Chapter 15

Construction and Screening of an Antigen-Derived Peptide Library Displayed on Yeast Cell Surface for CD4+ T Cell Epitope Identification

Fei Wen and Huimin Zhao

Abstract

Identification of T cell epitopes is a critical, but often difficult step in studying T cell function and developing peptide-based vaccines and immunotherapies. Unlike antibodies that recognize free soluble antigens, T cell receptor (TCR) recognizes its epitope bound to major histocompatibility complex (MHC) expressed on antigen presenting cells (APCs). In addition, the examination of T cell epitope activity requires the use of professional APCs, which are difficult to isolate, expand, and maintain. To address these issues, we have developed a facile, accurate, and high-throughput method for T cell epitope mapping by screening antigen-derived peptide libraries in complex with MHC protein displayed on yeast cell surface. Here, we use hemagglutinin and influenza A virus X31/A/Aichi/68 as examples to describe the key steps in identification of CD4+ T cell epitopes from a single antigenic protein and the entire genome of a pathogen, respectively. Methods for single-chain peptide-MHC complex vector design, yeast surface display, peptide library generation in *Escherichia coli*, and functional screening in *Saccharomyces cerevisiae* are discussed.

Key words CD4+ T cell epitope mapping, Peptide library, Major histocompatibility complex, Single-chain peptide-MHC complex, HLA-DR1, Yeast display, Influenza A virus, Flow cytometry, High throughput screening

1 Introduction

T cells respond to selected peptides (termed T cell epitopes) in complex with major histocompatibility complex (MHC) molecules expressed on antigen presentation cells (APCs) through their unique surface receptors (T cell receptors, TCRs) [1]. With their critical role in T cell development and activation, identification of T cell epitopes is very important in studying T cell lineages and phenotypes [2, 3], elucidating self-tolerance mechanisms [4], vaccine design and assessment [5, 6], tracking T cell *in vivo* [7], etc. Much effort has been devoted to identification of these T cell

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1061,
DOI 10.1007/978-1-62703-589-7_15,

epitopes in the last two decades. The most straightforward T cell epitope identification method is to use HPLC-MS to determine the sequence of the peptides extracted directly from the peptide-MHC (pMHC) complexes on the surface of professional APCs [8]. This method has met limited success due to the limited amount of peptides attainable and representation of antigenic peptides as nested sets with heterogeneous lengths.

When the antigenic protein sequence is known, the most commonly used T cell epitope identification method is to chemically synthesize overlapping peptides spanning the entire protein, which are then individually loaded onto professional APCs (usually irradiated peripheral blood mononuclear cells) and tested for their ability of activating T cells [9–11]. This method becomes impractical when the overall size of the antigenic protein increases (e.g., viruses with large genome size) and is not applicable when the antigen sequence is unknown. With the knowledge of MHC-binding motifs, computational algorithms, such as TEPITOPE [12], could be used first to pre-select MHC class II (MHCII)-binding peptides and reduce the number of peptides to be synthesized and tested, but may overlook some candidates due to the complexity of peptide-MHC interaction. As an alternative and more comprehensive approach, combinatorial synthetic peptide libraries with up to a trillion variants, such as positional scanning synthetic combinatorial libraries (PS-SCLs) [13] and bead-bound libraries [14], could be employed. These synthetic combinatorial peptide libraries are valuable especially when the relevant target antigen is unknown. However, generally speaking, all the chemistry-based epitope mapping methods are expensive and nonrenewable, thus several expression cloning strategies have been developed to generate DNA libraries encoding either random peptides or antigenic proteins.

When the restriction MHC binding motif is known, a random peptide library with fixed MHC anchor and variable TCR-contacting residues can be constructed using degenerate primers. As with the combinatorial peptide library-based methods described above, the library screening usually results in the identification of T cell mimotopes (peptides that structurally mimic the antigenic epitope) rather than the native epitopes due to the random nature of the peptide library. While the mimotopes could be used to elucidate the sequence of biologically relevant epitopes [15, 16], they could also have very little resemblance [17]. To directly identify T cell epitopes, several groups have developed methods by either expressing a cDNA library in engineered APCs [18] or expressing viral cDNA libraries in *E. coli* to identify the antigenic protein first [19], so that the epitopes could be identified by testing chemically synthesized overlapping peptides. These expression cloning methods usually require either extensive engineering of a cell line for efficient processing and presenting MHCII restricted antigens that are usually of exogenous origin, or the use of professional APCs

that are often difficult to isolate, expand and maintain. Therefore, the screening process is often time-consuming, laborious, and reagent-intensive.

To address some of the limitations discussed above, we have developed a facile and high throughput CD4+ T cell epitope mapping method by displaying pathogen-derived peptide libraries in complex with the restriction MHCII protein on yeast cell surface [20]. It presents the advantages of both combinatorial peptide libraries and expression cloning, and allows direct epitope identification from either known or unknown pathogens. Here, we use hemagglutinin and influenza virus X31/A/Aichi/68 as an example for each case. Both antigens contain $HA_{306\text{-}318}$ peptide (PKYVKQNTLKLAT) that has been well characterized in complex with the human MHCII allele, DR1. The design of the epitope mapping method is shown in Fig. 1. DNA encoding the pathogenic protein(s) is obtained by PCR or RT-PCR using sequence-specific or random hexamer primers (Fig. 1a) for known and unknown antigens, respectively. For RT-PCR, RNase H and DNA polymerase I are used to generate the second strand cDNA. The resulting double stranded DNA is then randomly digested by DNase I into fragments with size ranging from 30 to 60 nucleotides that encode all the possible peptides with 10–20 amino acids from the antigens, which are then purified and blunt-end polished (T4 DNA polymerase and Klenow fragment). Meanwhile, the yeast display vector containing the gene encoding the restriction single-chain MHCII molecule, e.g., DR1 (Fig. 1b), is digested with NotI/SpeI (*see* **Note 1**), blunt-end polished, and then ligated to the DNA fragments (Fig. 1c). Following gene expression, yeast cells displaying the peptide library in complex of DR1 are analyzed by fluorescence-activated cell sorting (FACS) to enrich those containing peptides with high DR1-binding affinity based on their surface expression levels (*see* **Note 2**). This enrichment cycle is repeated for ~2–4 times to obtain a small subset of yeast cells displaying DR1 in complex with good binding peptides, which are then directly screened for their ability to induce antigen-specific T cell activation as indicated by IL-2 secretion using T hybridoma cells transfected with the TCR of interest (*see* **Note 3**). DNA sequence analysis of the selected positive clones leads to the identification of antigenic peptides. If necessary, a series of DNA fragments encoding overlapping peptide sequences can be used to refine those antigenic peptide sequences.

