Journal of Immunological Methods 336 (2008) 37–44

Contents lists available at ScienceDirect

Journal of Immunological Methods

journal homepage: www.elsevier.com/locate/jim

Research paper

Rapid identification of CD4+ T-cell epitopes using yeast displaying pathogen-derived peptide library

Fei Wen [a], Olga Esteban [a], Huimin Zhao [a,b,*]

[a] Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States
[b] Departments of Chemistry, Biochemistry, and Bioengineering, Institute for Genomic Biology, Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

ARTICLE INFO

Article history:
Received 2 August 2007
Received in revised form 13 March 2008
Accepted 13 March 2008
Available online 14 April 2008

Keywords:
T-cell epitopes
Single-chain peptide–MHC complex
Yeast display
Flow cytometry

ABSTRACT

Identification of CD4+ T-cell epitopes is a critical step in studying and modulating the immune responses to tumors, infectious agents, and autoantigens. Here we report a facile, accurate, and high-throughput method for CD4+ T-cell epitope identification using yeast displaying pathogen-derived peptide library. A library of DNA fragments that encode all the possible peptides with 10–20 amino acids from the antigens (single antigenic proteins or pathogenic organisms) are fused to the gene encoding the restriction single-chain MHC class II molecule in a yeast display vector. The resultant library of recombinant yeast cells are analyzed by FACS to identify those containing peptides with high affinity towards the restriction MHC molecule, which are subsequently screened for their ability to induce antigen-specific T-cell activation. DNA sequence analysis of selected positive clones results in direct identification of the antigenic peptides. We show that this method can be used to rapidly pinpoint the $HA_{306–322}$ epitope from the haemagglutinin protein and the entire influenza virus X31/A/Aichi/68 genome, respectively.

1. Introduction

Class II major histocompatibility complex (MHCII) proteins bind processed antigenic peptides and present them to CD4+ T-cells through engagement of T-cell receptors (TCRs). Identifying these specific antigenic peptides, T-cell epitopes, is a critical but often slow and difficult step in studying the immune responses to tumors, infectious agents, and autoantigens, as well as in developing pertinent vaccines and therapeutic strategies (Rosenberg, 2001; Kwok and Nepom, 2003; Stone et al., 2005). Thanks to the introduction of T-cell cloning and propagation techniques, many antigen-specific T-cell clones have been identified, but their fine antigenic specificity within the antigens remain undefined.

A variety of computational and experimental methods have been developed to identify T-cell epitopes. When the potential antigenic protein(s) is known, T-cell epitopes are usually identified by testing a panel of overlapping peptides spanning the antigenic protein sequence (Gammon et al., 1991; Maecker et al., 2001; Novak et al., 2001). However, this method becomes impractical when the overall size of the potential antigenic proteins increases (e.g. viruses with large genome size), and is not applicable when the antigen is unknown. To address these issues, several expression cloning approaches have been introduced to identify the source protein first, so that the epitope could be identified using the conventional method described above. However, this strategy has met limited success with MHCII restricted antigens, of which the processing pathway favors exogenous proteins, and it requires either extensive engineering of a cell line for efficient MHCII antigen processing and presentation (Wang et al., 1999) or labor-intensive purification of the antigenic proteins (Koelle, 2003). As an alternative approach to this problem, random synthetic peptide library has been constructed, genetically or chemically, and screened using various methods, such as positional scanning synthetic combinatorial libraries (PS-SCLs) (Sospedra et al., 2003), bead-bound libraries (Hiemstra et al., 2000), and peptide libraries randomized at the TCR contacting residues (Boen et al., 2000; Crawford et al., 2006). Unfortunately, few of these methods were

* Corresponding author. Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States. Tel.: +1 217 333 2631; fax: +1 217 333 5052.
E-mail address: zhao5@uiuc.edu (H. Zhao).

0022-1759/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jim.2008.03.008

able to identify native epitopes (Nino-Vasquez et al., 2004; Wang et al., 2005), due to the "random" nature of the peptide library. In addition, most of these expression cloning and combinational peptide library-based approaches rely on the use of autologous antigen presenting cells (APCs) that are often difficult to isolate, expand, and maintain.

To address these limitations, we have developed a new CD4+ T-cell epitope identification method by using yeast display coupled with fluorescence activated cell sorting (FACS) and T-cell activation assay. Cell surface display has been widely used to engineer enzymes, antibodies, and TCRs (Georgiou et al., 1997; Wittrup, 2001; Sergeeva et al., 2006). However, this tool has not been used for T-cell epitope identification, mainly because antigenic peptides have to be expressed in the context of appropriate MHC proteins for recognition by TCRs, and most of the display platforms, such as *E. coli* and phage, are not amenable to heterologous expression of these complex dimeric eukaryotic proteins. The latter problem was partially solved by the introduction of eukaryotic display systems, such as yeast (Boder and Wittrup, 1997) and baculovirus (Ernst et al., 1998). We sought to use yeast display because it is a more robust, facile, and much less complex system than other eukaryotic systems. In this study, we used human MHCII protein, DR1, as a model to demonstrate the utility of the new CD4+ T-cell epitope identification method. A 17-residue epitope, $HA_{306–322}$, was rapidly pinpointed from the haemagglutinin protein, and the entire influenza virus X31/A/Aichi/68 genome, respectively.

2. Materials and methods

The sequences of all the PCR primers were listed in Supplemental Table S1 online and were synthesized by Integrated DNA Technologies (Coralville, IA). All restriction enzymes were obtained from New England Biolabs (Ipswich, MA). Unless otherwise indicated, all chemicals were purchased from Sigma (St. Louis, MO).

