Yeast Display for the Identification of Peptide-MHC Ligands of Immune Receptors

Citation: Huisman, Brooke D, Grace, Beth E, Holec, Patrick V and Birnbaum, Michael E. 2022. "Yeast Display for the Identification of Peptide-MHC Ligands of Immune Receptors." Methods Mol Biol, 2491.

As Published: 10.1007/978-1-0716-2285-8_15

Publisher: Springer US

Persistent URL: https://hdl.handle.net/1721.1/148130

Yeast display for the identification of peptide-MHC ligands of immune receptors

Brooke D. Huisman[1,2,*], Beth E. Grace[1,2,*], Patrick V. Holec[1,2], Michael E. Birnbaum[1,2]

[1]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[2]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

*These authors contributed equally.

Corresponding author email address: mbirnb@mit.edu

Running header: Yeast-Displayed Peptide-MHC Libraries

Yeast display for the identification of peptide-MHC ligands of immune receptors

Abstract

T cells detect peptide antigens presented by major histocompatibility complex (MHC) proteins via their T cell receptor (TCR). The sequence diversity of possible antigens, with trillions of potential peptide-MHC targets, makes it challenging to study, characterize, and manipulate the peptide repertoire of a given TCR. Yeast display has been utilized to study the interactions between peptide-MHCs and T cell receptors to facilitate high-throughput screening of peptide-MHC libraries. Here we present insights on designing and validating a peptide-MHC yeast display construct, designing and constructing peptide libraries, conducting selections, and preparing, processing, and analyzing peptide library sequencing data. Applications for this approach are broad, including characterizing peptide-MHC recognition profiles for a TCR, screening for high-affinity mimotopes of known TCR-binding peptides, and identifying natural ligands of TCRs from expanded T cells.

Key Words: yeast display, T cell receptor, peptide-MHC, ligand identification

1 Introduction

T cells play a critical role in adaptive immunity, surveilling peptide fragments presented on the cell surface by major histocompatibility complex proteins (MHCs). T cell responses play important roles in immune responses to infection and cancer, and can be detrimental in autoimmunity, allergies, and transplantation. Characterizing the peptide-MHC (pMHC) repertoire of a T cell receptor (TCR) or identifying its ligand(s) is, thus, useful for understanding the

target(s) of a T cell response and potentially for selectively increasing or dampening the T cell response.

However, interactions between TCRs and their pMHC ligands are challenging to study because of their low affinities and large number of possible ligands. Firstly, TCRs generally have low affinities for their ligands, on the order of 1-100 μM, much weaker than antibody affinities for antigens [1]. Furthermore, the space of potential peptide ligands is vast, with more than a trillion theoretically possible pMHC ligands for a given TCR, and thousands to millions of peptide sequences recognized by any given TCR [2, 3].

Yeast display has proven a powerful tool for studying TCR-pMHC binding [2, 4, 5]. In this technique, a library of peptides is displayed on the surface of yeast, covalently linked to a given MHC. Recombinantly expressed TCR protein binds to its pMHC ligands in the library and is used to enrich these yeast. The peptide sequences of enriched yeast are determined via DNA sequencing.

This protocol focuses on methods and considerations for designing, expressing, and selecting pMHC libraries as well as considerations for analyzing the selection results. While we focus on TCR-based selections, other pMHC-binding immune receptors can be studied similarly.

2 Materials

2.1 Cloning and DNA preparation

1. pYAL plasmid (Addgene)
2. pYAL standard primers for sequencing:

a. T7 Promoter: 5’ TAATACGACTCACTATAGGG 3’

b. pYAL Reverse: 3’ GGGATTTGCTCGCATATAGTTG 5’

3. Insert pMHC DNA
4. Phusion 5x HF buffer (NEB)
5. Dimethyl Sulfoxide (DMSO)
6. 40 mM dNTPs (10 mM each, GenScript)
7. Phusion High-fidelity DNA Polymerase (NEB)
8. PCR tubes and 96-well PCR plates
9. Thermocycler
10. Restriction enzymes: FastDigest *NheI* and *HindIII* (Thermo) (see **Note 1**)
11. Genemorph II random mutagenesis kit (Agilent)
12. GenBuilder Cloning Kit (Gibson assembly master mix, GenScript)
13. Competent DH5α cells
14. Corning SOC media (Fisher)
15. LB broth Miller (Fisher)
16. Luria agar (RPI)
17. Carbenicillin (Fisher): make stock solution at a concentration of 100 mg/mL
18. Agarose (Biorad)
19. 50x TAE buffer (RPI)
20. Gel Red (Biotium)
21. Gel combs, casters, and electrophoresis system
22. NucleoBond Xtra Midi Plus kit (Macherey-Nagel)
23. GeneJET PCR Gel Extraction kit (Thermo)
24. GeneJET Plasmid Miniprep kit (Thermo)

25. LB + carbenicillin plates: Make 1 L luria agar according to manufacturer's directions. After autoclaving, stir until cool enough to hold. Add 1 mL of 100 mg/mL carbenicillin, pour plates, and store at 4°C.
26. 14 mL culture tubes
27. 1.5 mL Eppendorf tubes

2.2 Yeast construct validation and library creation

1. Yeast strain EBY100 (ATCC)
2. Disposable electroporation cuvettes with 400 µL volume, 2 mm gap (VWR)
3. Electroporator (BTX ECM 630 Harvard Apparatus)
4. Instrument for measuring OD600
5. FACS buffer: For a 1 L final volume, combine concentrated PBS solution to 1x, EDTA to 2 mM, 5 g BSA, and deionized (DI) water. Sterile filter and store at 4°C.
6. 250 mL and 2.8 L baffled flasks
7. YPD broth (RPI)
8. YPD agar (RPI): autoclave, cool, and pour plates.
9. 0.22 µm sterile filters
10. SDCAA media pH 4.5 [6] (see **Notes 2, 3**): Measure 14.7 g sodium citrate, 4.29 g citric acid monohydrate, 6.7 g yeast nitrogen base, 5 g casamino acids, and 20 g dextrose. Add DI water to 1 L and dissolve. Sterile filter.
11. SDCAA plates [6]: Measure 5.4 g sodium phosphate dibasic, 8.56 g sodium phosphate monobasic, 182 g sorbitol, 15 g agar. Add DI water to 900 mL. Autoclave and cool to below 50°C. Measure 20 g dextrose, 6.7 g yeast nitrogen base, and 5 g casamino acids. Add DI water to 100 mL. Sterile filter and add to cooled autoclaved solution. Pour plates and store at 4°C.

12. SGCAA media pH 4.5 [6] (see **Notes 2, 3**): Measure 14.7 g sodium citrate, 4.29 g citric acid monohydrate, 6.7 g yeast nitrogen base, 5 g casamino acids, and 20 g galactose. Add DI water to 1 L and dissolve. Sterile filter.
13. E buffer: Measure 0.6 g Tris base, 91.09 g sorbitol, and 100 mg $MgCl_2$. Add DI water to 500 mL, sterile filter, and chill to 4°C.
14. Lithium acetate: make a 1 M stock with 1.02 g LiAc in 10 mL DI water
15. Tris-DTT: make a 1 M stock with 0.462 g DTT in 3 mL 1 M Tris pH 8.0--must be freshly made (or, for electrocompetent yeast prep, 2 M DTT in 1 M Tris: 3.09 g DTT in 10 mL 1 M Tris pH 8.0)
16. 0.5 M EDTA
17. 500 mL centrifuge tubes (Corning)
18. Flow cytometer: in the methods below, references are made to running samples on a flow cytometer to check yeast surface staining and to count yeast; counting may be done on an instrument that can make absolute counts or by using the event count within a gate drawn around the yeast population to calculate the total yeast in a larger sample.
19. Polystyrene culture tubes for flow cytometry samples (Fisher)

2.3 Magnetic selections

1. Epitope tag antibodies (Cell Signaling): e.g. anti-Myc antibody clone 9B11, anti-Flag antibody clone D6W5B, anti-HA antibody clone 6E2
2. Anti-human β2M antibody clone 2M2 (Biolegend)
3. Approximately 100 µg of high quality, site-specifically biotinylated, monodisperse TCR protein. These proteins have successfully been used from bacterial protein refolds [2, 4] and insect cell production [4]. Expression in other systems (such as mammalian expression) would likely work comparably, so long as the material is pure, biotinylated, and monodisperse.