2 Materials

All buffers and media were prepared using ultrapure water (ddH_2O) with a resistivity of 18.2 mΩ cm at room temperature, and were sterilized by autoclaving at 121 °C and 15 psi for 15 min or

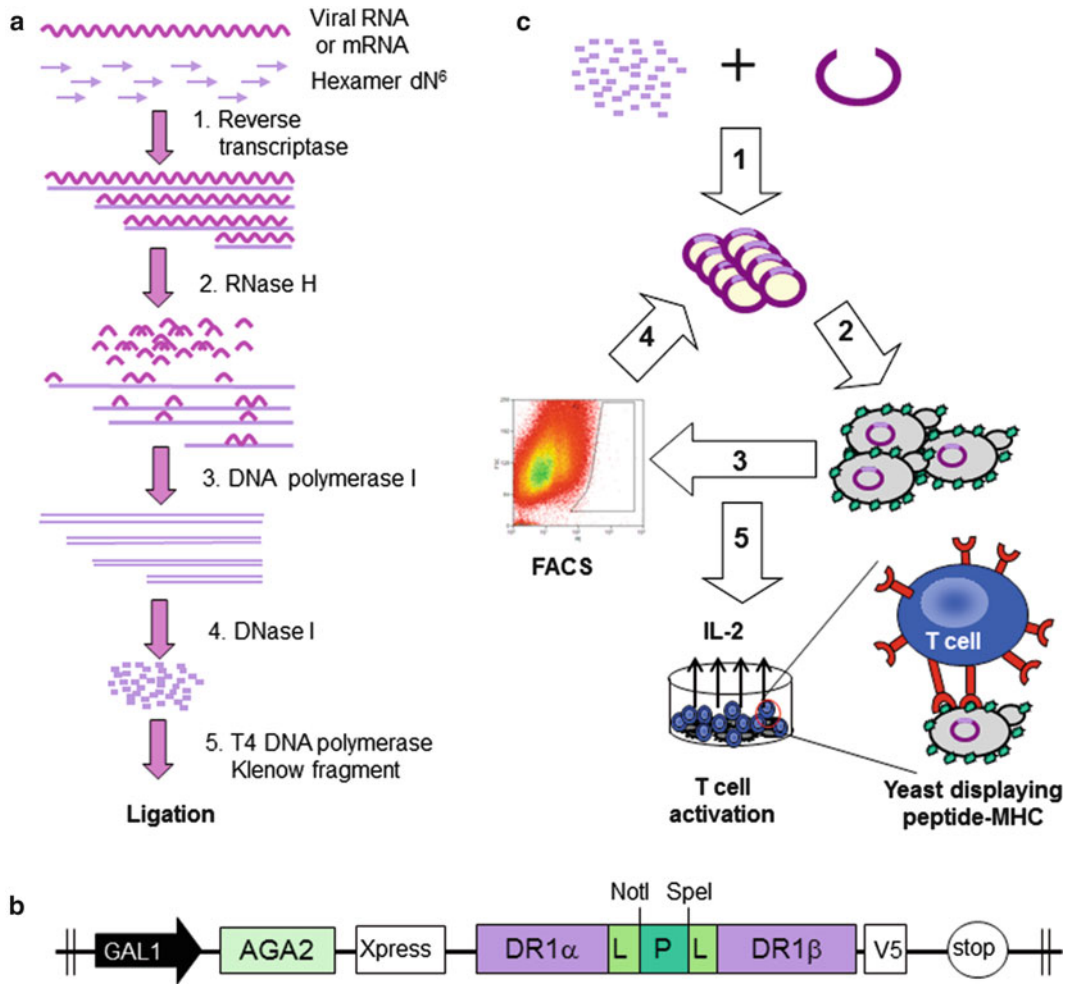

Fig. 1 Design of the CD4+ T cell epitope mapping method. (**a**) Generation of blunt-ended DNA fragments (~30–60 bp) encoding pathogen-derived peptides from genomic RNA or mRNA isolated from pathogens with unknown sequence. For antigens with known sequences, double stranded DNA could be readily amplified using specific primers by PCR and processed from *step 4* with DNase I treatment. (**b**) Schematic representation of the single-chain peptide-DR1 construct used for yeast display. GAL1 = yeast GAL1 promoter, AGA2 = an adhesion subunit of a-agglutinin of a-yeast-cells [25], Xpress = Xpress epitope, L = linker, P = peptide, V5 = V5 epitope. (**c**) Peptide library generation and screening. DNA fragments are ligated to the single-chain MHCII molecules in a yeast display vector (*step 1*). Following transformation and gene expression (*step 2*), yeast cells displaying the peptide library in complex of MHCII protein are analyzed by FACS (*step 3*) to identify a small subset of yeast cells containing peptides with high affinity toward the restriction MHCII. Plasmids are then recovered from these yeast cells (*step 4*) and analyzed for further enrichment. This enrichment cycle is usually repeated for ~2–4 rounds. Individual clones from the enriched library are then screened for their ability to induce antigen-specific T cell activation as indicated by IL-2 production using T cell hybridomas transfected with the TCR of interest (*step 5*) in a 96-well format

filtration through a 0.22 μm membrane. All restriction enzymes were obtained from New England Biolabs (Ipswich, MA) and stored at −20 °C and all antibodies were stored at 4 °C. All primers (listed in Table 1) were synthesized by Integrated DNA

Table 1
Primers used in cloning and peptide library creation

Name	Sequence
α-5BX	5′-GTACCAGGATCCAGTGTGGTGGAAAGGAAAGAAGAACATGTGATC-3′
β-3XH	5′-CCCTCTAGACTCGAGCTTGCTCTGTGCAGATTCAGAC-3′
pYD1For	5′-AGTAACGTTTGTCAGTAATTGC-3′
NotIRev	5′-TGCCAACTTCAGGGTGTTTTGCTTAACATACTTGGGGCGGCCGCCTCC TGAGCCTCCACC-3′
NotIFor	5′-GCAAAACACCCTGAAGTTGGCAACAGGTACCGGTGGCTCACTAG-3′
βRev73-67	5′-GGCCCGCCTCTGCTCCAGGA-3′
StfFor	5′-GGAGGCGGCCGCTTT TTG GATGGAGGAATTCATATG-3′
StfRev	5′-CTCACTAGTCGGGAAGACGTACGGGGTATACATGT-3′
AichiFor	5′-ATTCGCGGCCGCATGAAGACCATCATTGCTTTGAGCTACATTTTC-3′
AichiRev	5′-CTAATAACTAGTAATGCAAATGTTGCACCTAATGTTGCCTCTCTG-3′
2467For	5′-GGCCGCCCCAAGTATAGAAAGATGAACGCACGAAAGTTGGCAACAGGTA CCGGTGGCTCA-3′
2467Rev	5′-CTAGTGAGCCACCGGTACCTGTTGCCAACTTTCGTGCGTTCATCTT TCTATACTTGGGGC-3′
M4For	5′-GGCCGCGGAGGTTATAGACAGATGTCAGCACCAACTTTGGGAGGCGGTA CCGGTGGCTCA-3′
M4Rev	5′-CTAGTGAGCCACCGGTACCGCCTCCCAAAGTTGGTGCTGACATCTGTC TATAACCTCCGC-3′
YAKFor	5′-GGCCGCGCCGCATATGCCGCAGCGGCTGCCGCAAAGGCTGCCGCAGG TACCGGTGGCTCA-3′
YAKRev	5′-CTAGTGAGCCACCGGTACCTGCGGCAGCCTTTGCGGCAGCCGCTGCGG CATATGCGGCGC-3′
CIIFor	5′-GGCCGCGCTGGGTTTAAGGGGGAACAGGGACCTAAAGGAGAGCCTGG TACCGGTGGCTCA-3′
CIIRev	5′-CTAGTGAGCCACCGGTACCAGGCTCTCCTTTAGGTCCCTGTTCCCCCTT AAACCCAGCGC-3′
PKAFor	5′-GGCCGCCCCAAGGCTGTTAAGCAAAACACCCTGAAGTTGGCAACAGGT ACCGGTGGCTCA-3′
PKARev	5′-CTAGTGAGCCACCGGTACCTGTTGCCAACTTCAGGGTGTTTTGCTTAA CAGCCTTGGGGC-3′
LFor	5′-GGCCGCGGAGGTGGAGGCTCCGGAGGTGGAGGCTCAGGAGGTGGAGG TACCGGTGGCTCA-3′
LRev	5′-CTAGTGAGCCACCGGTACCTCCACCTCCTGAGCCTCCACCTCCGGA GCCTCCACCTCCGC-3′
MBPNotI	5′-GGAGGCGGCCGCGAAAACCCGGTTGTTCACTTCTTCAAAAACATCGTT ACCCCGCGTGGTACCGGTGGCTCACTAGTGA-3′

Technologies (Coralville, IA) and stored at −20 °C at a concentration of 100 μM. All other reagents were stored at room temperature unless indicated otherwise.