2.1. Vector construction

DNA encoding a single-chain polypeptide consisting of the α-chain of DR1, the peptide $HA_{306–318}$, and the β-chain of DR1 (Fig. 1b) was prepared by PCR amplification using primer pairs α-5BX and β-3XH and plasmid pJ3/238 (a gift from M. Mage, NIH, Bethesda, MD) as a template. The amplification product was cloned into pYD1 (Invitrogen) via BstXI and XhoI to generate pYD1scαHAβ.To facilitate plasmid construction with different peptides, a NotI restriction site was introduced by splicing overlapping extension PCR (SOE-PCR) (Horton et al., 1990) to directly precede $HA_{306–318}$ using primer pairs pYD1For/NotIRev and NotIFor/βRev73-67. The SOE-PCR product was cotransformed with the BstXI/SpeI fragment of pYD1scαHAβ into EBY100 strain (Invitrogen) to generate pYD1αHAβ. To eliminate the possibility that the epitope $HA_{306–318}$ identified from the peptide library is the undigested pYD1αHAβ, plasmid pYD1αSTFβ was constructed. The stuffer DNA (STF) is an unrelated gene (~1 kb) amplified from the phosphite dehydrogenase gene (Woodyer et al., 2003) using primers StfFor and StfRev. The PCR product was cloned into pYD1 via BstXI and XhoI to generate plasmid pYD1αSTFβ. Plasmids pYD1α2467β, pYD1αM4β, pYD1αYAKβ, pYD1αCIIβ, pYD1αPKAβ and pYD1αLβ were generated in two steps. First, the double-stranded DNA insert encoding a peptide was obtained by hybridization of the sense and anti-sense oligonucleotides (see Supplemental Table S2 online) encoding the peptide. Second, the resulting DNA insert was phosphorylated by Optikinase (USB, Cleveland, OH), and ligated into the same restricted pYD1αSTFβ. pYD1αMBPβ was generated in a slightly different way: the insert was obtained by annealing, extension and digestion of a self-annealing primer MBPNotI. All plasmid constructs were transformed into DH5α, recovered, and confirmed by DNA sequencing.

2.2. Yeast surface display and flow cytometric analysis

S. cerevisiae EBY100 clones transformed with different plasmid constructs were cultured and analyzed as described (Esteban and Zhao, 2004).

2.3. Library construction

Genomic RNA purified from influenza virus X31/A/Aichi/68 (Charles River Laboratories, Wilmington, MA) was used as a template for cDNA synthesis. First strand cDNA was synthesized using a first stand cDNA synthesis kit following the manufacturer's protocol (Roche, Indianapolis, IN). DNA polymerase I (5 U/20 μL reaction) and RNase H (0.16 U/20 μL reaction) were then used to synthesize double-stranded cDNA. The cDNA encoding the antigenic protein/proteins was fragmented using DNase I (0.075 U/μg DNA for 1 min) as described elsewhere (Zhao and Arnold, 1997). DNA fragments with size between 30 and 60 bp were collected and purified from agarose gel, blunt-end polished as described (Margulies et al., 2005), and ligated into SpeI and NotI restricted pYD1αSTFβ vector that was similarly blunt-end polished. The ligation mixture was transformed into ElectroMax DH5α competent cells (Invitrogen, Carlsbad, CA) and a library of 1.9×10^6 clones was obtained. Plasmids were then recovered and transformed into EBY100 and a library of 6.6×10^6 clones was generated. After two passages in selection media, the library was induced for protein expression and subjected to FACS. The haemagglutinin-derived peptide library was constructed similarly except that DNA encoding haemagglutinin was obtained by one-step RT-PCR using influenza A/Aichi/2/68 viral RNA (a kind gift from R. Donis at CDC, Atlanta, GA) as a template and AichiFor and AichiRev as primers.

2.4. FACS

The library of yeast clones displaying DR1 in complex with different peptides derived from the influenza A virus genome was enriched through two cycles of FACS. An aliquot of induced library culture containing 10^8 yeast cells was stained with LB3.1 followed by treatment with biotin-conjugated goat-anti-mouse (GAM) IgG and streptavidin-phycoerytrin (SA-PE) as described (Esteban and Zhao, 2004) and sorted by a Coulter 753 bench FACS at the Biotechnology Center of the University of Illinois (Urbana, IL). About 1.5% of the population with the highest fluorescence was collected and recovered in the first round of cell sorting. In the second round of cell sorting, cells from the top 0.5% of the population were sorted into four 96-well plates and analyzed individually for their