4. Fluorescently labeled streptavidin. We produce streptavidin site-specifically labeled with AlexaFluor 647, essentially as previously described [7]. However, commercially available streptavidin-fluorophore conjugates should work equally well, provided careful quantitation is conducted to ensure TCR is completely tetramerized.
5. Tube rotator (for Eppendorf and 15 mL conical tubes)
6. LS columns (Miltenyi Biotec)
7. QuadroMACS Separator (Miltenyi Biotec)
8. MACS Multistand (Miltenyi Biotec)
9. MACS 15 mL Tube Rack (Miltenyi Biotec)
10. Streptavidin MicroBeads (Miltenyi Biotec)
11. Anti-fluorophore MicroBeads (Miltenyi Biotec) (corresponding to the fluorophore used to make TCR tetramers, see #4 above)
12. 15 mL conical tubes

2.4 Next Generation Sequencing preparation

1. Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research)

3 Methods

3.1 Designing and validating a peptide-MHC construct for yeast display

3.1.1 Formatting MHCs for yeast expression

Yeast display allows for surface expression of proteins via fusion with the yeast agglutination factor Aga2, which forms disulfide bonds with Aga1 on the yeast surface [8]. Engineered proteins can be linked to the N- or C-terminus of Aga2; for MHC display on yeast, linkage of the

C-terminus of the MHC to the N-terminus of Aga2 is preferred, as it allows the C-terminus of the displayed peptide to be linked to the N-terminus of the MHC, which reduces the length of all required linkers. Displaying proteins from the N-terminus of Aga2 additionally de-enriches for proteins with frame-shift mutations or containing stop-codons, as they will not contain an in-frame Aga2 protein. **Figure 1** illustrates formatting of MHCI and MHCII constructs.

In MHCI constructs, peptide, beta-2-microglobulin (β2M) and α1-α3 heavy chain ectodomain are expressed as a single-chain trimer [4, 9, 10]. In the single chain trimer format, the N-terminus of the peptide is linked directly to a signal peptide to route the protein for surface expression. The C-terminus of the peptide is linked to the N-terminus of the β2M sequence via a flexible glycine-serine linker (see **Note 4**), which can also contain an epitope tag. The C-terminus of β2M is in turn linked to the N-terminus of the HLA heavy chain via a second glycine-serine linker. The C-terminus of the heavy chain is connected via a final glycine-serine linker (also frequently containing an epitope tag for expression checks) with the N-terminus of Aga2. Human β2M has been shown to stabilize murine MHCI heavy chains and promote peptide binding [11]; we have similarly seen improved expression of murine MHCs on yeast when using human β2M and suggest doing so.

To accommodate the flexible linker connecting the C-terminus of the peptide with the N-terminus of β2M, the peptide-binding groove is opened via a tyrosine to alanine mutation at position 84 of the heavy chain (Y84A) [10]. This mutation may affect the preferences of the MHC F pocket, but has been shown to largely recapitulate the wild-type MHCI binding properties [4].

For yeast-displayed MHCII molecules, we have used two strategies. First, we have linked a ‘mini-MHC’, consisting of the C-terminus of the peptide linked to the N-terminus of the MHCII β1

domain via a glycine-serine linker, which is in turn connected to the α1 domain via a short linker, before finally being linked to the N-terminus of Aga2 [2, 12]. In an alternative strategy, the full length MHC α and β chain ectodomains are expressed in a single transcript, separated by a self-cleaving P2A sequence [13]. In the full-length design, the C-terminus of the leader sequence connects to the N-terminus of the MHC α chain, which in turn adjoins the N-terminus of the P2A sequence via a flexible linker (frequently containing an epitope tag). The C-terminus of the P2A sequence connects to the N-terminus of a second leader sequence, which connects to the N-terminus of the peptide. The C-terminus of the peptide is joined to the N-terminus of the MHC β chain via an additional flexible linker, again frequently containing an epitope tag. The C-terminus of the MHC β chain joins to N-terminus of Aga2 via a final epitope-tag containing linker (see **Note 4**). Both strategies can be successful, with there possibly being preferences for one strategy over another for particular MHCII alleles.

In both MHCI and MHCII cases, the construct is cloned into the pYAL vector for expression in yeast driven by a galactose-inducible promoter. Because the MHCs are presented on the surface of yeast via the yeast Aga1-Aga2 system, the transmembrane and intracellular components of the MHC are removed. Furthermore, in the MHCI formatting, the leader sequence joins directly to the N-terminus of the peptide sequence. Cleavage of the leader sequence is imperfect and may result in truncation of the N-terminus of the peptide sequence (see **Note 5**). The cleavage site of a given leader-peptide combination can be predicted by models such as SignalP [14], Phobius [15], or SigCleave [16]. In some cases, we have found the endogenous Aga2 leader sequence to be incompatible with the expression of certain peptide-MHCs. For these cases, other secretory signal peptides can be used in their place.

Epitope tags can be included in linker regions to allow for rapid detection of MHC surface expression on yeast. Surface expression, however, does not always equate to proper folding on

the cell surface. Additional steps are necessary to validate proper folding of a construct on the yeast surface.

3.1.2 Validating MHC expression and folding on yeast: tetramer staining

Individual MHC constructs need to be verified for yeast-surface expression. This is because the simple protein-folding machinery in yeast can result in some MHCs, particularly class II molecules, not folding properly despite detectable surface expression [2, 12, 13]. **Figure 2** provides an overview of validating the expression of a yeast construct. While antibodies against the MHC can be used to assess fold, some antibodies may not be able to bind even if the construct is folded properly on the surface of the yeast; the antibody binding sites may be inaccessible due to the presence of the linkers in the MHC constructs designed for yeast display. Conversely, we have observed constructs for which antibodies bind yet TCRs do not, possibly representing constructs that are only partially folded. Thus, in our experience, the most effective way to validate MHC folding is via binding of a known peptide and TCR to that MHC. Repositories such as the IMGT 3D structure database or Protein Data Bank are good sources for known MHC-restricted TCRs and their peptide ligands [17, 18]. When selecting a TCR-peptide-MHC set for validating MHCI expression, confirm that the peptide-signal cleavage is predicted to be at the current junction between the yeast signal sequence and peptide using a prediction tool such as SignalP [14] or Phobius [15]. For MHCII, utilize a peptide with a favorable 9mer MHC binding core and peptide flanking region (PFR) extending several amino acids beyond the P1 and P9 pockets, as PFR sequences have been shown to affect peptide-MHC binding affinity [19–21] .

Recombinantly expressed soluble TCR protein can be used to confirm expression and fold of the yeast display-presented MHC through TCR-tetramer staining, in which TCRs are tetramerized using fluorescent streptavidin, and binding to the yeast-displayed MHC is

assessed by flow cytometry. Because of the low affinity of the TCR-pMHC interaction, it is necessary to increase the avidity of this interaction using tetrameric TCRs. An alternative approach for validating the construct is tag enrichment, described in Section 3.1.3.