2.1 Vector Construction and DNA Fragments Generation

1. pJ3/238: A gift from M. Mage (NIH, Bethesda, MD) was used as the PCR template for amplifying the DNA encoding a single-chain polypeptide consisting of the α-chain of DR1, the peptide $HA_{306\text{-}318}$, and the β-chain of DR1 (*see* Fig. 1b).
2. 5× Phusion HF Reaction Buffer and Phusion DNA polymerase (New England Biolabs, Beverly, MA, USA).
3. 40× dNTPs premix: 10 mM each nucleotide (Promega, Madison, WI, USA).
4. 0.5 M ethylenediamine tetraacetic acid (EDTA) solution (pH 8.0): For a 500 mL of stock solution of 0.5 M EDTA, weigh out 93.05 g of EDTA disodium salt (MW = 372.2) and dissolve it in 400 mL of deionized water. Adjust to pH 8.0 with NaOH and correct the final volume to 500 mL. EDTA will not be dissolved completely in water unless the pH is adjusted to about 8.0.
5. 50× Tris-acetate-EDTA (TAE) stock solution: Dissolve 242 g of Tris base (MW = 121.14) in approximately 750 mL of deionized water. Carefully add 57.1 mL of acetic acid and 100 mL of 0.5 M EDTA, and add deionized water to make a final volume of 1 L. The pH of this buffer is not adjusted and should be about 8.5.
6. Working solution of TAE buffer (1×): Dilute the stock solution by 50-fold with deionized water. Final solute concentrations are 40 mM Tris acetate and 1 mM EDTA.
7. 0.7 % (1 %, 2 %) Agarose gel in 1× TAE buffer: Add 0.7 g (1 g, 2 g) of agarose into 100 mL of 1× TAE buffer and microwave until agarose is completely melted. Cool the solution to approximately 70–80 °C. Add 5 μL of ethidium bromide into the solution and mix well. Pour 25–30 mL of solution onto an agarose gel rack with appropriate 2-well (for gel purification) or 8-well (for checking PCR products) combs.
8. QIAquick Gel Extraction Kit (QIAGEN, Valencia, CA, USA).
9. NanoDrop2000c: Used to measure the concentration of DNA (Thermo Scientific, Wilmington, DE, USA).
10. Precision Molecular Mass Standard (Bio-Rad, Hercules, CA).
11. pYD1: Obtained from Invitrogen (Carlsbad, CA) and encodes AGA2, an adhesion subunit of a-agglutinin of a-yeast-cells, which enables yeast surface display of a target protein.
12. QIAquick PCR Purification Kit (QIAGEN, Valencia, CA, USA).

13. T4 DNA ligase and 10× T4 DNA ligase buffer (New England Biolabs, Ipswich, MA).
14. DH5α chemical competent cells: Obtained from Media Preparation Facility (University of Illinois, Urbana, IL) for cloning.
15. 100 mg/mL ampicillin stock solution: Dissolve 1 g of ampicillin powder in 10 mL of ddH_2O and filter-sterilize.
16. LB broth: Add 10 g of bacto-tryptone, 5 g of yeast extract, 10 g of NaCl into 1 L of ddH_2O and autoclave.
17. LB-Amp agar plates: Autoclave LB-agar and when the solution cools down to 70–80 °C, add 1 mL of 100 mg/mL ampicillin to 1 L of LB-agar. Pour 20–25 mL into each Petri dish.
18. QIAprep Miniprep Kit (QIAGEN, Valencia, CA, USA).
19. *S. cerevisiae* EBY100 (MATa ura3-52 trp1 leu2Δ 1 his3Δ 200 pep4::HIS3 prb1Δ1.6R can1 GAL (pIU211:URA3)): Obtained from Invitrogen for yeast surface display of protein of interest and plasmid construction through homologous recombination.
20. Zymoprep II yeast plasmid miniprep (Zymo Research, Orange, CA, USA).
21. 50 mg/mL kanamycin stock solution: Dissolve 1 g of kanamycin powder in 20 mL of ddH_2O and filter-sterilize.
22. SD-CAA medium: Dissolve 20 g dextrose, 6.7 g yeast nitrogen base, 10 g casamino acids in 1 L of ddH_2O and sterilize by autoclaving.
23. SD-CAA agar plates: autoclave SD-CAA-agar and let cool down to 70–80 °C, add 50 μg/mL kanamycin. Pour 20–25 mL into each Petri dish.
24. Optikinase and 10× reaction buffer: Obtained from USB (Cleveland, OH) and was used to phosphorylate DNA inserts for ligation.
25. Influenza A/Aichi/2/68 viral RNA: A gift from R. Donis (CDC, Atlanta, GA) was used as template to amplify the hemagglutinin gene.
26. SuperScriptIII One-Step RT-PCR System with Platinum Taq High Fidelity (Invitrogen, Carlsbad, CA).
27. Transcriptor First Strand cDNA Synthesis Kit (Roche, Indianapolis, IN).
28. DNase I, DNA polymerase I, RNase H, T4 DNA polymerase, Klenow fragment, and NEBuffer 2 were obtained from New England Biolabs (Ipswich, MA) and stored at –20 °C.
29. QIAEX II Gel Extraction Kit (QIAGEN, Valencia, CA, USA).
30. 10 mg/mL bovine serum albumin (BSA) solution (New England Biolabs, Ipswich, MA).

31. Shrimp alkaline phosphatase (SAP) and 10× SAP buffer (Fermentas, Glen Burnie, Maryland).
32. ElectroMax DH5α competent cells: Obtained from Invitrogen (Carlsbad, CA) for peptide library construction by electroporation.
33. Influenza virus X31/A/Aichi/68 genomic RNA: Obtained from Charles River Laboratories (Wilmington, MA) and used as the template for cDNA synthesis.

2.2 Cell Growth and Transformation

1. YPAD medium: Dissolve 10 g of yeast extract, 20 g of peptone, 20 g of dextrose, and 100 mg of adenine hemisulfate in 1 L of ddH_2O and sterilize by autoclaving.
2. LiAc solutions: To prepare 50 mL of 1 M solution, dissolve 3.3 g lithium acetate (MW = 65.99 g/mol) in ddH_2O and sterilize by filtration. Prepare 0.1 M solution by mixing 5 mL of the 1.0 M solution with 45 mL sterile ddH_2O.
3. ssDNA stock solution: Dissolve 200 mg of deoxyribonucleic acid Sodium Salt Type III from Salmon Testes (Sigma, St. Louis, MO) in 100 mL of Tris–EDTA (TE) buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) by vigorous mixing on a magnetic stirrer. Aliquot the ssDNA in 1 mL and store at −20 °C (*see* **Note 4**).
4. 50 % (w/v) polyethylene glycol (PEG): Add 50 g PEG (MW = 3,350 g/mol) into 30 mL ddH_2O, heat the solution to ~60 °C with constant stirring to ensure quick dissolution of PEG and once completely dissolved, adjust volume with ddH_2O to 100 mL and sterilize by filtration.
5. Gene Pulser II and Pulse controller plus: Obtained from Bio-Rad (Hercules, CA) and used to transform plasmids or ligation mixture into *E. coli* through electroporation.
6. 1 M glucose solution: Dissolve 90 g of d-glucose in 400 mL of ddH_2O, adjust to a final volume of 500 mL, and filter-sterilize it.
7. SOC medium: Add 20 g of Bacto-tryptone, 5 g of yeast extract, 0.5 g of NaCl, 186.4 mg of KCl into 980 mL of ddH_2O. Adjust the pH to 7.0 with NaOH. Autoclave at 121 °C for 15 min. After the solution cools down to 70–80 °C, add 20 mL of sterile 1 M glucose.

