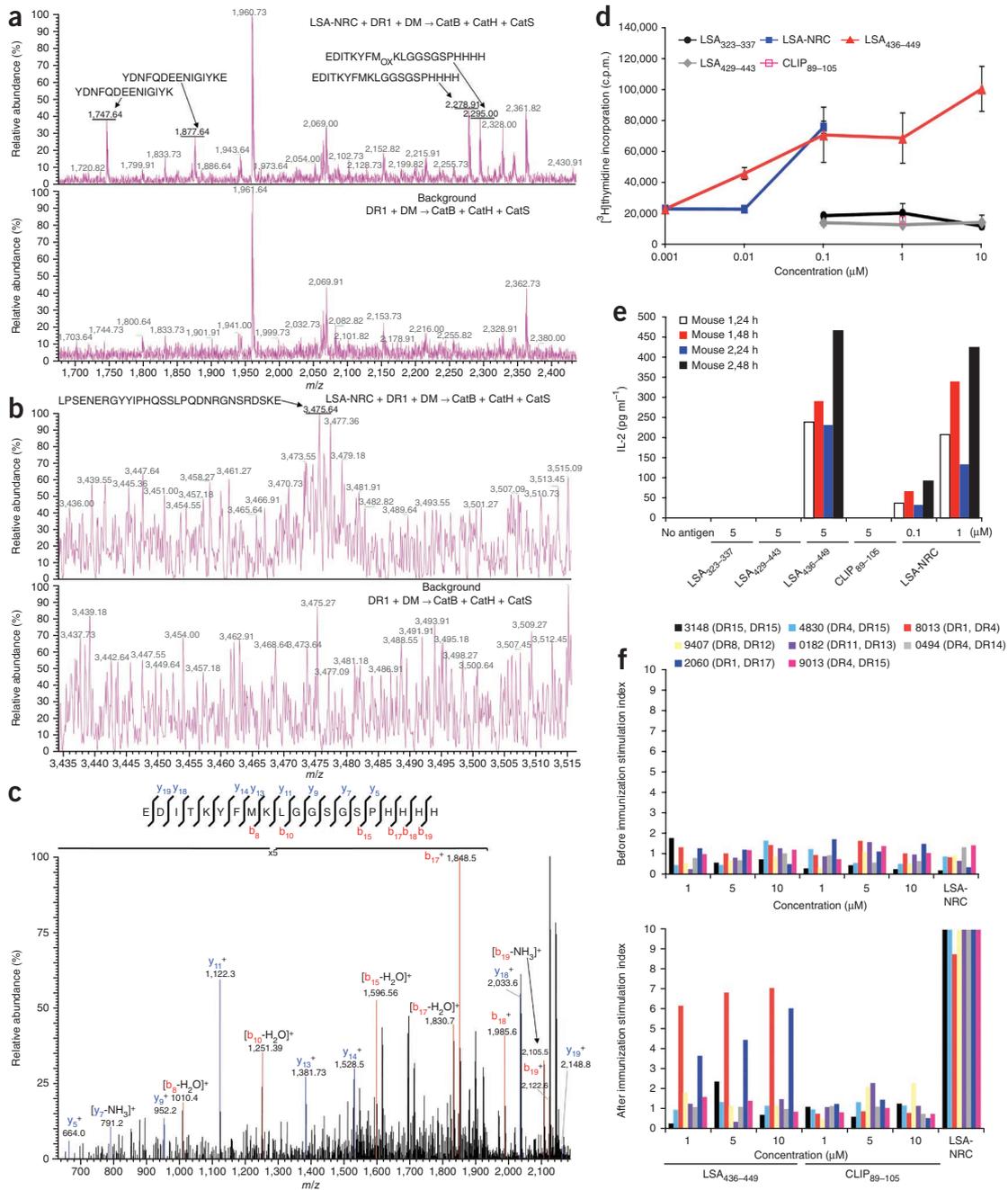


Figure 5 Identification of the DR1-restricted epitope of LSA-NRC and biological validation. **(a,b)** Mass spectra of LSA-NRC-derived peptides eluted from DR1. LSA-NRC (top) or no protein antigens (bottom) were incubated with DR1 and DM followed by addition of CatB, CatH and CatS. LSA-NRC-derived peptides captured by DR1 are underlined. **(a)** Expanded spectrum at 1,700–2,430 Da. **(b)** Spectrum at m/z 3,435–3,515 Da. The experiments were repeated more than five times. **(c)** Tandem mass spectrum of $m/z \sim 2,278$ Da. **(d,e)** LSA-NRC-immunized cells were incubated with LSA_{323–337} (VQYDNFQDEENIGIY), LSA_{429–443} (VDELSEDITKYFMKL), LSA_{436–449} (ITKYFMKLGSGSP), CLIP_{89–105} and LSA-NRC *in vitro*. **(d)** T cell proliferation, as measured by 3H thymidine incorporation (one out of three individual mice tested). **(e)** Cell culture supernatant from another two mice under a 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)** Proliferation results of PBMCs from vaccinated individuals and their HLA-DR types. PBMCs from eight volunteers were obtained before or after immunization with an LSA-NRC liposome-based vaccine and were stimulated with the intact LSA-NRC protein (0.2 μM) as a positive control or different concentrations of LSA_{436–449} and CLIP_{89–105} and were assayed by 3H thymidine incorporation. Samples of each individual are shown by different colors matching their HLA-DR haplotypes shown on top of the figure.

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

To confirm the potential clinical utility of this system, we used donor-derived T cells from human volunteers who, as part of a malaria vaccine trial, had been immunized twice with LSA-NRC in a liposome-based formulation²⁶. We tested samples from eight individuals whose CD4⁺ T cells had responded vigorously to the LSA protein *in vitro*. We obtained peripheral blood mononuclear cells (PBMCs) before immunization and after a second immunization with LSA-NRC. We used these PBMCs for proliferation by stimulation *in vitro* with various concentrations of LSA_{436–449}, CLIP_{89–105} and LSA-NRC. We found that whereas all eight post-immunization samples