1. Clone the MHC construct with validation peptide into the pYAL vector using Gibson assembly.
2. Thaw electrocompetent EBY100 cells on ice (see **Note 6**).
3. Add 1 µL plasmid DNA (~100-500 ng) to cells and flick the tube to mix.
4. Let this sit in ice while chilling a cuvette (~5 minutes).
5. Transfer the yeast to a chilled cuvette and electroporate with settings 500 V (LV mode), no resistance, 25 µF capacitance.
6. Resuspend yeast in 1 mL YPD and transfer to a 1.5 mL tube for the yeast to recover.
7. Shake at 250 rpm, 30°C for 1 hour.
8. Spin yeast at 5000xg for 1 minute at 4°C. Aspirate or decant the YPD from the cell pellet.
9. Resuspend cell pellet in 50-100 µL SDCAA media.
10. Plate 10% of resuspended yeast on SDCAA plate. Isolation streak and grow 2-3 days at 30°C until colonies are visible.
11. Pick a single colony. In a 14 mL culture tube, grow the colony in 3 mL SDCAA, shaking at 250 rpm at 30°C for 1-2 days until culture is confluent.
12. Measure OD600. Calculate volume of culture that will give enough yeast for OD 1 in a 3 mL culture (OD 1 = $1x10^7$ yeast/mL).
13. Spin this volume down at 5000xg for 1 minute at 4°C. Aspirate or decant the supernatant.
14. Resuspend in SGCAA media for yeast induction. In a 14 mL culture tube, grow in 3 mL of SGCAA for 2-3 days at 20°C, shaking at 250 rpm.

15. Check construct expression by staining 200 µL of yeast in FACS buffer with epitope tag and/or β2M antibodies and running the sample on a flow cytometer. We expect approximately 40-50% of the yeast to be positive, though we have observed constructs with between ~20-90% positive staining.
16. Create tetramers of TCR by making a 100 µL solution of 500 nM fluorescently labeled tetrameric streptavidin and 2500 nM biotinylated TCR (5:1 TCR:tetrameric streptavidin ratio) in FACS buffer. Incubate 5-10 minutes at 4°C to allow tetramers to assemble.
17. Measure OD of yeast culture. Remove $1x10^7$ yeast, wash in FACS buffer, and resuspend in the tetramer solution.
18. Incubate 30 minutes at 4°C in the dark.
19. Wash with FACS buffer and run on a flow cytometer to check for tetramer staining. Detectable staining indicates that the construct is expressing and properly folded on the surface of the yeast. Lack of staining suggests that the construct will need to be further optimized for expression on yeast; this process is described in greater detail below.

3.1.3 Validating MHC expression and folding on yeast: tag enrichment

If the TCR-pMHC interaction is sufficiently low affinity, binding may not be detectable by tetramer staining (described in Section 3.1.2) and may necessitate using higher avidity streptavidin-coated MACS MicroBeads. This method does not directly determine TCR binding, but in our experience, it can also indicate whether folding of the pMHC has been successful. Perform steps 1-15 in Section 3.1.2 before proceeding with the steps below.

1. Take 1e6 yeast induced to express the desired pMHC with Myc tag and $1x10^7$ uninduced yeast. Mix these two populations of yeast to obtain a sample with a ~1:10 ratio of induced to uninduced yeast.

2. Spin at 5000xg for 1 minute at 4°C. Discard supernatant and resuspend the mixture in 50 µL FACS buffer.
3. Add 1:50 dilution of Myc antibody and stain for 20 minutes at 4°C.
4. Spin at 5000xg for 1 minute at 4°C. Discard supernatant and resuspend in 500 µL FACS buffer.
5. Take 50 µL to run on a flow cytometer to determine a starting percentage of Myc-stained yeast (this should be around 10%).
6. In 50 µL of streptavidin MicroBeads, add TCR to a final **biotinylated** TCR concentration of 400 nM.
7. Add the TCR-loaded beads to the remaining 450 µL of yeast mixture from step 4. Rotate at 4°C for 2 hours.
8. Spin down (do not decant yet) and keep on ice.
9. In the 4°C cold room, place an LS column on the magnet stand. Place a 15 mL conical tube below. Add 5 mL FACS buffer to the column for equilibration. Dump this out from the 15 mL conical tube once it runs through the column.
10. Decant yeast tube and resuspend in 5 mL FACS buffer. Load solution onto the column.
11. Once all of the liquid has flowed through the column, add 3 mL FACS buffer. Repeat twice more for a total of 3x3 mL washes.
12. At this point, any yeast that are bound to the streptavidin beads are magnetically trapped in the column. The rest of the yeast are in the 15 mL conical tube. Cap and remove the tube of unselected yeast.
13. Remove the LS column and place it into a new 15 mL conical tube. Add 5 mL FACS buffer and use the plunger to force the liquid through the column. Take a 500 µL sample. Run the sample on a flow cytometer to determine the percentage of yeast that are Myc-tag positive. If the pMHC is correctly folded, the column-eluted population will be majority epitope tag positive, compared to the ~10% positive population prior to the column. Lack

of enrichment of the Myc-tagged population suggests that the pMHC construct will need to be further optimized for expression on yeast; this process is described in greater detail below. Examples of successful and unsuccessful tag enrichment are shown in **Figure 3**.

3.1.4 Optimizing MHCs for expression on yeast

In the event that the MHC is not properly folded on the surface of yeast, stabilizing mutations may be necessary to improve surface expression and fold. These stabilizing mutations can be identified via error-prone mutagenesis of the MHC and creation of a random mutant library, followed by selection for binding with the cognate TCR. The steps for generating this library are similar to generating a peptide library which will be discussed in more detail below and have also been described elsewhere [6]. Once an error prone library is created, we suggest checking the sequences of 10-12 individual colonies (see Section 3.3.8) to determine an average error rate (e.g. mutations/kb).

To ensure any acquired mutations minimally affect identified peptide sequences, if possible the mutations on the MHC should not directly interact with peptide or with the TCR itself. Here, it is helpful if the validation TCR-pMHC has a reported structure, as the interaction interfaces and the mutation positions can be directly compared. Absent a structure of the given pMHC-TCR complex, the compendium of other structures can give insight into MHC residues that are commonly involved in either peptide or TCR binding [22].

3.2 Designing and constructing libraries

When designing a peptide library, peptide length and extent of diversity should be considered. First, the peptide lengths most commonly displayed by the MHC should be identified. Generally, MHCI displays peptides 8-10 residues in length, and MHCII displays peptides 13-25 residues in length with 9 amino acids binding directly in the peptide binding groove [23]. Since peptides of

different lengths are differentially oriented in the MHCI binding groove, each peptide length of interest must be assessed separately.

Next, consideration must be made for how to balance sequence diversity and library coverage. For example, if the goal is TCR deorphanization, all residues of the peptide can be randomized using degenerate codons. For a randomized position, use of degenerate codon NNK is suggested over NNN to eliminate 2 of 3 potential stop codons (N=A,C,G,T; K=G,T). Amino acid constraint of anchor residues for the MHC being used can help ensure that the created library consists mainly of peptides able to bind to the MHC, increasing the effective diversity of the library. In this case, degenerate codons are chosen which restrict the amino acids that can be encoded at the anchor residue positions. For a well-studied MHC, anchor residue amino acids can be selected from the literature or other resources, such as NetMHC Motif Viewer, IEDB, and SYFPEITHI [24–27]. For example, for murine MHCI H-2D^{b}, based on literature [28] and online resources mentioned above, 9mer and 10mer peptides are most commonly displayed, with position 5 (P5) often being asparagine and the C terminal residue (PΩ) often being leucine, isoleucine, methionine, or valine. In this case, P5 could be fixed to asparagine, and PΩ could be encoded by degenerate codon VTR to allow only those four amino acids (V=A,C,G; R=A,G). In contrast, if the goal is to identify mimotopes of previously identified peptides, specifically chosen residues can be randomized. For example, to search for a peptide mimotope with improved binding affinity for a known TCR, putative TCR contact residues can be specifically targeted for mutagenesis.

For MHCII libraries, 9 amino acids sit in the MHC binding groove, although peptide flanking regions extending outside of the groove can affect peptide-MHC binding affinity [19–21]. It is important to note that the peptide-binding register is not predetermined with a randomized library. Register can be biased by constraining the identity of anchor positions or the peptide

flanking residues [13]. Additionally, if the first amino acid of the linker sequence is a preferable position 9 amino acid, the peptide may bind with the linker binding in the position 9 groove.