2.3 Expression and Function Analysis of pMHC Complexes on Yeast Cell Surface

1. YPG medium: Dissolve 10 g yeast extract, 20 g peptone and 20 g galactose in 1 L of ddH_2O and sterilize by autoclaving.
2. Phosphate-buffered saline (PBS) (VWR, Radnor, PA).
3. Bovine serum albumin (BSA): Obtained from Sigma (St. Louis, MO) and used at a final concentration of 0.5 % in PBS for all antibody staining and washing steps.

4. Antibodies: anti-V5 (Invitrogen, Carlsbad, CA), LB3.1 (American Tissue Culture Collection (ATCC), Manassas, VA), and goat-anti-mouse (GAM) IgG (Rockland, Gilbertsville, PA).
5. Streptavidin–phycoerythrin (SA-PE) conjugate (eBiosciences, San Diego, CA).
6. HA1.7 hybridoma T cells: A gift from J. Bill [17] was used as an indicator cell line for productive DR1-peptide-TCR interactions that produce IL-2 upon receptor engagement.
7. Complete Iscove's Modified Dulbecco's Media (IMDM) for HA1.7 culture: IMDM (Invitrogen, Carlsbad, CA) supplemented with 10 % fetal bovine serum (Biomeda, Forster City, CA) and Penicillin (10 U/mL), streptomycin (10 μg/mL).
8. Murine IL-2 enzyme-linked immunosorbent assay (ELISA) kit (eBiosciences, San Diego, CA).

3 Methods

3.1 Vector Construction for Yeast Surface Display of Single-Chain pMHC Complexes

1. Amplify DNA encoding single-chain DR1α-linker-$HA_{306-318}$-linker-DR1b using α-5BX and β-3XH as primers (*see* Table 1) and plasmid pJ3/238 as template. Set up the reaction mixture as following: 10 ng of DNA template, 1× HF Phusion buffer, 50 pmol of forward primer, 50 pmol of reverse primer, 0.25 mM (each) dNTPs, 2 U of Phusion DNA polymerase, and add ddH_2O to make up a final volume of 100 μL. Unless otherwise specified, this recipe was used for all PCR reactions.
2. PCR thermocycler program: an initial denaturation of 2 min at 98 °C, followed by 25–30 cycles of denaturation at 98 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 60 s, and a final 10-min elongation at 72 °C. Unless otherwise specified, this program was used for all PCR reactions with the addition of extension time of 30 s per additional 1 kb long gene.
3. Load the PCR reaction mixture onto 0.7 % agarose gels and perform electrophoresis in 1× TAE buffer at 120 V for 20 min.
4. Gel-purify the DNA band with correct size using the QIAquick Gel Extraction Kit and check the concentration using NanoDrop, as per the manufacturer's instructions (*see* **Note 5**).
5. Digest 1 μg of the purified PCR product (insert) and 1 μg of plasmid pYD1 (vector) with BstXI and XhoI and purify using the QIAquick PCR Purification Kit.
6. Ligate the digested insert and vector to generate plasmid pYD1scαHAβ at 16 °C overnight. Set up the ligation mixture as following: 50 ng of vector, 100 ng of insert, 1× T4 ligase buffer, 1 μL of T4 ligase, and add ddH_2O to make up a final volume of 20 μL.

7. Transform 5 μL of the ligation mixture into chemical competent DH5α cells (*see* **Note 6**) and select transformants on a LB-Amp+ agar plate in a 37 °C incubator.
8. Pick a single colony into 3 mL LB medium with 100 μg/mL ampicillin, grow at 37 °C with 250 rpm agitation overnight, isolate the plasmid using the QIAprep Miniprep Kit, as per the manufacturer's instructions, and confirm the sequence of pYD1scαHAβ by DNA sequencing.
9. Introduce a NotI restriction site directly upstream of $HA_{306\text{-}318}$ (*see* Fig. 1b) to facilitate swapping of different peptides in the single-chain pMHC construct as follows: Perform two PCR reactions using pYD1scαHAβ as template and pYD1For/NotIRev and NotIFor/βRev73-67 as primers (*see* Table 1), respectively. Purify the PCR products using QIAquick PCR Purification Kit and splice them as described elsewhere [21]. Co-transform the spliced DNA and pYD1scαHAβ digested with BstXI and SpeI into *S. cerevisiae* EBY100 to generate pYD1αHAβ. Refer to Subheading 3.2 for detailed yeast transformation protocol. Isolate pYD1αHAβ from a 3 mL yeast culture in SD-CAA using the Zymoprep II yeast plasmid kit.
10. Transform 5 μL of the yeast plasmid into chemical competent DH5α cells (*see* **Note 7**), and isolate the plasmid for DNA sequencing as described in **steps 7** and **8** to confirm the sequence of pYD1αHAβ.
11. Construct plasmid pYD1αSTFβ that contains a segment of stuffer DNA in place of the $HA_{306\text{-}318}$ peptide in pYD1αHAβ (*see* **Note 1**): The stuffer DNA (STF) is an unrelated gene (~1 kb) amplified from the phosphite dehydrogenase gene [22] using primers StfFor and StfRev (*see* Table 1). The PCR product was cloned into pYD1αHAβ via NotI and SpeI to generate plasmid pYD1αSTFβ.
12. To establish the correlation between the yeast surface expression level of the peptide-DR1 complexes and the peptide-binding affinity (*see* **Note 2**), plasmids pYD1α2467β, pYD1αM4β, pYD1αYAKβ, pYD1αCIIβ, pYD1αPKAβ, pYD1αLβ and pYD1αMBPβ were constructed in two steps. First, phosphorylate the oligonucleotides by Optikinase (USB, Cleveland, OH). Set up the phosphorylation reaction as follows: 2 μL of oligonucleotides (2.5 μM), 1× Optikinase buffer, 2.5 μL of dATP (10 mM), 1 μL of Optikinase and add ddH_2O to make a final volume of 25 μL. Incubate the mixture at 37 °C for 2.5–3.5 h and 65 °C for 15 min. Second, generate the double stranded DNA insert encoding the peptide by hybridization of the forward and reverse oligonucleotides (*see* Table 1). For pYD1αMBPβ, generate the insert using a self-annealing oligonucleotide MBPNotI. Ligate the resulting insert into pYD1αSTFβ digested by NotI and SpeI.