responded to LSA protein, only two individuals (subject 8013 and subject 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⁺ individuals (DR B1*0101) (Fig. 5f). The six nonresponding samples were from DR1⁻ individuals carrying DR4, DR8, DR11, DR12, DR13, DR14, DR15 or DR17 variants. These results indicate that LSA_{436–449} is a DR1-restricted epitope, consistent with the fact that the reductionist system used to identify this epitope used DR1. Notably, these results demonstrate that the reductionist system is capable of predicting epitopes that human T cells recognize.

DISCUSSION

We have established a reductionist antigen processing system composed of DR1, DM and cathepsins B, H and S that identifies physiologically relevant immunodominant epitopes. We confirmed immunodominance in immunized DR1-transgenic mice and in humans. We confirmed the immunodominance of *de novo* epitopes of H5N1-rHA1 or LSA-1 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 MHC II during antigen processing.

The finding that so few components of MHC II antigen processing can process full-length antigens and yield the physiologically selected immunodominant epitopes from protein antigens is notable. 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). In contrast, a combination of cathepsins B, H and S was sufficient for all four antigens studied here. Cathepsins B and H in the absence of cathepsin S led to the capture of HA_{306–318} and HA_{298–317} from our model influenza antigen. There has been little appreciation for the roles of the exopeptidases, cathepsin B and cathepsin H, in antigen processing. Cathepsin B-deficient mice have been shown not to have any major defects in antigen processing²⁷. Also, little information, if any, is available regarding a role for cathepsin H in antigen processing. As such, it was quite unexpected that the combined use of cathepsins B and H led to the proper processing of protein antigens and the selection of immunodominant epitopes,

whereas the presence or absence of cathepsin S did not change the overall outcome of epitope selection for rHA1 from the A/Texas1/77 strain. Thus, we here define a new key role for cathepsin H in conjunction with cathepsin B.