Depending on the library design and purpose, different methods can be used to prepare the library insert. Error prone mutagenesis, described in Section 3.2.1, is used when randomizing the entire MHC construct. The method in Section 3.2.2 is used when only a select portion of the construct should be randomized, such as creating a library of unique peptides.

3.2.1 Insert preparation: error prone mutagenesis

Error prone mutagenesis is used when randomizing the whole MHC construct; if the MHC is not folding properly upon initial validation testing, error-prone mutagenesis can be used to introduce mutations to the entire construct. Selections can then be performed using a known cognate TCR for the peptide-MHC to enrich yeast displaying properly folding MHCs. These yeast can be sequenced to determine the stabilizing mutations needed for proper folding.

1. Prepare a 50 µL PCR reaction:
    a. 41.5 µL DI water
    b. 5 µL 10x Mutazyme II reaction buffer
    c. 1 µL 40 mM dNTP mix
    d. 0.5 µL of primer mix (250 ng/µL each primer)
    e. 1 µL Mutazyme II DNA polymerase
    f. 1 µL template DNA (100 ng/µL)
2. Run sample on thermocycler:
    a. 95°C for 2 minutes
    b. 30x 95°C for 30 seconds, 56°C for 30 seconds, 72°C for 1 minute
    c. 72°C for 10 minutes
    d. 12°C hold

3. While the PCR is running, prepare a 1% agarose TAE gel by adding 0.50 g agarose to 50 mL 1x TAE solution. Microwave until agarose has dissolved. Add 1:10000 dilution of Gel Red, mix, and pour the gel with an 8-well comb.
4. Add 10 µL 6x loading dye to the 50 µL PCR sample. Transfer to a well on the gel. Add DNA ladder to another well and run the gel at 110 V for 30 minutes.
5. Image the gel, extract PCR product, and purify DNA via gel extraction kit.

3.2.2 Insert preparation: peptide library

This method is used when only a select portion of the construct should be randomized; for example, when creating a library of unique peptides. We recommend beginning with template DNA containing a stop codon in the peptide sequence or containing no peptide at all (Aga2 leader directly followed by GGGS linker and β2M, etc.). This ensures that any uncut vector appearing in the final library does not encode a functional pMHC sequence, as contamination of a given peptide sequence may result in that peptide overtaking selections, especially if trying to identify new peptide epitopes for a TCR with an already identified ligand. When cloning a peptide-less or stop codon-containing vector, changing the epitope tag can aid in quickly identifying any contamination of validation construct in the library. Reference primer sequences are provided in **Note 7**.

1. Prepare a 50 µL PCR reaction:
    a. 35 µL DI water
    b. 10 µL 10x HF buffer
    c. 1.5 µL DMSO
    d. 1 µL 40 mM dNTPs
    e. 1 µL 10 uM forward primer
    f. 1 µL 10 uM reverse primer
    g. 0.5 µL insert template DNA

h. 0.5 µL Phusion polymerase

2. Run sample on thermocycler:
    a. 98°C for 3 minutes
    b. 30x 98°C for 10 seconds, 56°C for 30 seconds, 72°C for 30 seconds
    c. 72°C for 5 minutes
    d. 12°C hold
3. While the PCR is running, prepare a 1% agarose TAE gel by adding 0.50 g agarose to 50 mL 1x TAE solution. Microwave until agarose has dissolved. Add 1:10000 dilution of Gel Red, mix, and pour the gel with an 8-well comb.
4. Add 10 µL 6x loading dye to the 50 µL PCR sample. Transfer to a well on the gel. Add DNA ladder to another well and run the gel at 110 V for 30 minutes.
5. Image the gel, extract PCR product, and purify DNA via gel extraction kit.

3.2.3 Insert preparation: scale up PCR

To create large libraries, ideally 50µg of insert DNA is needed [6] (see **Note 8**). To prepare this insert DNA, a scaled up PCR reaction is performed that is 100x a typical 50 µL PCR reaction. Primers for this step are designed to anneal to the PCR amplification product from Section 3.2.1 or 3.2.2 and to extend the fragment to include ~50 nucleotides of overlap with the digested vector. A small amount of the new PCR product is checked on a gel to ensure that the reaction is clear and the desired band size is present, and then the rest of the scaled up PCR is typically purified via a PCR purification kit.

1. Prepare a 5 mL master mix:
    a. 3.5 mL DI water
    b. 1 mL 5x HF buffer
    c. 150 µL DMSO
    d. 100 µL 40 mM dNTPs

    e. 10 µL 100 uM forward primer
    f. 10 µL 100 uM reverse primer
    g. 50 µL Phusion polymerase
2. Mix well and aliquot 100 µL/well in a 96 well PCR plate.
3. Cover and run on thermocycler:
    a. 98°C for 3 minutes
    b. 30x 98°C for 10 seconds, 56°C for 30 seconds, 72°C for 30 seconds
    c. 72°C for 5 minutes
    d. 12°C hold
4. While PCR is running, prepare a 50 mL 1% agarose TAE gel with 1:10000 dilution of Gel Red and an 8-well comb.
5. Pool solution from each well into a 50 mL conical tube. Mix well.
6. Remove 15 µL and add to 3 µL 6x loading dye. Run this sample on a gel at 110 V for 30 minutes.
7. If the band is clean and of expected size (see **Note 9** for troubleshooting), purify the remaining PCR solution using 6 columns. Combine elutions and check concentration. At least 50 µg of insert is ideal for library creation.

3.2.4 Vector preparation

For library creation, ideally 10 µg of linearized vector DNA is needed. Thus, vector is typically prepared at midiprep scale. The resulting plasmid is then digested with appropriate restriction enzymes and purified.

1. Thaw 100 µL DH5α competent bacteria on ice.
2. Add ~100 ng of plasmid to the bacteria.
3. Incubate on ice for 15 minutes.
4. Heat shock at 42°C for 45 seconds.

5. Rescue with 500 µL SOC media.
6. Shake at 250 rpm at 37°C for 45-90 minutes.
7. Add directly to 100-200 mL LB (a larger culture should result in enough DNA to allow for multiple library creation attempts, if needed) with a 1:1000 dilution of carbenicillin in a sterile baffled flask. Grow overnight at 37°C while shaking at 250rpm.
8. Midiprep using NucleoBond Xtra Midi Plus kit.
9. Digest based on volume eluted from midiprep:
    a. Volume <340 µL:
        i. *X* µL vector DNA
        ii. 340 - *X* µL water
        iii. 40 µL 10x green FD buffer
        iv. 10 µL FastDigest *NheI*
        v. 10 µL FastDigest *HindIII*
    b. Volume >340 µL:
        i. *X* µL vector DNA
        ii. 430 - *X* µL water
        iii. 50 µL 10x green FD buffer
        iv. 10 µL FastDigest *NheI*
        v. 10µL FastDigestD *HindIII*
10. Incubate at 37C for 90 minutes.
11. During incubation, prepare a 100 mL 1% agarose gel with 1:10000 Gel Red in a wide gel box with a single comb.
12. Run the entire sample on gel at 110 V for 30-40 minutes or until adequate separation of vector and insert bands is achieved.
13. Extract the large vector band into a 15 mL conical tube.

14. Extract DNA from gel using 4-6 columns of a gel extraction kit. Elute DNA and pool. 10 ug of vector DNA is optimal for library creation.

3.2.5 Library creation

Yeast are transformed using the insert and vector DNA generated above, following a strategy similar to that described by Chao et al [6].