3.2 Yeast Transformation

1. Inoculate a single colony of *S. cerevisiae* EBY100 into 3 mL of YPAD medium and grow overnight in a shaker at 30 °C and 250 rpm (*see* **Note 8**).
2. Measure the OD_{600} of the overnight culture and inoculate the appropriate amount to 50 mL (this is enough for ten transformations, scale up or down proportionally for more or less transformations) of fresh YPAD medium to obtain an OD_{600} of 0.2.
3. Continue to grow the 50 mL of culture for approximately 4–5 h to obtain an OD_{600} of 0.8 (*see* **Note 9**).
4. Wash the cells once with 25 mL of sterile ddH_2O (3200 rcf, 5 min), once with 1 mL of 0.1 M LiAc (3350 rcf, 30 s), add 400 μL of 0.1 M LiAc and resuspend (it should result in a total of 500 μL of resuspended cells), and aliquot 50 μL into sterile Eppendorf tubes. Each tube is used for one transformation.
5. In the meantime, boil 1 mL of ssDNA stock solution in a water bath for 5 min and chill immediately on ice.
6. Spin down the 50 μL cells (3350 rcf, 15 s) and discard the supernatant.
7. Prepare the transformation mixture immediately as the following: Add 240 μL of 50 % PEG, 36 μL of 1 M LiAc, 50 μL of boiled ssDNA stock solution, 0.1–10 μg of plasmid DNA and add sterile ddH_2O to make a final volume of 360 μL (*see* **Note 10**). For homologous recombination, 5–20 ng of the digested vector and a 10–20-fold molar excess of the insert are used.
8. Mix the transformation mixture extensively by vigorous vortexing, incubate at 42 °C for 40 min, spin down the cells (3350 rcf, 30 s), and remove the transformation mixture carefully.
9. Resuspend the cell pellet in 1 mL ddH_2O gently with pipetting.
10. Plate 50–200 μL of the resuspension on a SD-CAA agar plate and incubate at 30 °C for 2–3 days until colonies appear.

3.3 Expression Analysis of pMHC Complexes on Yeast Surface

1. Grow *S. cerevisiae* EBY100 clones transformed with different plasmid constructs in 3 mL of SD-CAA at 30 °C until OD600 reaches ~5.
2. Wash the yeast cells twice (3350 rcf, 1 min) with 1 mL YPG medium.
3. To induce AGA2 fusion protein expression, resuspend the cells to an OD_{600} ~1.0 in 3 mL of YPG medium supplemented with 50 μg/mL kanamycin and incubate in a refrigerated rotary shaker at 20 °C and 250 rpm for ~48 h. The induced yeast cells could be stored at 4 °C up to 4 months for repeated analysis (*see* **Note 11**).

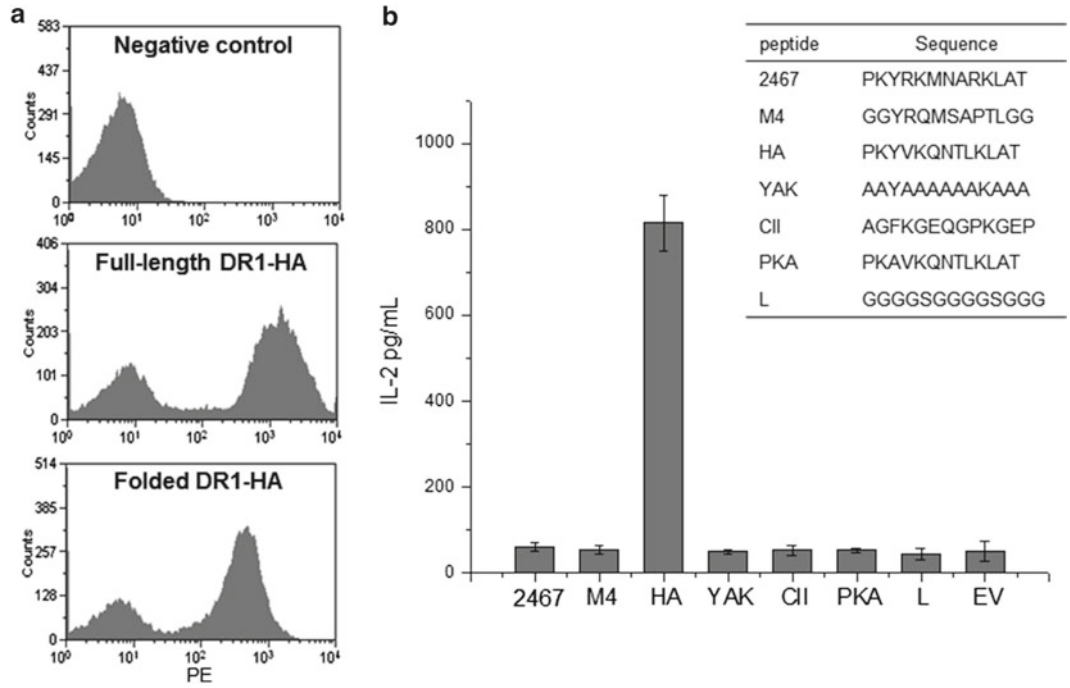

Fig. 2 Expression and function analysis of peptide-DR1 complexes displayed on yeast cell surface. (**a**) Yeast cells transformed with pYD1-scDR1αHAβ were stained with anti-V5 antibody to detect full-length protein expression, or with LB3.1 antibody to detect correctly folded complexes. Yeast cells stained only with a secondary antibody were used as a negative control. (**b**) Yeast cells displaying the peptide-DR1 complexes activated HA1.7 hybridoma in an antigen-specific manner. Yeast cells containing the empty plasmid pYD1 (EV) were cultured, induced, and analyzed in the same way as yeast displaying other constructs. Sequences of peptides used in the scDR1αpepβ constructs are shown in the *inset*

4. For flow cytometric analysis, wash ~2.5×10^6 cells in a 96-well V-bottom plate with 180 μL of PBS containing 0.5 % BSA and incubate with the primary antibody (anti-V5 for full-length protein detection and LB3.1 for correctly folded DR1 detection, *see* Fig. 2a) at 4 °C for 1 h at a 100-fold dilution rate.
5. Wash cells once with 180 μL of PBS + 0.5 % BSA and incubate with biotinylated GAM IgG (1:100 dilution) at 4 °C for 1 h.
6. Wash cells once with 180 μL of PBS + 0.5 % BSA and incubate with SA-PE (1:100 dilution) at 4 °C for 30 min.
7. Wash cells three times with PBS + 0.5 % BSA to remove unbound SA-PE.
8. Analyze fluorescently labeled yeast cells on a Coulter Epics XL flow cytometer at the Biotechnology Center of University of Illinois at Urbana-Champaign as per facility protocols (Urbana, IL) (*see* **Note 12**).

3.4 Stimulation of T Cell Hybridoma and IL-2 Detection

1. Aliquot ~10^6 yeast cells induced with YPG into 300 μL of PBS in a 96-well R-bottom tissue-culture plate. Prepare a triplicate for each pMHC construct.
2. Wash once with PBS and resuspend in 300 μL of PBS.
3. Incubate the plate at 4 °C overnight to allow the yeast cells to attach to the surface.
4. Wash away unbound yeast cells three times with 350 μL of PBS.
5. Wash HA1.7 hybridoma cells once with 10 mL of warm complete IMDM medium and resuspend to a density of 10^5 per 300 μL complete IMDM.
6. Add 300 μL of the HA1.7 hybridoma cell resuspension to each well and incubate for ~24 h at 37 °C with 5 % CO_2.
7. Spin down the cells and test the supernatant for IL-2 production using the murine IL-2 ELISA kit. IL-2 should only be detected in the supernatant from the wells where the yeast cells displaying functional peptide-MHCs specific for HA1.7 TCR, i.e., DR1-$HA_{306\text{-}318}$ (*see* Fig. 2b).