Our system relies on mass spectrometry for identifying epitopes, a technology that has been previously used for characterizing antibody epitopes²⁸ and 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 crucial in activation of CD4⁺ and CD8⁺ T cells by MHC II³² and MHC I³³, 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 with only HLA-DR1, 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 and cancer. The identified epitopes can be incorporated into MHC class II multimers³⁴ 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 the design of peptide vaccines or for evaluation of proper recombinant vaccine designs before human trial studies to ascertain that artifacts, such as additional sequences not present naturally, introduced to the recombinant vaccines do not interfere with the recognition of the wild-type epitopes.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturemedicine/>.

Note: Supplementary information is available on the Nature Medicine website.

ACKNOWLEDGMENTS

We wish to thank S. Landry, A. Hamad, J. Pomerantz and K. Narayan for reading 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 *Escherichia coli* transformed with an expression vector encoding influenza HA, D. Zaller (Merck) for the original DR1-transgenic mice, L. Rein for testing the human samples and the US National Institutes of Health Tetramer Facility for providing DR1 and CLIP monomers. This work was supported by R01 grants AI063764 and GM053549, a grant from Johns Hopkins Malaria Research Institute to S.S.-N., a US National Science Foundation predoctoral award to I.Z.H. and a Johns Hopkins University Rheumatology T32 Fellowship to A.K.

AUTHOR CONTRIBUTIONS

I.Z.H. and A.K. designed and executed experiments, analyzed data and wrote the paper; R.J.C., W.G., T.B. and R.N.C. provided mass spectrometry data and analyses; K.W. cloned DM; S.K.D. contributed to *in vivo* testing; R.S., D.E.L. and U.K. did LSA-1 experiments in humans; and S.S.-N. designed experiments, supervised the project, obtained funding and wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturemedicine/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