1. Isolation streak EBY100 yeast on a YPD plate. Grow 2 days at 30°C.
2. Inoculate 4-6 x 4 mL YPD with swipes of EBY100 colonies. Grow overnight at 30°C.
3. In the morning, inoculate 2 x 100 mL YPD to an optical density (OD600) of 0.2 in sterile 250 mL baffled flasks.
4. Grow at 30°C, shaking at 250 rpm, until OD600 = 1.5-1.67. This usually takes 4-6 hours.
5. While waiting, chill centrifuges and ~25-30 cuvettes. Prepare and filter the following:
    a. 500 mL E buffer (chilled to 4°C)
    b. 10 mL 1 M Lithium Acetate (LiAc)
    c. 1L SDCAA
    d. In the last hour, 3 mL 1M Tris-DTT--must be fresh.
6. When OD600 measures 1.5-1.67, add 1 mL Tris-DTT, 5 mL LiAc, and 200µL 0.5M EDTA to each flask.
7. Shake at 250 rpm, 30°C for 15 minutes.
8. While the flasks are shaking, combine 50 µg insert with 10 µg vector. Top up to 500 µL with cold E buffer, and sterile filter the DNA solution.
9. Divide the cultures into four 50 mL conical tubes.
10. Spin at 2500xg for 3 minutes at 4°C. Dispose of supernatants.
11. Keeping the cells on ice, resuspend each pellet in 50 mL cold E buffer.
12. Spin at 2500xg for 3 minutes at 4°C. Dispose of supernatants.

13. Resuspend two pellets with 25 mL cold E buffer each and combine into two conical tubes (ending with two tubes instead of four).
14. Spin at 2500xg for 3 minutes at 4°C. Dispose of supernatants.
15. Resuspend one pellet with 10 mL cold E buffer and use to resuspend both pellets.
16. Spin at 2500xg for 3 minutes at 4°C. **Aspirate** supernatant.
17. Resuspend pellet fully in filtered DNA solution from step 8.
18. Taking 50 µL at a time (see **Note 10**) and ensuring the yeast are well mixed, electroporate in chilled cuvettes with settings 500 V (LV mode), no resistance, 25uF capacitance. Rescue with 2 x 1 mL YPD per cuvette into a fresh 250 mL sterile baffled flask. Time constants should be ~40-60 ms.
19. Shake at 250 rpm, 30°C for 1 hour after rescue. Allow 4 SDCAA plates to warm up.
20. Pour into one 50 mL tube and spin at 2500xg for 3 minutes at 4°C. **Aspirate** supernatant.
21. Resuspend in 10 mL SDCAA.
22. Titer onto four warm SDCAA plates. Change pipette tip for each dilution.
    a. 10^4: 990 µL SDCAA + 10 µL library (from 10 mL in step 21). Mix well.
    b. 10^5: 900 µL SDCAA + 100 µL 10^4 above, mix well.
    c. 10^6: 900 µL SDCAA + 100 µL 10^5 above, mix well.
    d. 10^7: 900 µL SDCAA + 100 µL 10^6 above, mix well.
23. Spread 100 µL of dilution onto each plate. Incubate 2-3 days at 30°C until colonies are large enough to count.
24. Pour the remaining library into a sterile 2.8 L baffled flask with 490 mL SDCAA, using some SDCAA to wash the conical tube. Shake at 250 rpm, 30°C.
25. The following day, split the library flask 1:6 by pouring ~100 mL of library into the bottle containing ~500 mL SDCAA. Pour out the remaining library and pour the split library back into the flask.

26. 2-3 days after library creation, check the plates. Count colonies on the 10^6 and 10^7 plates. Use the smaller count as library size (e.g. if there are 12 colonies on the 10^7 plate and 150 colonies on the 10^6 plate, record the library size as $1.2x10^8$). See **Note 11** regarding library size.
27. 1-2 days after the library is split, remove it from the shaker and save 4x50 mL conical tubes of the library. Store at 4°C labeled with titer and OD for up to a month (see **Note 12** for long-term storage tips). Remaining library can be induced immediately for use in selections or discarded.

3.3 Conducting selections

Once a library of adequate size has been created, selections can be performed using the library. First, the library must be grown and induced to express the pMHC on the yeast surface.

3.3.1 Inducing the library

1. If starting from a freshly made or freshly thawed and passaged library, skip to step 5.
2. If starting from yeast stored at 4°C, check the OD and calculate the volume needed to obtain 10-20x more yeast than the size of the library to ensure no overall library diversity is lost via sampling (e.g. if the library is $1x10^8$, start selections on $1\text{-}2x10^9$ yeast).
3. In a sterile baffled flask, dilute this amount of yeast down to OD 1 in SDCAA.
4. Grow for 1 day at 30°C, shaking at 250 rpm.
5. Check the OD. Beginning with at least 10-20x the number of yeast as the titered library size, grow in SGCAA for 2-3 days at 20°C, shaking at 250 rpm.

3.3.2 Assessing induction

1. Use 200 µL of induced yeast culture for an unstained control and 200 µL to stain with the relevant epitope tag or β2M antibody.

2. Spin at 5000xg for 1 min at 4°C. Remove supernatant and wash with 500 µL of FACS buffer.
3. Resuspend in 50 µL FACS buffer + 1:50 dilution of antibody. Incubate 20 minutes at 4°C.
4. Wash twice with FACS buffer.
5. Run samples on a flow cytometer to assess percent of yeast that are induced (displaying protein on surface). For an initial library, 15-50% induction is typical. Much lower, something may be wrong with expression. Much higher may suggest a contaminant.

Next, design the selection strategy. If the goal is to select for peptide binders of a given TCR, a general strategy would be to do 3 rounds of selection with TCR on streptavidin beads and then increase the stringency by switching to tetramers of TCR for rounds 4+. Additional considerations might include: tetramer concentration (lower to increase stringency further), negative selections with another TCR for which no binding is expected or desired (to select for specificity), or rounds of selections with multiple TCRs for which peptide cross-reactivity is desirable. Bead-based selections are conducted using streptavidin-coated magnetic microbeads which are flowed through a column attached to a magnet. Sample selection data is shown in **Figure 4**.

3.3.3 Starting selections: removing non-specific binders

Especially in early rounds of selection, non-specific binders should be removed from the library by incubating yeast with streptavidin beads only. Any yeast that bind non-specifically to the beads will be retained in the column, and the TCR-specific selection can be performed on the cleared yeast.

1. To begin selections, check the OD600 of the induced library. Calculate the volume of culture needed to examine 10-20x yeast as compared to the starting library size. Spin down this amount of yeast at 5000xg for 1 minute at 4°C.
2. Wash with 10 mL of FACS buffer. Spin at 5000xg for 1 minute at 4°C.
3. Resuspend in 5 mL and transfer to 15 mL conical tube.
4. Add 250 µL streptavidin beads. Place on a rotator at 4°C for 1 hour.
5. Spin down (do not decant yet) and keep on ice.
6. In the cold room, place an LS column on the magnet stand. Place a 15 mL conical tube below. Add 5 mL FACS buffer to the column for equilibration. Dump this out from the 15 mL conical tube once it runs through the column.
7. Decant yeast tube and resuspend in 5 mL FACS buffer. Load solution onto the column.
8. Once all of the liquid has flowed through the column, add 3 mL FACS buffer. Repeat twice more for a total of 3x3 mL washes.
9. At this point, any yeast that are non-specifically bound to the streptavidin beads are magnetically trapped in the column. The rest of the yeast are in the 15 mL conical tube (cleared library). Cap and save the tube of the cleared library--these are the yeast to continue working with for TCR-specific selection.
10. Remove the LS column and place it into a new 15 mL conical tube. Add 5 mL FACS buffer and use the plunger to force the liquid through the column. Take a 500 µL sample and run it on the flow cytometer to count how many non-specific binders are present. Record this number for future reference to compare specific and non-specific enrichment.

3.3.4 TCR-specific selection: round 1

1. With the 15 mL conical tube of cleared yeast, spin at 5000xg for 1 minute at 4°C. Resuspend the yeast in 5 mL of FACS buffer. Keep on ice.