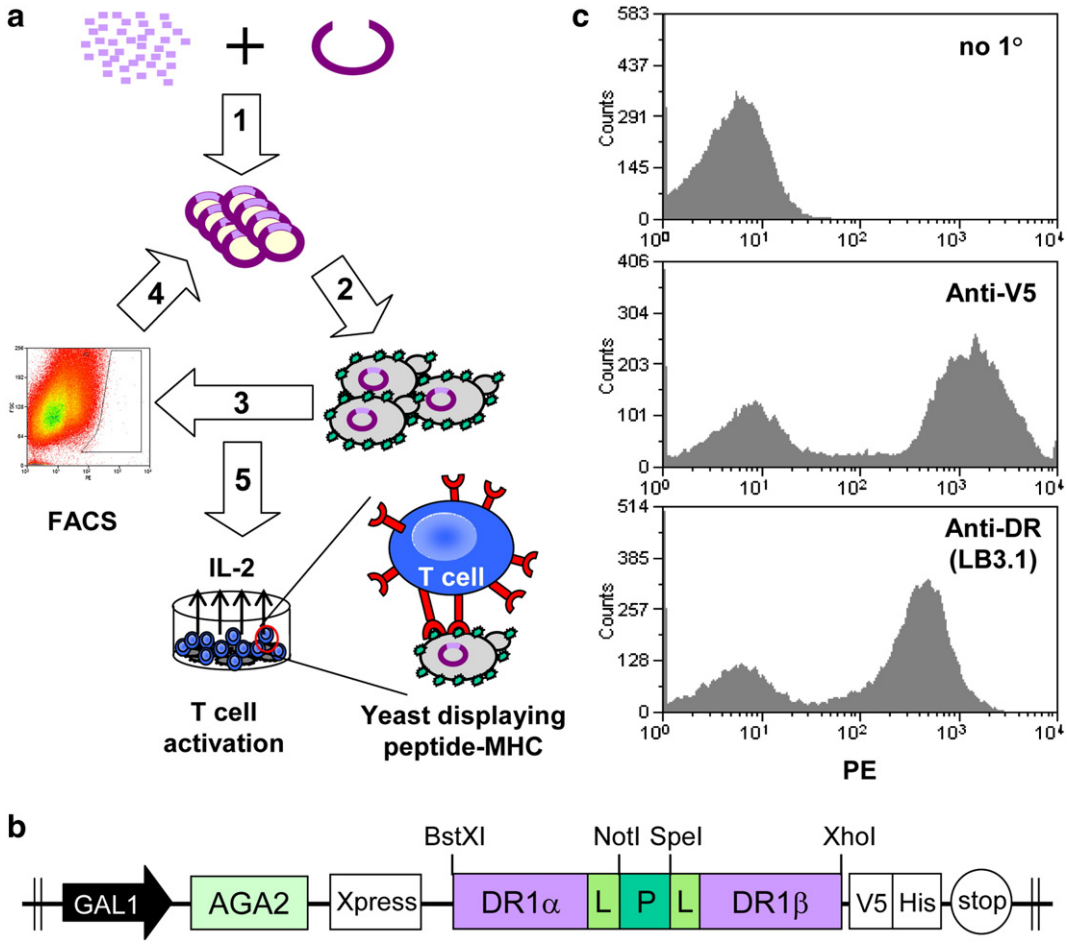

Fig. 1. Design of a novel CD4+ T-cell epitope identification method. (a) Fragments (~30–60 bp) generated from antigenic DNA material are fused to the single-chain MHCII molecules in a yeast display vector (step 1). Following transformation and gene expression (step 2), yeast cells displaying the entire library of peptide–MHC complexes are analyzed by FACS (step 3) to identify a small subset of yeast cells containing peptides with high affinity towards the restriction MHCII molecule. Plasmids are then recovered from these yeast cells (step 4) and analyzed for further enrichment. This enrichment cycle is repeated for ~2–4 rounds. Individual clones from the enriched library are then screened for their ability to induce antigen-specific T-cell activation and produce IL-2 using T-cell hybridomas transfected with the TCR of interest (step 5). (b) Schematic representation of the single-chain HLA–DR1–peptide construct used in yeast display. Xpress = Xpress epitope, L = linker, P = peptide, V5 = V5 epitope. (c) Flow cytometric analysis of single-chain HLA–DR1–$HA_{306–318}$ expression on the yeast surface. Yeast cells transformed with pYD1αHAβ were stained with anti-V5 antibody to detect full-length protein expression or with LB3.1 to detect correctly folded peptide–DR1 complexes. Yeast cells stained with secondary antibody only (no 1°) were used as a control.

ability to stimulate T-cell hybridoma. The haemagglutinin-derived peptide library was analyzed similarly.

2.5. Stimulation of HA1.7 T-cell hybridomas and IL-2 detection

The HA1.7 hybridoma T-cells (Boen et al., 2000) were used as indicator cell lines for productive MHC–peptide–TCR interactions that produce IL-2 upon receptor engagement. They were maintained in IMDM (Invitrogen, Carlsbad, CA) with 10% fetal bovine serum (FBS) in a 5% CO_2 environment at 37 °C. About 10^6 yeast cells in PBS were loaded into each well of a 96-well tissue-culture plate. After incubation at 4 °C for at least 24 h, unattached yeast cells were washed off with PBS. HA1.7 hybridoma cells were added to wells (10^5 cells/well) and incubated for ~24 h at 37 °C with 5% CO_2. The supernatant was tested for IL-2 production using a murine IL-2 ELISA kit (eBiosciences, San Diego, CA).

3. Results

3.1. Design of a new CD4+ T-cell epitope identification method

To overcome some of the above-mentioned issues presented by existing CD4+ T-cell epitope identification methods, we have developed a new approach that is both technically simple and conceptually novel. As shown in Fig. 1a, DNA encoding the pathogenic protein(s) obtained by PCR/RT-PCR using sequence-specific primers and random hexamer primers for known and unknown antigens, respectively, is randomly digested into fragments using DNase I. A library of DNA fragments ranging from 30 to 60 nucleotides that encode all the possible peptides with 10–20 amino acids from the antigens (single antigenic proteins or pathogenic organisms such as parasites, tumors, and viruses) is purified and fused to the gene encoding the

restriction single-chain MHCII molecule in a yeast display vector. Following gene expression, yeast cells displaying the entire library of peptide–MHC complexes are analyzed by FACS to identify a small subset of yeast cells containing peptides with high affinity towards the restriction MHC molecule according to their surface expression level of the peptide–MHC complexes. This small subset of yeast cells are then screened for their ability to induce antigen-specific T-cell activation using T-cell hybridomas transfected with the TCR of interest. DNA sequence analysis of selected active clones results in direct identification of the antigenic peptides. If necessary, a series of DNA fragments encoding overlapping peptide sequences can be used to refine those antigenic peptide sequences.

The success of this approach hinges on the development of recombinant yeast cells that are capable of displaying peptide–MHCII complexes, discriminating peptide binding affinities, and activating T-cells of interest. To demonstrate the concept, we sought to use an important human MHCII protein, DR1, and its known T-cell epitope, $HA_{306–318}$ as a model system.