1. Sercarz, E.E. *et al.* Dominance and crypticity of T cell antigenic determinants. *Annu. Rev. Immunol.* **11**, 729–766 (1993).
2. Sercarz, E.E. Immune focusing vs diversification and their connection to immune regulation. *Immunol. Rev.* **164**, 5–10 (1998).
3. Lamb, J.R., Eckels, D.D., Lake, P., Woody, J.N. & Green, N. Human T-cell clones recognize chemically synthesized peptides of influenza haemagglutinin. *Nature* **300**, 66–69 (1982).
4. Rosloniec, E.F. *et al.* An *HLA-DR1* transgene confers susceptibility to collagen-induced arthritis elicited with human type II collagen. *J. Exp. Med.* **185**, 1113–1122 (1997).
5. Jensen, P.E. Recent advances in antigen processing and presentation. *Nat. Immunol.* **8**, 1041–1048 (2007).
6. Amigorena, S., Drake, J.R., Webster, P. & Mellman, I. Transient accumulation of new class II MHC molecules in a novel endocytic compartment in B lymphocytes. *Nature* **369**, 113–120 (1994).
7. Kroshofer, H. *et al.* Editing of the HLA-DR-peptide repertoire by HLA-DM. *EMBO J.* **15**, 6144–6154 (1996).
8. Chou, C.L. & Sadegh-Nasseri, S. HLA-DM recognizes the flexible conformation of major histocompatibility complex class II. *J. Exp. Med.* **192**, 1697–1706 (2000).
9. Riberty, J.M., Newcomb, J.R., Surman, M.J., Barbosa, J.A. & Cresswell, P. HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. *Nature* **360**, 474–477 (1992).
10. Sloan, V.S. *et al.* Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* **375**, 802–806 (1995).
11. Sadegh-Nasseri, S., Chen, M., Narayan, K. & Bouvier, M. The convergent roles of tapasin and HLA-DM in antigen presentation. *Trends Immunol.* **29**, 141–147 (2008).
12. Plüger, E.B. *et al.* Specific role for cathepsin S in the generation of antigenic peptides *in vivo*. *Eur. J. Immunol.* **32**, 467–476 (2002).
13. Riese, R.J. *et al.* Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. *Immunity* **4**, 357–366 (1996).
14. Chapman, H.A. Endosomal proteases in antigen presentation. *Curr. Opin. Immunol.* **18**, 78–84 (2006).
15. Pisoni, R.L., Acker, T.L., Lisowski, K.M., Lemons, R.M. & Thoene, J.G. 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.* **110**, 327–335 (1990).
16. Maric, M. *et al.* Defective antigen processing in GILT-free mice. *Science* **294**, 1361–1365 (2001).
17. Boja, E.S. & Fales, H.M. Overalkylation of a protein digest with iodoacetamide. *Anal. Chem.* **73**, 3576–3582 (2001).
18. Schey, K.L. & Finley, E.L. Identification of peptide oxidation by tandem mass spectrometry. *Acc. Chem. Res.* **33**, 299–306 (2000).
19. Narayan, K. *et al.* Staphylococcal enterotoxin A induces small clusters of HLA-DR1 on B cells. *PLoS ONE* **4**, e6188 (2009).
20. Van den Steen, P.E. *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.* **16**, 379–389 (2002).
21. Van den Steen, P.E. *et al.* Generation of glycosylated remnant epitopes from human collagen type II by gelatinase B. *Biochemistry* **43**, 10809–10816 (2004).
22. Stern, L.J. & Wiley, D.C. The human class II MHC protein HLA-DR1 assembles as empty $\alpha\beta$ heterodimers in the absence of antigenic peptide. *Cell* **68**, 465–477 (1992).
23. Krzych, U. *et al.* T lymphocytes from volunteers immunized with irradiated *Plasmodium falciparum* sporozoites recognize liver and blood stage malaria antigens. *J. Immunol.* **155**, 4072–4077 (1995).
24. Fidock, D.A. *et al.* *Plasmodium falciparum* liver stage antigen-1 is well conserved and contains potent B and T cell determinants. *J. Immunol.* **153**, 190–204 (1994).
25. Hillier, C.J. *et al.* Process development and analysis of liver-stage antigen 1, a preerythrocyte-stage protein-based vaccine for *Plasmodium falciparum*. *Infect. Immun.* **73**, 2109–2115 (2005).
26. Cummings, J.F. *et al.* Recombinant liver stage antigen-1 (LSA-1) formulated with AS01 or AS02 is safe, elicits high titer antibody and induces IFN- γ /IL-2 CD4+ T cells but does not protect against experimental *Plasmodium falciparum* infection. *Vaccine* **28**, 5135–5144 (2010).
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* **95**, 4516–4521 (1998).
28. Parker, C.E., Papac, D.I., Trojak, S.K. & Tomer, K.B. Epitope mapping by mass spectrometry: determination of an epitope on HIV-1 IIIB p26 recognized by a monoclonal antibody. *J. Immunol.* **157**, 198–206 (1996).
29. Chicz, R.M. *et al.* Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* **358**, 764–768 (1992).
30. Hunt, D.F. *et al.* Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **255**, 1261–1263 (1992).
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.* **7**, 2703–2711 (2008).
32. Ireland, J., Herzog, J. & Unanue, E.R. Cutting edge: unique T cells that recognize citrullinated peptides are a feature of protein immunization. *J. Immunol.* **177**, 1421–1425 (2006).
33. Beltrami, A. *et al.* Citrullination-dependent differential presentation of a self-peptide by HLA-B27 subtypes. *J. Biol. Chem.* **283**, 27189–27199 (2008).
34. Gebe, J.A. & Kwok, W.W. Tracking antigen specific CD4+ T-cells with soluble MHC molecules. *Methods Mol. Med.* **136**, 39–50 (2007).

ONLINE METHODS

Peptides. HA_{306–318} peptide (PKYVKQNTLKLAT), human short CLIP_{89–105} peptide (KMRMATPLLMQALPM) and CII_{280–294} (AGFKGEQGPKEGPGP) were synthesized by Global Peptide. HA_{298–317} peptide (INSSLPYQNIHPV TIGECPKY), HA_{259–274} peptide (SNGNFIAPEYAYKIVK) and LSA_{436–449} (ITKYFMKLGSGSP) were synthesized by Elim Biopharmaceuticals at >85% purity as analyzed by reverse-phase HPLC, and their identities were confirmed by mass spectrometry. Overlapping peptides LSA_{323–337} (VQYDNFQDEENIGIY) and LSA_{429–443} (VDELSIEDITKYFMKL) 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}. rHA1 was produced in *E. coli* transformed with an expression vector for a dual 6× histidine-tagged influenza HA. The protein contains residues 12–340 of the HA of influenza strain A/PR/8/34 HA protein with a MRGSHHHHHHTDPSSRSA tag on the amino terminus and a ACPKYVKQNTLKLATGMRKLHHHHHHN tag on the carboxy terminus (the underlined residues comprise the HA_{306–318} epitope from influenza strain A/Texas/1/77). After affinity purification from bacterial lysate with Ni-NTA-charged agarose resin (Ni-NTA Superflow, Qiagen), the protein was refolded by stepwise dialysis into PBS plus 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 previously²⁵.