3.5 Peptide Library Construction from Hemagglutinin

1. Amplify DNA encoding hemagglutinin by one-step RT-PCR using influenza A/Aichi/2/68 viral RNA as the template and AichiFor and AichiRev as primers (Table 1). Set up the reaction mixture as follows: 2 μL of RNA, 1× reaction buffer, 0.2 μM of AichiFor, 0.2 μM of AichiRev, 1 μL of SuperscriptIII, and add ddH_2O to make up a final volume of 50 μL.
2. The thermocycler program: 55 °C for 30 min, an initial denaturation of 2 min at 94 °C, followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, and extension at 68 °C for 2 min, and a final 10-min elongation at 68 °C.
3. Gel-purify the DNA band with correct size using the QIAquick Gel Extraction Kit and quantify the concentration using NanoDrop, as per the manufacturer's instructions.
4. Fragmentize the purified DNA using DNase I as the following: mix 1 μg of DNA, 5 μL of 500 mM Tris–HCl, pH 7.4, 2.5 μL of 200 mM MnCl2, and ddH_2O to a total volume of 42.5 μL. Incubate the mixture at 15 °C for 15 min. Freshly dilute DNase I with ddH_2O to 0.01 U/μL and add 7.5 μL to the DNA mixture and incubate at 15 °C for 1 min. Heat inactivate at 90 °C for 10 min (*see* **Note 13**).
5. Run the DNase I digestion mixture immediately on a 2 % agarose gel and purify the DNA fragments with size between 30 and 60 bp using the QIAEX II Gel Extraction Kit, as per manufacturer's instructions.

6. Blunt-end polish the purified DNA fragments as follows: Mix 0.5 μL of 10 mg/mL BSA, 1× NEBuffer 2, 5 μL of 1 mM dNTP, 100 ng of DNA fragments, 0.1 U of T4 DNA polymerase and add ddH_2O to a final volume of 50 μL. Incubate at 25 °C for 10 min and add 1.25 μL of Klenow fragment (5 U/μL). Incubate at 25 °C for 10 min and 16 °C for 105 min.
7. Gel-purify the blunt-end polished DNA fragments using the QIAEX II Gel Extraction Kit.
8. In the meantime, blunt-end polish pYD1αSTFβ digested by SpeI and NotI in the same way as described in **step 5**, gel-purify using QIAquick Gel Extraction Kit, and dephosphorylate using 1 μL of SAP per 1 μg of DNA at 37 °C for 90 min. Heat inactivate at 70 °C for 10 min and clean up the mixture using the QIAquick PCR Purification Kit.
9. Ligate the DNA fragments from **step 6** into the vector from **step 7** as follows (*see* **Note 14**): Mix 1 μL of T4 DNA ligase buffer, 50 ng of vector, twofold molar excess of DNA fragments, 1.5 μL of PEG 8000, 1 μL of T4 DNA ligase, and add ddH_2O to a final volume of 10 μL. Incubate the ligation mixture at 16 °C for 16–20 h.
10. Clean up the ligation mixture as follows: add *n*-butanol to make a final volume of 500 μL, mix vigorously by vortexing for 30 s, centrifuge at the maximum speed in a benchtop centrifuge for 10 min, immediately remove the supernatant as much as possible, place the tube in a chemical hood for~1 h to dry, and resuspend the ligation mixture in 2 μL of ddH_2O (this is enough for one electroporation).
11. Transform the ligation mixture into ElectroMax DH5α competent cells by electroporation using Gene Pulser II.
12. After electroporation, immediately recover cells in 1 mL of pre-warmed SOC medium at 37 °C with 250 rpm agitation for 1 h.
13. For library creation, perform multiple electroporations, pool the cells into 400 mL LB medium supplemented with 100 μg/mL ampicillin, and grow overnight at 37 °C with 250 rpm agitation. For example, a library of 4.6×10^5 clones was obtained with ten electroporations in our study (*see* **Note 15**).
14. Isolate the plasmids using the QIAprep Miniprep Kit, as per manufacturer's instructions.
15. Perform plasmid transformations into the EBY100 competent yeast cells as described in Subheading 3.2 and pool the cells into 400 mL of SD-CAA medium. With 40 transformations, we obtained a library of 2.1×10^6.
16. After two passages in 400 mL of SD-CAA medium, induce protein expression using YPG as described in Subheading 3.3, and continue with FACS screening.

3.6 Peptide Library Construction from Influenza A Virus

1. Synthesize first strand cDNA using Transcriptor First Strand cDNA Synthesis Kit with random priming. Set up the reaction as follows: 2 μL of X31/A/Aichi/68 genomic RNA, 2 μL of hexamer, and 9 μL of PCR grade H_2O. Incubate the mixture at 65 °C for 10 min and put on ice immediately. Add 0.5 μL of PCR grade H_2O, 4 μL reaction buffer, 2 μL of dNTP, and 0.5 μL of Transcriptor RTase. Incubate the mixture at 25 °C for 10 min, 55 °C for 30 min, 85 °C for 5 min, and then cool down to 4 °C.
2. Synthesize the second strand DNA by incubating the cDNA with 10 units of DNA polymerase I, 0.32 units of RNase H, 0.25 mM dNTP, and 1× NEBuffer 2 in a final volume of 40 μL at 15 °C for 90 min.
3. Run the double stranded DNA on a 1 % agarose gel and purify the bands with different size ranges separately: <500 bp, 500 bp to 1 kb, and >1 kb.
4. Quantify the purified DNA using the Precision Molecular Mass Standard as per the manufacturer's instructions.
5. Fragmentize the DNA with DNase I using the same method as described in **step 4** of Subheading 3.6 except for approximately 6, 0.75, 1.7 units per 1 μg of DNA with size <500 bp, 500 bp to 1 kb, and >1 kb, respectively (*see* **Note 13**).
6. Generate the yeast display library for FACS screening using the same method as described in **steps 5–16** of Subheading 3.6.

3.7 Yeast Display Library Screening

1. Take an aliquot of the of induced peptide library culture (e.g., 7.4×10^6 yeast cells for the hemagglutinin library and 10^8 for the influenza A virus library) and stain with LB3.1 antibody followed by biotinylated GAM IgG and SA-PE as described in Subheading 3.3.
2. Sort the cells on a Coulter 753 bench FACS sorter (Flow Cytometry Facility, University of Illinois at Urbana-Champaign) and collect ~1.5 % of the population with the highest fluorescence (*see* Figs. 3a and 4a) in SD-CAA medium.
3. After protein expression induction in YPG, perform another two rounds of cell sorting (*see* **Note 16**) in the same way as described in **steps 1–2** except that the top 0.5 % of the population is collected into SD-CAA medium.
4. Induce protein expression in YPG and sort the top 0.5 % of the population into SD-CAA medium in a 96-well plate with no more than one cell in a well.
5. For hemagglutinin-derived peptide library, randomly pick ten of these cells and examine surface protein display level and their ability to stimulate T cell hybridoma HA1.7 as described in Subheadings 3.3 and 3.4 (*see* Fig. 3b, c). To determine the sequence of the peptide insert, perform DNA sequencing analysis.