3.2. Functional expression of wild-type single-chain HLA–DR1–HA on yeast cell surface

Based on our previous attempt (Esteban and Zhao, 2004), we constructed a single polypeptide chain, scDR1αHAβ, in which peptide $HA_{306–318}$ was covalently attached to the C-terminus of the α-chain of DR1 and the N-terminus of the β-chain of DR1 through two flexible linkers, and fused it to the AGA2 protein in the yeast pYD1 display vector (Fig. 1b). Analysis of the V5 epitope expression by flow cytometry allowed the estimation of the expression level of the full-length protein on the yeast cell surface. A positive staining yeast cell population was detected with anti-V5 antibody (Fig. 1c) and the relative mean fluorescence units (MFU) of ~100 indicated a high surface expression level of the wild-type scDR1αHAβ protein. Expression of properly folded wild-type scDR1αHAβ protein was also detected using a conformational antibody, LB3.1, that can only recognize correctly assembled DR heterodimers (Stern and Wiley, 1992) (Fig. 1c). Taken together, these data suggested that the wild-type scDR1αHAβ protein can be displayed on yeast surface in their native state. Note that wild-type human class II DR4 molecules with a covalently attached epitope have also been displayed on yeast surface in their native state, but they are expressed as a noncovalently associated heterodimer and are incapable of stimulating antigen-specific T-cells (Boder et al., 2005). Similarly, a single-chain murine class II MHC protein mutant without a covalently attached peptide was displayed on the yeast surface in its native state, but no peptide-binding or T-cell activation studies were reported (Starwalt et al., 2003).

3.3. A positive correlation between surface expression level and peptide binding affinity

To evaluate whether the recombinant yeast could discriminate peptide binding affinity towards DR1, we designed a series of scDR1αpepβ constructs containing peptides with different affinity towards DR1 molecule (Figs. 1b and 2a, inset). Peptide 2467 was designed to have optimal residues at anchor positions P2, P4, P6 and P7 (Hammer et al., 1992) and was expected to have the highest affinity. Peptide M4 was chosen because it has a higher binding affinity than $HA_{306–318}$ (Hammer et al., 1992), but its affinity should be lower than that of 2467 since it does not have the optimal anchor residue combination. Peptides YAK and PKA have been reported to have lower affinity towards DR1 than $HA_{306–318}$ (Sato et al., 2000). Peptide $CII_{261–273}$ has been reported to have at least 10 times lower binding affinity than $HA_{306–318}$ (Rosloniec et al., 2002). A negative control, a peptide linker $(G_4S)_2G_3$, was also included. These constructs were analyzed for their protein surface expression levels by flow cytometry. As shown in Fig. 2a, the protein surface expression level of the peptide-MHC complex increases with increasing affinity of the covalently attached peptide, indicating a positive correlation between the peptide binding affinity and the protein surface expression level. Therefore, the yeast surface expression level of the scDR1αpepβ protein, as measured by DR-specific antibody (LB3.1) staining, may be used as a proxy screening variable for high affinity peptides that are candidate CD4+ T-cell epitopes. By means of the new expression cloning strategy described above, a library of scDR1αpepβ proteins containing up to 10^9 different pathogen-derived peptides may be generated within days and examined individually for protein surface expression level by FACS within hours. Single yeast cells of interest can also be readily sorted into 96-well plates for further analysis. However, to identify the antigenic CD4+ T-cell epitopes from the resulting pool of binding peptides, the corresponding scDR1αpepβ proteins need to be evaluated for their ability to induce T-cell activation, which unfortunately requires the use of professional APCs as in many of the current epitope identification methods. To overcome this key limitation associated with most CD4+ T-cell epitope discovery methods, we decided to examine whether the engineered yeast could support direct activation of T-cells.

3.4. Recombinant yeast activate T hybridoma cells in a peptide-specific manner

In the T-cell activation assay, we used HA1.7 hybridoma cells that are specific for DR1–$HA_{306–318}$ or DR4–$HA_{306–318}$, and were developed to obviate the need to establish and maintain functional T-cell clones (Boen et al., 2000). Similar to the previous report (Boder et al., 2005), no IL-2 secretion was detected when HA1.7 hybridoma cells were stimulated with yeast displaying the DR1–$HA_{306–318}$ complexes in suspension (data not shown). This inability to induce T-cell activation may be caused by the lack of costimulatory molecules on the yeast surface. Without the aid of costimulatory molecules, TCR signaling is weak. In addition, the spatially restricted MHC–peptide complexes displayed on yeast surface may not reach the density threshold that is important for stable interaction with TCRs (Grakoui et al., 1999). Therefore, we postulated that immobilization of yeast cells on plastic surface might increase the relative density of MHC–peptide complexes available for TCR interaction and consequently enable T-cell activation.

To test this hypothesis, yeast cells displaying various scDR1αpepβ constructs were immobilized in the wells of a 96-well plate and incubated with HA1.7 hybridoma cells. As shown in Fig. 2b, IL-2 secretion was detected for yeast cells displaying the DR1–$HA_{306–318}$ complexes. More importantly, the hybridoma cells only responded to yeast cells displaying the DR1–$HA_{306–318}$