Experimental setup 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 (L243, American Type Culture Collection). Bound peptides were eluted with trifluoroacetic acid³⁶, filtered through a 10-kDa molecular weight cut-off Microcon filter (Millipore) and lyophilized.

Mass spectrometry. Lyophilized samples were resuspended in 5–10 µl of 50% ethanol, 50% water and 0.2% TFA. We spotted 0.5 µl of resuspended sample onto a MALDI target plate, dried it and covered it with 0.5 µl matrix (40–50 mg ml^{–1} 2,5-dihydroxybenzoic acid or 2.5 mg ml^{–1} α-cyano-4-hydroxycinnamic acid in 50% ethanol, 50% water and 0.15%TFA). Samples were analyzed on the vMALDI-LTQ mass spectrometer (ThermoFisher), using Tune Plus 2.2 Xcalibur 2.0 SR2 vMALDI LTQ 2. Full mass spectrometry scans (*m/z* 1, 100–4,000 Da) were acquired to select peptides of interest for identification by CID fragmentation (MS²).

Full mass spectrometry spectra were manually inspected to find the peptides of interest. MSⁿ 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 MS¹, mass tolerance of 1Da for MS² and MS³, and with allowances for variable carboamidomethylated of cysteine and oxidation of methionine. For CII, 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-transgenic mice¹⁹ were immunized with 50 µg proteins in CFA in the base of the tail. All experiments were performed in accordance with protocols approved by the Animal Care and Use Committee of the Johns Hopkins University School of Medicine. After 8–10 d, the draining lymph nodes were collected, and the cells were incubated with a range of peptide and protein concentrations for 3 d before adding [³H]thymidine (Amersham). The cells were collected and counted after a further incubation of 18–20 h, and the incorporated radioactivity was measured by a 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 ELISA (R&D Systems). 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 peripheral blood mononuclear cell proliferation assay. Details of the screening and enrollment of the volunteers, preparation of LSA-NRC, the adjuvant used for immunization and other procedures have been previously described²⁶. The study was conducted according to Good Clinical Practices under a protocol approved by the Human Use Review Committee of the Walter Reed Army Institute of Research, The Western Institutional Review Board and the US Army Medical Research and Materiel Command's Human Subjects Research Review Board under an Investigational New Drug application filed at the US Food and Drug Administration. This study was registered at www.ClinicalTrials.gov under reference identifiers NCT00312702 and NCT00312663. PBMCs both before immunization and after immunization with LSA-NRC²⁶ were collected and tested for proliferation. PBMCs from eight different donors were incubated with several concentrations of peptides or protein for 3 d before adding [³H]thymidine. The cells were collected and counted after a further incubation of 18–20 h, and then the incorporated radioactivity was measured. The stimulation index was calculated by dividing antigen-induced proliferation by background proliferation (medium alone). HLA-DRB1 typing of donors was done in the immunogenetics laboratory at Johns Hopkins University by PCR.

35. Narayan, K. *et al.* HLA-DM targets the hydrogen bond between the histidine at position β81 and peptide to dissociate HLA-DR-peptide complexes. *Nat. Immunol.* **8**, 92–100 (2007).
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.* **168**, 1235–1243 (2002).
37. Natarajan, S.K., Stern, L.J. & Sadegh-Nasseri, S. Sodium dodecyl sulfate stability of HLA-DR1 complexes correlates with burial of hydrophobic residues in pocket 1. *J. Immunol.* **162**, 3463–3470 (1999).