2. Incubate 250 µL streptavidin beads with protein of interest to bring the beads to a final **biotinylated** protein concentration of 400 nM. The degree of biotinylation can be assessed through various methods, such as a gel shift assay [29], and a low degree of biotinylation requires more total protein in selections. For example, if a TCR is only 50% biotinylated, calculate the amount of protein for 400 nM and then use double that amount to account for incomplete biotinylation. Mix well and incubate 5 minutes on ice.
3. Add the beads to the resuspended yeast. Rotate at 4°C for 2-3 hours.
4. Repeat steps 5-8 described for clearing non-specific binders above (Section 3.3.3). Now, the yeast that flowed through the column are undesirable yeast--they did not bind the TCR. Cap and remove the 15 mL conical tube.
5. The yeast that have putatively bound to the protein of interest are magnetically bound to the column. Elute from the column as above (Section 3.3.3, step 10) and save the elution containing the selected yeast. Take a 250 µL sample and run on a flow cytometer to count selected yeast.
6. Spin down the selected yeast at 5000xg for 1 minute at 4°C. Discard supernatant and resuspend in 5 mL SDCAA to serve as a wash to remove residual FACS buffer.
7. Spin at 5000xg for 1 minute at 4°C. Discard supernatant and resuspend in 3 mL SDCAA. Transfer to a 14 mL culture tube and grow at 30°C overnight, shaking at 250 rpm.
8. After growing overnight, if continuing to further rounds of selection immediately (see **Note 13**), check OD and induce 10-20x more yeast than were selected in round 1. Induce in SGCAA at 20°C for 2-3 days, shaking at 250 rpm.

3.3.5 TCR-specific selection: rounds 2-3

When checking induction, the percent of yeast that stain with an epitope tag antibody typically increases in successive rounds of selection (see **Figure 4**). The number of non-specific binders should decrease in further rounds (see **Note 14**).

1. For rounds 2-3, use the protocol above (Sections 3.3.3 and 3.3.4) but with the yeast in 500 µL instead of 5 mL for incubation with beads, and use a 50 µL volume of streptavidin beads instead of 250 µL.
2. Continue to start with at least 10-20x more yeast than were selected in the previous round. If that calculated number is impractically small to manipulate, there is no detriment in using a larger number of yeast (such as $3x10^7$, which can be cultured in 3 mL). After round 2, it may take more than 24 hours for yeast to grow back up in SDCAA if the number selected was low.

3.3.6 TCR-specific selection: rounds 4+

We find that for most selections, if a library contains pMHC binders, they will be apparent after round 3 at latest. There are also often significant numbers of pMHCs that may specifically interact with highly avid TCR-coated microbeads yet not with lower-avidity multimers such as tetramers. Therefore, for round 4, consider increasing the stringency of selection by incubating the yeast with tetramers of TCR rather than streptavidin beads to separate out pMHC sequences with higher affinity. Note that since the physiological affinity of the pMHC-TCR interaction is low (~1-50 µM), few if any library sequences would be expected to bind if monovalent TCRs were used for selection. Tetramer-positive yeast are enriched using magnetic beads decorated with antibodies to the tetramer fluorophore. Alternatively, for small libraries, these selections could be conducted using fluorescence-activated cell sorting.

1. Create tetramers of TCR by making a 100 µL solution of 500 nM fluorescently labeled tetrameric streptavidin and 2.5 uM biotinylated TCR in FACS buffer (5:1 TCR:tetrameric streptavidin ratio, ensuring complete tetramer formation with margin for error). Incubate 5-10 minutes at 4°C to allow tetramers to assemble.
2. Wash starting yeast with FACS buffer; spin down and remove supernatant. Resuspend in 100 µL of tetramer solution. Place on a rotator at 4°C for 2-3 hours in the dark.

3. Wash twice with 500 µL of FACS buffer.
4. Resuspend in 500 µL FACS buffer. Take a 50 µL sample to run on a flow cytometer and assess the percentage tetramer-positive yeast pre-column.
5. Add 50 µL anti-fluorophore beads to the remaining 450 µL of yeast. Incubate 20 minutes at 4°C.
6. Run selection on LS column as described above for other rounds of selection (Section 3.3.3, steps 5-8). Elute selected yeast in 5 mL FACS buffer (similar to Section 3.3.3, step 10). Run a 250 µL sample on a flow cytometer to count selected yeast and assess tetramer-positive percentage--there should be some enrichment of tetramer positive yeast compared to the pre-column sample.
7. Wash yeast with 5 mL SDCAA. Resuspend in 3 mL SDCAA and grow overnight as above.
8. For further rounds, tetramer concentration can be lowered to increase selection stringency even further. It may be helpful to titrate tetramer on samples of yeast that have grown up following the previous round of selection to determine an optimal tetramer concentration.

3.3.7 Growing and sequencing selected yeast: individual yeast colonies

Once there is an enriched population of TCR-binding yeast, it can be advantageous to sequence individual yeast clones to estimate the sequence diversity of the enriched pool and to ensure that the sequences are well behaved (in frame, predicted to bind to MHC, and not contaminants from previous experiments). To ‘spot check’ hits before deep sequencing, individual yeast colonies can be grown to test tetramer staining or to sequence the peptides (see **Note 15** for troubleshooting). An alternative is described in Section 3.3.8, which provides peptide sequences faster though does not allow for tetramer staining.

1. After the yeast have grown up post-selection, plate 20 µL on an SDCAA plate and isolation streak. Grow 2-3 days at 30°C until colonies are visible.
2. Add 1 mL SDCAA into wells of a 2 mL 96-well block. Pick colonies from the plate and add to wells. Cover with breathable film.
3. Grow at 30°C for 1 day.
4. From here, either or both of these steps can be performed:
    a. Option 1: Conduct a yeast miniprep on 100 µL of culture as described by manufacturer. Transform 10 µL elution in DH5α bacteria. Grow overnight at 37°C with 1:1000 dilution of carbenicillin. Miniprep bacteria and submit DNA for sequencing to see peptide sequences.
    b. Option 2: spin plate, dump supernatants, and resuspend yeast in SGCAA. Shake at 20°C at 250 rpm for 2-3 days. Test individual colonies for tetramer staining.

3.3.8 Growing and sequencing selected yeast: bacteria colonies

As an alternative to growing and sequencing individual yeast colonies (Section 3.3.7), individual bacteria colonies transformed with plasmid purified from the enriched pool of yeast binders can be sequenced.

1. After the yeast have grown up post-selection, take a sample with 10x the amount selected and perform a yeast mini prep as described by the manufacturer.
2. Transform into DH5α bacteria and plate.
3. Pick a dozen or so colonies and grow overnight in 3 mL LB with 1:1000 dilution of carbenicillin.
4. Mini prep and send for sequencing.
5. If desired, DNA can be transformed back into yeast for further tetramer staining or analysis. Transformation of DNA into yeast was described above in Section 3.1.2.

3.4 Preparing, processing, and analyzing NGS data

After three rounds of TCR-based selections, the peptide library is likely highly converged such that sequencing individual colonies will provide a sense of what is in the library. However, depending on the diversity and size of the starting library, Next Generation Sequencing (NGS) may provide important insight into the content of the library, including identifying low-abundance sequences that may bind with lower affinity, covariation between residues, and relative enrichment of peptide sequences. Sequencing can be performed on the Illumina MiSeq (Illumina Incorporated; Irvine, CA), which can accommodate limited-diversity amplicon sequencing while providing sufficient coverage for an enriched yeast display library with ~10 million reads.

3.4.1 Preparing sample for NGS

To prepare the sample for NGS, first perform a yeast miniprep to extract plasmid DNA from yeast, following manufacturer kit instructions (see **Note 15** for troubleshooting). Sequencing each round of selection -- not just the final round -- is recommended to gain insight into the convergence of the library as selections progress. Ensure that the number of yeast utilized in the miniprep exceeds the expected diversity in that round. The exception is with the unselected library, which is frequently intentionally not sequenced to full depth. For example, if the library starts at 100 million members, sequencing to full depth would require more reads than a standard MiSeq run. Additionally, because each individual member is present at low frequency, it will be challenging to distinguish between true members and PCR errors. Typically, including 10 million - 50 million yeast into the mini prep for each round is sufficient.