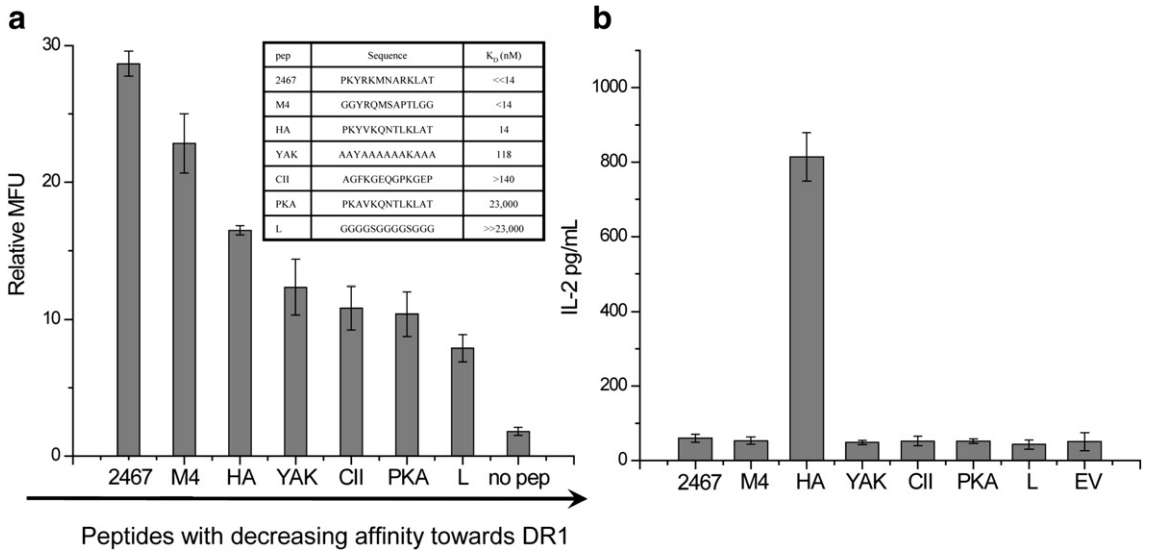

Fig. 2. Analysis of yeast displaying peptide–DR1 complexes. (a) A positive correlation between yeast surface display level and peptide binding affinity (inset) towards HLA–DR1 molecule was observed. Peptides on the *x*-axis were ranked in a decreasing order of their binding affinity towards DR1. The expression level of scDR1αβ that has no peptide in the linker between α chain and β chain was reported previously (Esteban and Zhao, 2004). MFU = mean fluorescence unit. (b) Yeast displaying HLA–DR1-peptide acted as artificial APCs and activated HA1.7 hybridoma in an antigen-specific manner. Yeast cells containing empty plasmid pYD1 (EV) were cultured, induced, and analyzed in the same way as yeast displaying other constructs.

complexes, but not to those displaying the scDR1αpepβ constructs harboring other peptides. These data clearly suggested that yeast displaying scDR1αpepβ proteins, when immobilized, can activate hybridoma T-cells in a peptide specific manner, and thus could be used directly in the activity screening.

3.5. Epitope identification from a single antigenic protein

To demonstrate the utility of our method, we first used a known antigen, influenza A virus haemagglutinin containing the CD4+ T-cell epitope, $HA_{306–318}$. To create a library of peptides containing all the possible peptides derived from haemagglutinin, the DNA encoding influenza A virus haemagglutinin was obtained by RT-PCR and randomly digested by DNase I. The DNA fragments with size ranging from 30 bp to 60 bp (corresponding to 10 to 20 amino acids) were purified from the agarose gel after electrophoresis, and cloned into the yeast display vector encoding HLA-DR1 protein using blunt-end ligation, resulting in a library of $\sim 4.6\times10^5$ clones. Note that since the haemaglutinin gene is ~1.7 kb in size, the library created here has ~100% chance of covering all possible ~20-residue-peptides based on a reported algorithm (Patrick et al., 2003). Yeast displaying scDR1αpepβ proteins harboring high affinity binding peptides were enriched by FACS (Fig. 3a). After four rounds of FACS enrichment, ten clones were randomly picked and evaluated for their protein surface expression levels and T-cell activation ability. All ten clones showed higher or comparable expression levels as the yeast displaying the DR1–$HA_{306–318}$ complexes (data not shown) and five of them (clone 3, 4, 5, 7, and 9), representing 50% of the enriched library, were capable of stimulating HA1.7 hybridoma T-cells (Fig. 3b). DNA sequencing analysis showed that all these five clones had an identical peptide PKYVKQNTLKLATGMRNVPEKQT, which contained the $HA_{306–318}$ epitope (underlined).

To probe the reason why the other five peptides were also enriched in the library, the peptide-encoding region was sequenced and the peptide sequence was deduced and analyzed by TEPITOPE (Bian et al., 2003). Four out of five peptides contained predicted DR1-binding peptides (see Supplemental Table S3 online), indicating that 90% of the enriched peptides are, in theory, capable of binding HLA–DR1. This result suggested that our proposed method can be used efficiently to enrich MHC-binding peptides and identify T-cell epitopes from a single known antigenic protein.

3.6. Epitope identification from a pathogenic organism

To further evaluate whether our method can be used to directly identify CD4+ T-cell epitopes from pathogenic organisms such as viruses and tumor cells, we used the influenza A virus as a model antigen. The genome of influenza A virus consists of eight separate segments of negative-sense single-stranded RNA (Lamb and Choppin, 1983), thus cDNA has to be synthesized to create a peptide library. Although the genome sequence of the influenza virus strain used in this research (X31/A/Aichi/68) is available, usage of sequence information is avoided purposely to demonstrate that our method is generally applicable, especially when studying a newly emerging virus or an unknown tumor cell. Unfortunately, conventional cDNA synthesis method using poly-dT oligo is not applicable in the case of influenza viruses, since there is no poly-dA tail in the viral genome (Lamb and Choppin, 1983). Therefore, random hexamer primer was used to synthesize the first strand cDNA. Following RNaseH and DNA polymerase I treatment, double-stranded cDNA was obtained. A library of $\sim 1.9\times10^6$ yeast clones displaying scDR1αpepβ complexes, which should contain all possible 20-residue-peptides derived from influenza A virus based on the reported algorithm (Patrick et al., 2003), was