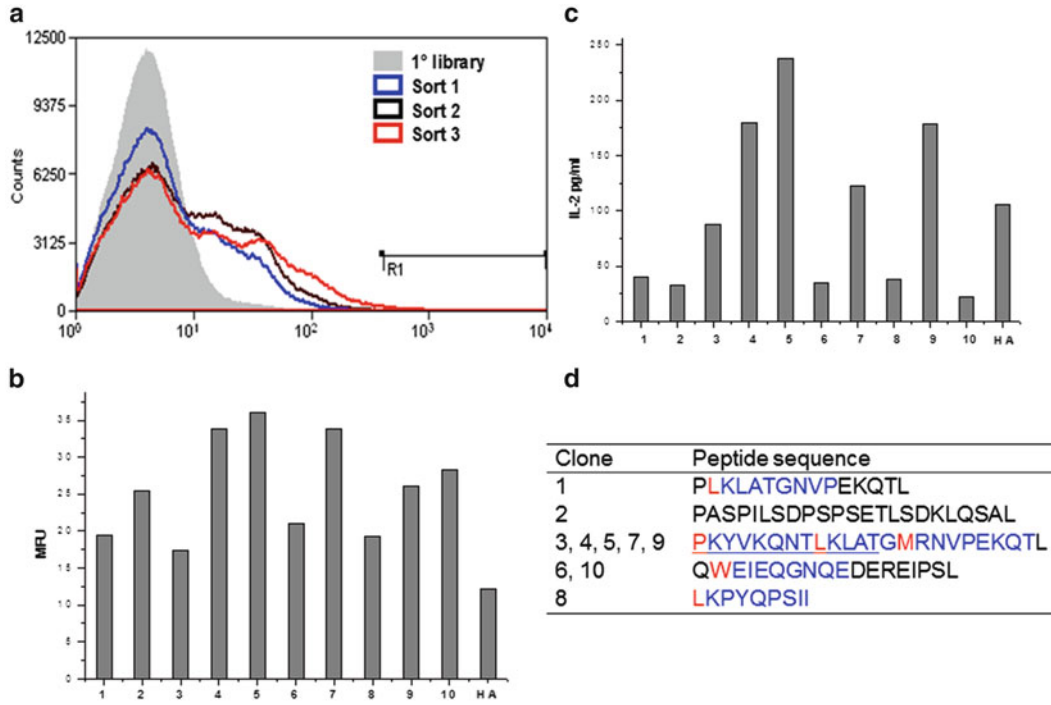

Fig. 3 Direct epitope identification from a single antigenic protein with known sequence—hemagglutinin. (**a**) FACS enrichment of potential good binders from the hemagglutinin-derived peptide library. LB3.1 antibody was used to stain cells as a measurement of surface expression level of correctly folded peptide-DR1 complexes. Surface expression (**b**), T cell hybridoma activation (**c**), and DNA sequence (**d**) analysis of ten clones randomly picked from the library after three rounds of cell sorting. The predicted binding 9-residue peptides are shown in blue with red letters corresponding to the amino acid residue at position P1

6. For influenza A virus-derived peptide library, sort four 96-well plates of cells from the top 0.5 % of the population and analyze them in the same way as described in **step 5** (*see* Fig. 4b, c).
7. Align peptide sequences from the active clones to identify epitope sequences using ClustalW (http://embnet.vital-it.ch/software/ClustalW.html) (*see* Figs. 3d and 4d).

4 Notes

1. To eliminate the possibility that the epitope $HA_{306\text{-}318}$ identified from the peptide library is derived from the undigested pYD1αHAβ, it is important to use plasmid pYD1αSTFβ that contains a segment of unrelated stuffer DNA in place of the $HA_{306\text{-}318}$ peptide as the backbone for blunt end ligation. In addition, the 1 kb stuffer DNA enables clear separation of digested plasmid by DNA electrophoresis.

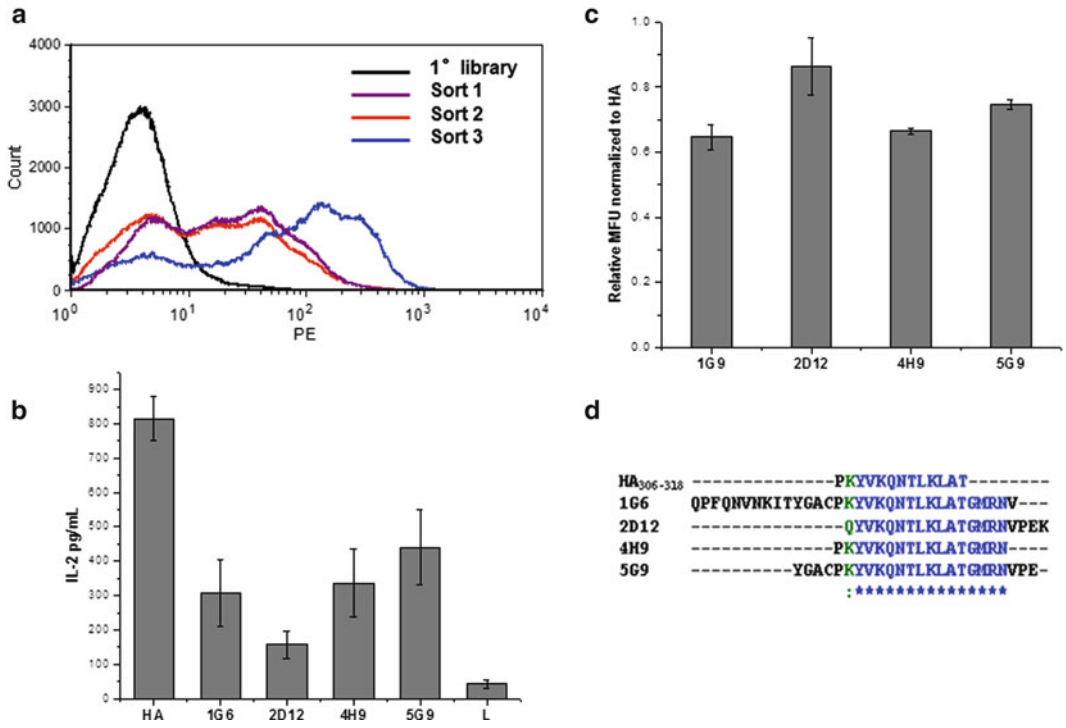

Fig. 4 Direct epitope identification from a pathogen with unknown sequence—influenza A virus. (**a**) FACS enrichment of potential DR1-binding peptides using LB3.1. (**b**) T-cell activation analysis of the four clones identified from the enriched peptide library. (**c**) The four active clones showed comparable surface expression levels as yeast displaying scDR1αHAβ. The relative mean fluorescence unit (MFU) was normalized to HA for direct comparison. (**d**) The peptide sequences from the four clones were aligned with the HA306-318 epitope sequence. *Asterisk* = fully conserved residues, *colon* = conservation of strong groups

2. A positive correlation between the yeast surface expression level of the peptide-DR1 complexes and the peptide-binding affinity was established [20]. Therefore, the yeast surface expression level of the single chain DR1-peptide protein, as measured by DR-specific antibody (LB3.1) staining, can be used as a proxy screening variable for DR1-binding peptides, significantly reducing the number of clones required in the function screening assay. This stabilizing effect of a binding peptide was also observed for DR4 [20] and DR2 [23].
3. The ability of yeast cells displaying DR1-$HA_{306\text{-}318}$ to activate T hybridoma cell HA1.7 greatly simplified the functional screening for epitope identification. This could be generally applicable to MHCII alleles, such as DR4 [20]. However, due to the polygenic and polymorphic nature of MHC proteins and the diversity of the binding peptides, it is difficult to establish that yeast cells could display all or most of pMHC complexes in a functional form to engage specific TCRs, such as low affinity self-reactive ones (e.g., DR2-$MBP_{85\text{-}99}$-Ob.1A12 [23]).