In a standard NGS prep, PCR primers are designed to incorporate three components: sequence matching desired sequencing primers, i5 and i7 paired-end handles, and an index barcode

unique to each sample being included in the multiplex (e.g. one index barcode for each round of selection). Depending on the sequencer, it may also be beneficial to include a short randomized sequence for cluster identification on Illumina sequencers [30]. These additions are usually made in two rounds of PCR as illustrated in **Figure 5**. Typically, a 150+150 paired-end read will sufficiently cover the peptide sequence, although the primer positions can be moved to cover additional sequences of interest, such as signal peptides, flanking linkers, or epitope tags.

3.4.2 Processing NGS data

Paired end reads can be assembled utilizing one of many available tools, such as FLASH [31] or PandaSeq [32], and can increase sequence fidelity by correcting for sequencing errors. Additional filtering steps include removing sequences with errors in regions flanking the peptide or in constant regions of a partially-randomized peptide. Additionally, the most probable library contaminants are the constructs previously used for validation, which may need to be excluded.

During each NGS processing step, low frequency errors may arise, such as polymerase misincorporation during PCR amplification. Because these errors are typically stochastic and low frequency, they can be identified by processing each sample in replicate and comparing peptide presence and frequency across replicates.

3.4.3 Analyzing NGS data

From the sequencing run, there are two pieces of information with which to understand the library: peptide sequences and the read count for each peptide. TCR selection data is frequently highly hierarchical, with ~10% of the peptides that are present after round three of selection accounting for >90% of the sequencing counts. As a result, read counts may be important to fully understanding the peptide motifs.

Additionally, contaminating sequences and noise will account for some low-frequency peptides in the NGS data. Noise can be filtered by excluding peptides with fewer than a set number of reads. An arbitrary read count may prove helpful (e.g. greater than 10 reads), although read count cutoffs can be rationally set by considering factors like known noise such as stop-codon containing peptides. While it's not uncommon for the stop-codon template plasmid to be a contaminant at higher frequency, stop-codon sequences can be more frequently found at low read counts.

We have found that the false positive rate is low for pMHC-TCR binders. For inferring negative binders, the opposite is true: a peptide that did not enrich may be absent due to stochastic drop-out in early rounds of selection when each member is present at low frequency or due to poor processing in yeast such as incorrect signal peptide cleavage. When extending yeast display-enriched peptides to T cell phenotypic assays, it's important to note that a peptide that binds to a TCR of interest will not always result in T cell activation or phenotypic response [33]. Additionally, yeast display does not incorporate peptide processing pathways present in mammalian cells, meaning a peptide identified through yeast display may not be a natural ligand.

The data resulting from sequencing these yeast can be used for many exciting applications, such as identifying natural TCR ligands [2, 4], determining peptide mimotopes [12], characterizing TCR motifs [2, 13], and training prediction algorithms [2, 4, 13, 34].

4 Notes

1) Restriction sites used for cloning may differ depending on the construct design and/or epitope tags used (e.g. for a construct including a Myc epitope tag, there is a *HindIII* site within the Myc tag used for cloning).
2) As described by Chao et al [6], SDCAA and SGCAA can also be prepared with sodium phosphate dibasic and sodium phosphate monobasic. However, use of sodium citrate and citric acid monohydrate instead are favorable for reducing growth of bacterial contaminants. We have additionally noted that for some MHC constructs, foldedness is improved via low pH induction.
3) If there is persistent contamination in yeast cultures, ensure good aseptic technique. It may be helpful to add Penicillin-Streptomycin at 1:100 to SDCAA and SGCAA media. Following the addition of Penicillin-Streptomycin, store media at 4°C. Filter tips may also reduce the frequency or severity of contamination.
4) When designing MHCI constructs, we typically use the following linker lengths: 3xGGGS between peptide and β2M, 4xGGGS between β2M and MHCI heavy chain, and an epitope tag followed by 3xGGGS between MHCI heavy chain and Aga2. Similarly for MHCII: GSGS connecting MHCIIα to an epitope tag and P2A, GGSGGG--epitope tag--GG--optional protease site--GGSG between peptide and MHCIIβ (protease cut site can be used for peptide-MHC binding selections as done by Rappazzo et al [13]), and GS--epitope tag--3xGGGGS between MHCIIβ and Aga2. If primers may need to bind to part of a linker sequence during any cloning, we suggest varying the DNA sequences of the linkers to avoid multiple primer binding sites.
5) Signal peptide cleavage can impact peptide preferences, specifically by causing the absence of proline as peptide position 1, which may be improperly cleaved from an N-terminally adjacent signal peptide.

6) To prepare electrocompetent yeast, follow steps 1-17 in Section 3.2.5 with the following changes: in step 3, inoculate a 1 L YPD culture in a sterile 2.8 L baffled flask; in step 5, prepare 500 mL E buffer and 10 mL 2 M DTT in Tris; in step 6, add only 10 mL Tris-DTT per liter of cells; in step 9, use 500 mL centrifuge tubes; in step 17, resuspend in 15 mL E buffer, aliquot 50 µL/tube in sterile Eppendorf tubes, and slow freeze to -80°C.
7) Reference primer sequences for a randomized MHCI 9mer peptide library:
    a) Forward--PCR 1 (peptide randomization):
    CAATATTTTCTGTTATTGCTAGCGTTTTGGCTNNKNNKNNKNNKNNKNNKNN
    KNNKNNKGGTGGAGGAGGTTCTGGAGGTG
    b) Forward--PCR 2 (scale up) (homology past NheI site):
    TTCAATTAAGATGCAGTTACTTCGCTGTTTTTCAATATTTTCTGTTATTGCTAG
    CGTTTTGGCT
8) For library preparation, if DNA yields are low, a library can be made with less DNA, although library size will decrease. For example, we have successfully made a $1x10^7$ library using 5 ug of insert DNA. Proportionally scale the amount of vector DNA used.
9) Troubleshooting scale-up PCR: one common issue for the scale-up PCRs is a non-specific PCR product that appears on the quality control gel. If there is a non-specific band, optimization of PCR conditions such as altering PCR melting temperatures or DMSO concentration may remove the non-specific band; we recommend performing these optimizations at a small scale (50 µL reaction). If PCR optimizations fail, we recommend gel purifying all of the DNA, rather than PCR purifying, to remove the non-specific amplification product. If there are challenges scaling a 50 µL reaction to 5 mL, a medium scale (~1 mL) reaction may be successful instead.
10) When aliquoting yeast for electroporation, we have found aliquoting individual cuvettes for immediate electroporation or aliquoting 5-10 cuvettes at a time for subsequent electroporation works better than aliquoting all yeast into cuvettes before beginning

electroporation. Time constants are more consistent when the yeast are well mixed and do not settle much before electroporation.

11) Regarding library size: for a fully randomized peptide MHCI library for TCR deorphanization, our goal is typically a library size of at least $1x10^8$, which can be achieved with the protocol in Section 3.2. Other libraries may not need to be so large; for example, a mimotope library with four peptide residues randomized would include 20^4 = 160,000 theoretical peptides, so even a library size of $1x10^7$ would provide excess coverage.
12) Yeast grown in SDCAA may be stored at 4°C. For long term storage of a yeast library, freeze an oversampling of the library. Protocols are available for long term yeast storage and revival: see Supplementary Methods by Chao et al [6] or various other chapters of this volume.
13) Yeast induce most robustly after growing for at least a day. For library selections, we suggest inducing yeast in SGCAA directly after they have been growing in SDCAA. Thus, if pausing between rounds of selection, uninduced yeast can be stored at 4°C, but reculture the library in SDCAA media for a day before inducing in SGCAA.
14) In selection rounds 3+, it may not be necessary to continue removing non-specific binders first, as there should be very few. However, to do so in tetramer rounds, yeast can be incubated first in 100 µL of 500 nM fluorescently labeled tetrameric streptavidin, then with anti-fluorophore beads, and then run on an LS column. The cleared library can be incubated with TCR tetramers as described above for TCR-specific selections in tetramer rounds.
15) Troubleshooting yeast mini prep: one common challenge in performing colony sequencing of post-selection libraries is having no colonies appear after transforming the yeast mini prep into DH5α bacteria. Note that yeast genomic DNA will make the plasmid DNA from a yeast mini prep impure, so it is important to amplify in bacteria and to use

most or all of the yeast mini prep output DNA. When eluting from the yeast mini prep column, allow the water to sit on the column for several minutes. If there continue to be no or few colonies, consider varying the number of yeast utilized in the yeast mini prep.