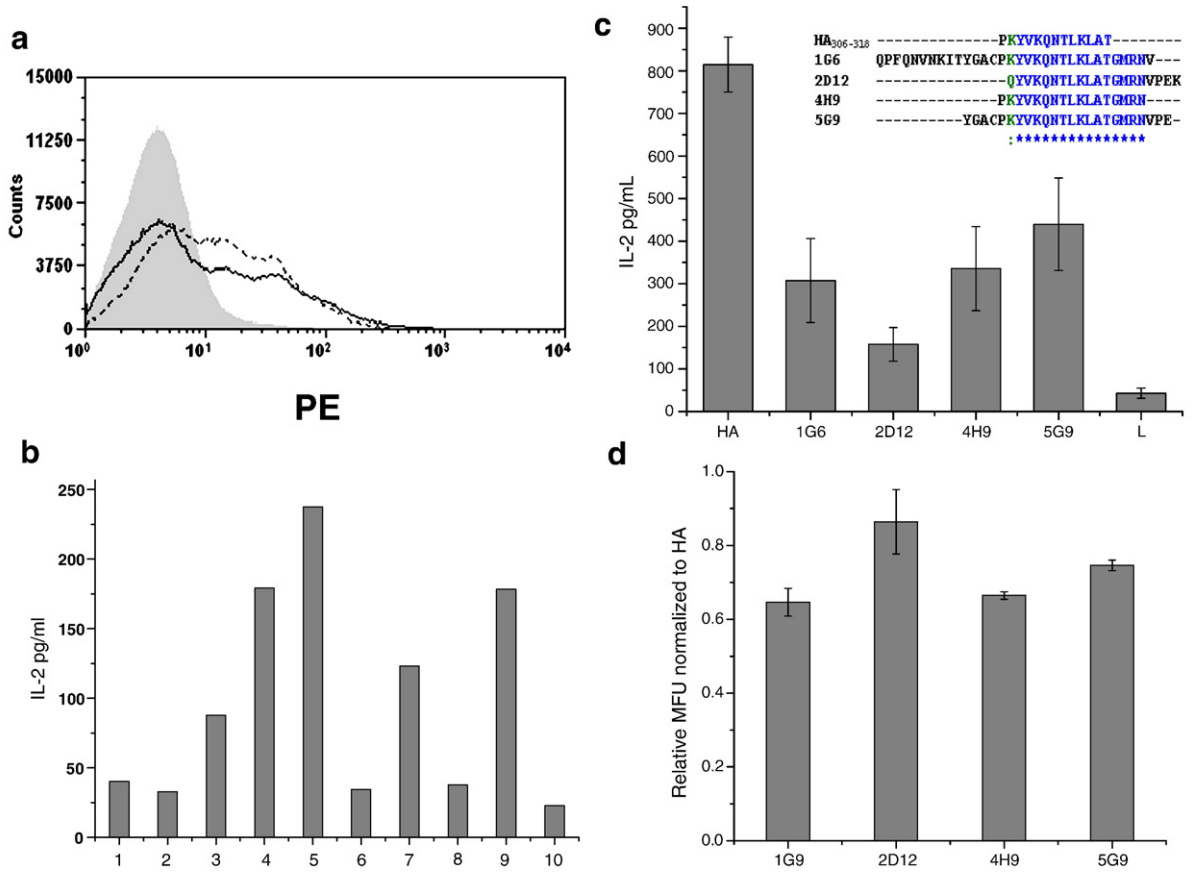

Fig. 3. Identification of $HA_{306–318}$ epitope from peptide libraries. (a) FACS enrichment of potential good binders. LB3.1, a DR-specific conformation-sensitive antibody, was used to stain cells as a measurement of surface expression level of correctly folded peptide-DR1 complexes. Profiles of haemagglutinin-derived peptide library before enrichment and after three rounds of enrichment, and influenza A virus genome-derived peptide library after two rounds of enrichment are shown in shade, solid line, and dashed line, respectively. (b) T-cell activation analysis was performed using the ten randomly picked clones and five of them (clone 3,4,5,7, and 9) activated T hybridoma cell, HA1.7. (c) T-cell activation analysis of the four clones identified from the influenza A virus genome-derived peptide library. The peptide sequences from the four clones were aligned with the $HA_{306–318}$ epitope sequence (inset). * = fully conserved residues,: = conservation of strong groups. (d) The four active clones selected from the influenza A virus genome-derived peptide library showed comparable surface expression level as yeast displaying scDR1αHAβ. The relative MFU was normalized to that of HA for direct comparison.

created using the same method described above. To probe the diversity of the peptides in the primary library, twenty clones were randomly picked, sequenced, and mapped to the viral genome. There was at least one peptide derived from each of the eight pieces of viral genomic RNA and displayed in either sense or anti-sense orientations (data not shown), indicating a good representation of the whole viral genome.

After two rounds of FACS enrichment, four plates of yeast clones (with no more than one cell per well) from the top 0.5% population were collected and screened for their ability to activate HA1.7 hybridoma T-cells. Four clones (1G6, 2D12, 4H9, 5G9), representing ~1% of the enriched library, activated T-cells (Fig. 3c). Compared to yeast displaying scDR1αHAβ, the four identified clones showed comparable surface expression level (Fig. 3d) and reduced antigenicity by ~2–3 fold. This might be due to the reduced binding capacity engendered by the extra amino acids at the N- and/or C-terminus (O'Sullivan et al., 1991). Sequence alignment of the four peptides revealed a T-cell epitope, PKYVKQNTLKLATGMRN with the $HA_{306–318}$ epitope sequence underlined, and a T-cell mimotope QYVKQNTLKLATGMRN that differs from the epitope by two residues: proline at the P-2 position is absent and a substitution of lysine with glutamine at the P-1 position (Fig. 3c, inset). The presence of the mimotope, not unexpectedly, was a result of an open reading frame fusion at the junction of the vector-encoded and virus-encoded DNA regions as described (Koelle, 2003). Two more rounds of FACS screening further enriched the epitope occurrence frequency to ~6%. This result clearly suggested that our method was able to rapidly pinpoint a 17-residue CD4+ T-cell epitope from the entire influenza A virus genome, and it might be used to identify T-cell epitopes from an unknown antigen.

4. Discussion

We have described a simple and efficient approach that can be used to rapidly identify CD4+ T-cell epitopes from both known and unknown antigens by using yeast displaying pathogen-derived peptide library. By means of a new expression cloning method, a pathogen-derived peptide library was displayed on yeast surface in the context of wild-type single-

chain DR1. The ability of the recombinant yeast cells to distinguish the peptide binding affinity towards DR1 allowed efficient enrichment of DR1-binding peptides using a DR-specific antibody, LB3.1, and the epitope occurrence frequency was enriched to ~50% and ~6% of the enriched library for the haemagglutinin- and influenza viral genome-derived peptide library, respectively. In the subsequent T-cell activity screening assay, the recombinant yeast cells could be used directly to activate T hybridoma cells in a peptide specific manner, enabling the identification of the epitope. This study represents the first demonstration that yeast surface peptide–MHCII display can be used to identify T-cell epitopes.