4. It is not desirable to freeze–thaw the carrier ssDNA frequently and usually it is discarded after 3–4 times of thawing.
5. When the DNA concentration is low (below 20 ng/μL), run 3–5 μL on a 0.7 % agarose gel and use the Precision Molecular Mass Standard for a more accurate estimation according to the manufacturer's instructions.
6. DH5α was used for DNA cloning in our experiments. However, any *E. coli* strain with recA and endA, such as Top10 and JM109, can also be used.
7. The number of DH5α transformants could vary from a few to several thousands. This is mainly due to the low quality of the isolated yeast plasmids. Sometimes, no colony was even obtained. In such cases, repeat the DH5α transformation.
8. Yeast competent cells need to be freshly prepared each time.
9. Normally, the doubling time for a *S. cerevisiae* laboratory strain is approximately 2 h.
10. It is important to follow the order by which the transformation mixture components are added, especially for 50 % PEG, which shield the competent cells from the toxic 1 M LiAc.
11. The storage time might vary depending on the stability of the target protein displayed on yeast cell surface. For the pMHC complexes in our study, no significant degradation was observed within 4 months.
12. To compare surface expression levels of different pMHC complexes, it is important to induce the protein expression, and perform the antibody staining in a single experiment run to minimize variation.
13. Since DNase I has very high activity, it is necessary to test and optimize the digestion conditions for different target DNA preparations and for different lots of DNase I.
14. The difficulties in cloning small pieces of DNA include: (a) optimization of the DNase I digestion step to have majority of the digested DNA fragments with a desired size range; (b) intramolecular circularization, which results in a high ligation background; (c) occurrence of multiple inserts, which results in splicing peptides that are not present in the natural peptide repertoire. Accordingly, when constructing the peptide library, it is important to (a) optimize the DNase I concentration and the digestion time; (b) dephosphorylate the vector and include 15 % PEG in the ligation mixture to promote intermolecular ligation over intramolecular circularization [24]; and (c) optimize the insert to vector molar ratio.
15. It is important to sample the diversity of the primary peptide library created in *E. coli*. To do so, an aliquot of the library is selected on a LB-Amp agar plate to obtain individual colonies

(this also allows an estimate of the library size) and 20 clones are randomly picked, sequenced, and mapped to the viral genome. There should be at least one peptide derived from each of the eight pieces of influenza viral genomic RNA displayed in either sense or antisense orientations, indicating a good representation of the whole viral genome.

16. The number of sorting cycles varies depending on how much improvement of the expression the enriched library has over the previous round. For example, there was a steady increase in the fluorescence intensity over the sorting rounds in Fig. 3a, so a total of four rounds of enrichment (the fourth round is indicated by region R1) were performed.

References

1. Germain RN (1994) MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. Cell 76:287–299
2. Altman JD, Moss PA, Goulder PJ, Barouch DH, McHeyzer-Williams MG, Bell JI, McMichael AJ, Davis MM (1996) Phenotypic analysis of antigen-specific T lymphocytes. Science 274:94–96
3. Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang YH, Wang Y, Hood L, Zhu Z, Tian Q, Dong C (2005) A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. Nat Immunol 6: 1133–1141
4. Klein L, Kyewski B (2000) "Promiscuous" expression of tissue antigens in the thymus: a key to T-cell tolerance and autoimmunity? J Mol Med 78:483–494
5. Coulie PG, Karanikas V, Lurquin C, Colau D, Connerotte T, Hanagiri T, Van Pel A, Lucas S, Godelaine D, Lonchay C, Marchand M, Van Baren N, Boon T (2002) Cytolytic T-cell responses of cancer patients vaccinated with a MAGE antigen. Immunol Rev 188: 33–42
6. Huang CH, Peng S, He L, Tsai YC, Boyd DA, Hansen TH, Wu TC, Hung CF (2005) Cancer immunotherapy using a DNA vaccine encoding a single-chain trimer of MHC class I linked to an HPV-16 E6 immunodominant CTL epitope. Gene Ther 12:1180–1186
7. Jenkins MK, Khoruts A, Ingulli E, Mueller DL, McSorley SJ, Reinhardt RL, Itano A, Pape KA (2001) In vivo activation of antigen-specific CD4 T cells. Annu Rev Immunol 19:23–45
8. Henderson RA, Cox AL, Sakaguchi K, Appella E, Shabanowitz J, Hunt DF, Engelhard VH (1993) Direct identification of an endogenous peptide recognized by multiple HLA-A2.1-specific cytotoxic T cells. Proc Natl Acad Sci USA 90:10275–10279
9. Novak EJ, Liu AW, Gebe JA, Falk BA, Nepom GT, Koelle DM, Kwok WW (2001) Tetramer-guided epitope mapping: rapid identification and characterization of immunodominant CD4+ T cell epitopes from complex antigens. J Immunol 166:6665–6670
10. Babon JA, Cruz J, Orphin L, Pazoles P, Co MD, Ennis FA, Terajima M (2009) Genome-wide screening of human T-cell epitopes in influenza A virus reveals a broad spectrum of CD4(+) T-cell responses to internal proteins, hemagglutinins, and neuraminidases. Hum Immunol 70:711–721
11. Richards KA, Chaves FA, Sant AJ (2009) Infection of HLA-DR1 transgenic mice with a human isolate of influenza a virus (H1N1) primes a diverse CD4 T-cell repertoire that includes CD4 T cells with heterosubtypic cross-reactivity to avian (H5N1) influenza virus. J Virol 83:6566–6577
12. Bian H, Reidhaar-Olson JF, Hammer J (2003) The use of bioinformatics for identifying class II-restricted T-cell epitopes. Methods 29: 299–309
13. Sospedra M, Pinilla C, Martin R (2003) Use of combinatorial peptide libraries for T-cell epitope mapping. Methods 29:236–247
14. Hiemstra HS, Duinkerken G, Benckhuijsen WE, Amons R, de Vries RR, Roep BO, Drijfhout JW (1997) The identification of CD4+ T cell epitopes with dedicated synthetic peptide libraries. Proc Natl Acad Sci USA 94:10313–10318
15. Nino-Vasquez JJ, Allicotti G, Borras E, Wilson DB, Valmori D, Simon R, Martin R, Pinilla C

(2004) A powerful combination: the use of positional scanning libraries and biometrical analysis to identify cross-reactive T cell epitopes. Mol Immunol 40:1063–1074
16. Wang Y, Rubtsov A, Heiser R, White J, Crawford F, Marrack P, Kappler JW (2005) Using a baculovirus display library to identify MHC class I mimotopes. Proc Natl Acad Sci USA 102:2476–2481
17. Boen E, Crownover AR, McIlhaney M, Korman AJ, Bill J (2000) Identification of T cell ligands in a library of peptides covalently attached to HLA-DR4. J Immunol 165: 2040–2047
18. Wang RF, Wang X, Atwood AC, Topalian SL, Rosenberg SA (1999) Cloning genes encoding MHC class II-restricted antigens: mutated CDC27 as a tumor antigen. Science 284: 1351–1354
19. Koelle DM (2003) Expression cloning for the discovery of viral antigens and epitopes recognized by T cells. Methods 29:213–226
20. Wen F, Esteban O, Zhao H (2008) Rapid identification of CD4+ T-cell epitopes using yeast displaying pathogen-derived peptide library. J Immunol Methods 336:37–44
21. Horton RM, Cai ZL, Ho SN, Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques 8:528–535
22. Woodyer R, van der Donk WA, Zhao H (2003) Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. Biochemistry 42:11604–11614
23. Wen F, Sethi DK, Wucherpfennig KW, Zhao H (2011) Cell surface display of functional human MHC class II proteins: yeast display versus insect cell display. Protein Eng Des Sel 24:701–709
24. Upcroft P, Healey A (1987) Rapid and efficient method for cloning of blunt-ended DNA fragments. Gene 51:69–75
25. Boder ET, Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol 15:553–557