5 References

1. Matsui K, Boniface J, Reay P, et al (1991) Low affinity interaction of peptide-MHC complexes with T cell receptors. Science 254:1788–1791

2. Birnbaum ME, Mendoza JL, Sethi DK, et al (2014) Deconstructing the peptide-MHC specificity of T cell recognition. Cell 157:1073–1087

3. Wooldridge L, Ekeruche-Makinde J, van den Berg HA, et al (2012) A Single Autoimmune T Cell Receptor Recognizes More Than a Million Different Peptides*. Journal of Biological Chemistry 287:1168–1177

4. Gee MH, Han A, Lofgren SM, et al (2018) Antigen Identification for Orphan T Cell Receptors Expressed on Tumor-Infiltrating Lymphocytes. Cell 172:549–563.e16

5. Adams JJ, Narayanan S, Liu B, et al (2011) T Cell Receptor Signaling Is Limited by Docking Geometry to Peptide-Major Histocompatibility Complex. Immunity 35:681–693

6. Chao G, Lau WL, Hackel BJ, et al (2006) Isolating and engineering human antibodies using yeast surface display. Nature Protocols 1:755–768

7. Ramachandiran V, Grigoriev V, Lan L, et al (2007) A robust method for production of MHC tetramers with small molecule fluorophores. J Immunol Methods 319:13–20

8. Boder ET, Dane Wittrup K (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nature Biotechnology 15:553–557

9. Hansen T, Lawrence Yu YY, Fremont DH (2009) Preparation of Stable Single-Chain Trimers Engineered with Peptide, β2 Microglobulin, and MHC Heavy Chain. Current Protocols in Immunology 87

10. Lybarger L, Lawrence Yu YY, Miley MJ, et al (2003) Enhanced Immune Presentation of a Single-chain Major Histocompatibility Complex Class I Molecule Engineered to Optimize Linkage of a C-terminally Extended Peptide. Journal of Biological Chemistry 278:27105–27111

11. Pedersen LØ, Stryhn A, Holtet TL, et al (1995) The interaction of beta 2-microglobulin (β2m) with mouse class I major histocompatibility antigens and its ability to support peptide binding. A comparison of human and mouse β2m. European Journal of Immunology 25:1609–1616

12. Fernandes RA, Li C, Wang G, et al (2020) Discovery of surrogate agonists for visceral fat Treg cells that modulate metabolic indices in vivo. eLife 9

13. Rappazzo CG, Huisman BD, Birnbaum ME (2020) Repertoire-scale determination of class II MHC peptide binding via yeast display improves antigen prediction. Nature Communications 11

14. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, et al (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37:420–423

15. Kall L, Krogh A, Sonnhammer ELL (2007) Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. Nucleic Acids Research 35:W429–W432

16. Heijne G von, von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Research 14:4683–4690

17. Sussman JL, Lin D, Jiang J, et al (1998) Protein Data Bank (PDB): database of three-dimensional

structural information of biological macromolecules. Acta Crystallogr D Biol Crystallogr 54:1078–1084

18. Kaas Q (2004) IMGT/3Dstructure-DB and IMGT/StructuralQuery, a database and a tool for immunoglobulin, T cell receptor and MHC structural data. Nucleic Acids Research 32:208D–210

19. O'Brien C, Flower DR, Feighery C (2008) Peptide length significantly influences in vitro affinity for MHC class II molecules. Immunome Res 4:6

20. Zavala-Ruiz Z, Strug I, Anderson MW, et al (2004) A polymorphic pocket at the P10 position contributes to peptide binding specificity in class II MHC proteins. Chem Biol 11:1395–1402

21. Lovitch SB, Pu Z, Unanue ER (2006) Amino-terminal flanking residues determine the conformation of a peptide-class II MHC complex. J Immunol 176:2958–2968

22. Rossjohn J, Gras S, Miles JJ, et al (2015) T cell antigen receptor recognition of antigen-presenting molecules. Annu Rev Immunol 33:169–200

23. Wieczorek M, Abualrous ET, Sticht J, et al (2017) Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. Frontiers in Immunology 8

24. Andreatta M, Nielsen M (2016) Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 32:511–517

25. Nielsen M, Lundegaard C, Worning P, et al (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 12:1007–1017

26. Vita R, Mahajan S, Overton JA, et al (2019) The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Research 47:D339–D343

27. Rammensee H, Bachmann J, Emmerich NP, et al (1999) SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50:213–219

28. Falk K, Rötzschke O, Stevanović S, et al (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351:290–296

29. Fairhead M, Howarth M (2015) Site-specific biotinylation of purified proteins using BirA. Methods Mol Biol 1266:171–184

30. Mitra A, Skrzypczak M, Ginalski K, Rowicka M (2015) Strategies for achieving high sequencing accuracy for low diversity samples and avoiding sample bleeding using illumina platform. PLoS One 10:e0120520

31. Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963

32. Masella AP, Bartram AK, Truszkowski JM, et al (2012) PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics 13:31

33. Sibener LV, Fernandes RA, Kolawole EM, et al (2018) Isolation of a Structural Mechanism for Uncoupling T Cell Receptor Signaling from Peptide-MHC Binding. Cell 174:672–687.e27

34. Dai Z, Huisman BD, Zeng H, et al (2021) Machine learning optimization of peptides for presentation by class II MHCs. Bioinformatics. https://doi.org/10.1093/bioinformatics/btab131

Figure Captions

Figure 1. Formatting of yeast display constructs a) in plasmids and b) on yeast surface.
Abbreviations: SP = signal peptide; MHC-HC = MHCI heavy chain; β2M = beta-2-microglobulin

Figure 2. Flow chart for validating a yeast display construct.

Figure 3. Example flow cytometry plots showing Myc tag staining pre- and post-selection with 400nM TCR-loaded streptavidin beads. The left column shows a successfully enriched population of pMHC-expressing yeast, suggesting that the construct is properly folded on the surface of the yeast. The right column shows an unsuccessful enrichment, suggesting that the construct is not properly folded.

Figure 4. Sample selection data for TCR deorphanization. a) Myc epitope tag staining shows enrichment over successive rounds of selection (gated on unstained yeast). b) Selection scheme and data for five rounds of selection on a randomized 9mer library displayed by mouse MHCI H-2D^{b}. In this example, 'spot check' of a dozen colonies following round 3 indicated enrichment of three peptides.

Figure 5. Amplicon sequencing formatting. Short randomized sequences for cluster identification are indicated as "NNNNNN".
Abbreviations: F primer = Forward NGS primer; R primer = Reverse NGS primer; BC = Index Barcode; "N" indicates any nucleotide

Acknowledgements

We would like to thank K. Christopher Garcia (Stanford University) for generous sharing of reagents, and Christine Devlin for aiding in the creation of our yeast display protocols. This work was supported by National Science Foundation Graduate Research Fellowships to B.D.H, B.E.G., and P.V.H., and a Melanoma Research Alliance grant, the AACR-TESARO Career Development Award for Immuno-oncology Research (17-20-47-BIRN), Schmidt Futures, and the National Institutes of Health (P30CA14051 and 5U19AI110495) to M.E.B..