Compared to other existing methods (Kwok and Nepom, 2003), our approach is faster, simpler, less reagent-intensive, and offers the following advantages. First, unlike other peptide libraries, the expression cloning peptide library created in this study consists, in theory, all possible peptides derived from a pathogen. Thus it allows direct identification of T-cell epitopes, rather than mimotopes that are less biological relevant to the natural TCRs. Second, the fragmentation of pathogenic protein(s) is carried out at DNA level in vitro and these peptides are genetically fused to yeast surface protein, therefore, no complex MHCII peptide processing and presentation pathways are required for artificial antigen presentation, which is the major obstacle for conventional expression cloning libraries (Wang et al., 1999; Koelle, 2003). Third, the occurring frequency of the epitope is greatly increased after enrichment by FACS, only a small number of yeast cells (100–200) need to be examined for their ability to activate T-cells, obviating the need for brute-force screening. Finally, the recombinant yeast cells can be used directly in the T-cell activity assay, thus obviating the use of autologous APCs.

To achieve rapid and accurate identification of antigenic CD4+ T-cell epitopes, the screening method described in this study requires *two-step*: a high-throughput FACS based screening and a low throughput T-cell activation assay. This is mainly because the peptides are covalently attached to MHCII proteins through flexible linkers, which might enable binding of MHCII molecules with peptides that normally do not bind or bind weakly. These "false–positive" binders can be eliminated in the secondary screening step using T-cells that bear TCRs specific for the antigenic peptide. However, this method on its own is not sufficient for more complex situations, such as epitope identification from unknown antigens using T-cell lines, of which the antigen specificity is not well defined. In this case, peptides that do not normally activate T-cells might also be selected from the library after two steps of screening, because the local concentration of covalently attached peptides in the single-chain construct is much higher than that found under normal physiological conditions. To address this issue, further examination of the isolated peptides should be conducted by using conventional T-cell function analysis, which involves testing chemically synthesized peptides and using professional APCs.

Although the CD4+ T-cell epitope identification method described above was only demonstrated with DR1 allele, the generality of this system was partially demonstrated by the functional expression of another human MHCII allele, DR4, on yeast cell surface that also activated HA1.7 hybridoma T-cells in a peptide specific manner (see Supplemental Fig. S1 online). Nonetheless, since MHC molecules are highly polygenic and polymorphic, it might be difficult to functionally display every target MHCII allele on yeast surface. Fortunately, this limitation may be overcome by introduction of a few point mutations via directed evolution into the target MHC protein (Starwalt et al., 2003; Esteban and Zhao, 2004). However, there may be concerns about the effect of the introduced mutations on the biological activity of MHC molecules. Recently, we have analyzed one of the DR1 mutants we identified earlier, which has a leucine to histidine mutation at position 11 of β-chain (Esteban and Zhao, 2004). And it was found that this mutant actually elicited a much stronger T-cell response than the wild-type DR1 did (data not shown). This is probably a result of the increased surface expression level of the mutant, which increased the ligand density for TCR recognition. Therefore, it is possible that mutations only improve the solubility of MHC proteins, while leave their biological function intact.

Finally, although this method was demonstrated by identifying a viral epitope, it should be generally applicable to identification of T-cell epitopes from other systems such as cancers and autoimmune diseases. For such applications, the sensitivity of the yeast acting as "artificial" APCs needs to be improved so that they could be used to stimulate T-cell lines or primary T-cells, which are much more fastidious in their activation requirements than T hybridoma cells. Currently, we are trying to include the costimulatory molecules and accessory molecules in yeast, which have been shown to be important for activation of primary T-cells or T-cell lines (Kim et al., 2004). It should be noted that this yeast display system is amenable to engineering by powerful directed evolution approaches. As such, it provides a technology platform for engineering the displayed MHC molecules for the development of various MHC-based reagents for basic and applied biomedical research.

Acknowledgements

We thank M. Mage, J. Bill, R. Donis, and J. Hammer for kindly providing the pJ3/238 plasmid, HA1.7 hybridoma cells, influenza A/Aichi/2/68 viral RNA, and the TEPITOPE program, respectively; B. Pilas and B. Montez at the Biotechnology Center of the University of Illinois for flow cytometry and FACS assistance. This work was partly supported by the Department of Chemical and Biomolecular Engineering of University of Illinois at Urbana-Champaign and National Science Foundation CAREER Award (BES-0348107).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jim.2008.03.008.

References

Bian, H., Reidhaar-Olson, J.F., Hammer, J., 2003. The use of bioinformatics for identifying class II-restricted T-cell epitopes. Methods 29, 299.

Boder, E.T., Wittrup, K.D., 1997. Yeast surface display for screening combinatorial polypeptide libraries. Nat. Biotechnol. 15, 553.

Boder, E.T., Bill, J.R., Nields, A.W., Marrack, P.C., Kappler, J.W., 2005. Yeast surface display of a noncovalent MHC class II heterodimer complexed with antigenic peptide. Biotechnol. Bioeng. 92, 485.

Boen, E., Crownover, A.R., McIlhaney, M., Korman, A.J., Bill, J., 2000. Identification of T cell ligands in a library of peptides covalently attached to HLA-DR4. J. Immunol. 165, 2040.

Crawford, F., Jordan, K.R., Stadinski, B., Wang, Y., Huseby, E., Marrack, P., Slansky, J.E., Kappler, J.W., 2006. Use of baculovirus MHC/peptide display libraries to characterize T-cell receptor ligands. Immunol. Rev. 210, 156.

Ernst, W., Grabherr, R., Wegner, D., Borth, N., Grassauer, A., Katinger, H., 1998. Baculovirus surface display: construction and screening of a eukaryotic epitope library. Nucleic Acids Res. 26, 1718.

Esteban, O., Zhao, H., 2004. Directed evolution of soluble single-chain human class II MHC molecules. J. Mol. Biol. 340, 81.

Gammon, G., Geysen, H.M., Apple, R.J., Pickett, E., Palmer, M., Ametani, A., Sercarz, E.E., 1991. T cell determinant structure: cores and determinant envelopes in three mouse major histocompatibility complex haplotypes. J. Exp. Med. 173, 609.

Georgiou, G., Stathopoulos, C., Daugherty, P.S., Nayak, A.R., Iverson, B.L., Curtiss III, R., 1997. Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. Nat. Biotechnol. 15, 29.

Grakoui, A., Bromley, S.K., Sumen, C., Davis, M.M., Shaw, A.S., Allen, P.M., Dustin, M.L., 1999. The immunological synapse: a molecular machine controlling T cell activation. Science 285, 221.

Hammer, J., Takacs, B., Sinigaglia, F., 1992. Identification of a motif for HLA-DR1 binding peptides using M13 display libraries. J. Exp. Med. 176, 1007.

Hiemstra, H.S., Drijfhout, J.W., Roep, B.O., 2000. Antigen arrays in T cell immunology. Curr. Opin. Immunol. 12, 80.

Horton, R.M., Cai, Z.L., Ho, S.N., Pease, L.R., 1990. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques 8, 528.

Kim, J.V., Latouche, J.B., Riviere, I., Sadelain, M., 2004. The ABCs of artificial antigen presentation. Nat. Biotechnol. 22, 403.

Koelle, D.M., 2003. Expression cloning for the discovery of viral antigens and epitopes recognized by T cells. Methods 29, 213.

Kwok, W.W., Nepom, G.T., 2003. T-cell epitope mapping. Methods 29, 211–212.

Lamb, R.A., Choppin, P.W., 1983. The gene structure and replication of influenza virus. Annu. Rev. Biochem. 52, 467.

Maecker, H.T., Dunn, H.S., Suni, M.A., Khatamzas, E., Pitcher, C.J., Bunde, T., Persaud, N., Trigona, W., Fu, T.M., Sinclair, E., Bredt, B.M., McCune, J.M., Maino, V.C., Kern, F., Picker, L.J., 2001. Use of overlapping peptide mixtures as antigens for cytokine flow cytometry. J. Immunol. Methods 255, 27.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376.

Nino-Vasquez, J.J., Allicotti, G., Borras, E., Wilson, D.B., 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.

Novak, E.J., Liu, A.W., Gebe, J.A., Falk, B.A., Nepom, G.T., Koelle, D.M., Kwok, W.W., 2001. Tetramer-guided epitope mapping: rapid identification and characterization of immunodominant CD4+ T cell epitopes from complex antigens. J. Immunol. 166, 6665.

O'Sullivan, D., Sidney, J., Del Guercio, M.F., Colon, S.M., Sette, A., 1991. Truncation analysis of several DR binding epitopes. J. Immunol. 146, 1240.

Patrick, W.M., Firth, A.E., Blackburn, J.M., 2003. User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. Protein Eng. 16, 451.

Rosenberg, S.A., 2001. Progress in human tumour immunology and immunotherapy. Nature 411, 380.

Rosloniec, E.F., Whittington, K.B., Zaller, D.M., Kang, A.H., 2002. HLA–DR1 (DRB1*0101) and DR4 (DRB1*0401) use the same anchor residues for binding an immunodominant peptide derived from human type II collagen. J. Immunol. 168, 253.

Sato, A.K., Zarutskie, J.A., Rushe, M.M., Lomakin, A., Natarajan, S.K., Sadegh-Nasseri, S., Benedek, G.B., Stern, L.J., 2000. Determinants of the peptide-induced conformational change in the human class II major histocompatibility complex protein HLA–DR1. J. Biol. Chem. 275, 2165.

Sergeeva, A., Kolonin, M.G., Molldrem, J.J., Pasqualini, R., Arap, W., 2006. Display technologies: application for the discovery of drug and gene delivery agents. Adv Drug Deliv Rev 58, 1622.

Sospedra, M., Pinilla, C., Martin, R., 2003. Use of combinatorial peptide libraries for T-cell epitope mapping. Methods 29, 236.

Starwalt, S.E., Masteller, E.L., Bluestone, J.A., Kranz, D.M., 2003. Directed evolution of a single-chain class II MHC product by yeast display. Protein Eng. 16, 147.

Stern, L.J., Wiley, D.C., 1992. The human class II MHC protein HLA–DR1 assembles as empty alpha beta heterodimers in the absence of antigenic peptide. Cell 68, 465.

Stone, J.D., Demkowicz, W.E., Stern, L.J., 2005. HLA-restricted epitope identification and detection of functional T cell responses by using MHC–peptide and costimulatory microarrays. Proc. Natl. Acad. Sci. U. S. A. 102, 3744.

Wang, R.F., Wang, X., Atwood, A.C., Topalian, S.L., Rosenberg, S.A., 1999. Cloning genes encoding MHC class II-restricted antigens: mutated CDC27 as a tumor antigen. Science 284, 1351.

Wang, Y., Rubtsov, A., Heiser, R., White, J., Crawford, F., Marrack, P., Kappler, J.W., 2005. Using a baculovirus display library to identify MHC class I mimotopes. Proc. Natl. Acad. Sci. U. S. A. 102, 2476.

Wittrup, K.D., 2001. Protein engineering by cell-surface display. Curr. Opin. Biotechnol. 12, 395.

Woodyer, R., van der Donk, W.A., Zhao, H., 2003. Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. Biochemistry 42, 11604.

Zhao, H., Arnold, F.H., 1997. Optimization of DNA shuffling for high fidelity recombination. Nucleic Acids Res. 25, 1307.