Video Article

Semi-automated Biopanning of Bacterial Display Libraries for Peptide Affinity Reagent Discovery and Analysis of Resulting Isolates

Deborah A. Sarkes[1], Justin P. Jahnke[1], Dimitra N. Stratis-Cullum[1]

[1]Sensors and Electron Devices Directorate, US Army Research Laboratory

Correspondence to: Deborah A. Sarkes at deborah.a.sarkes.civ@mail.mil

Abstract

Biopanning bacterial display libraries is a proven technique for peptide affinity reagent discovery for recognition of both biotic and abiotic targets. Peptide affinity reagents can be used for similar applications to antibodies, including sensing and therapeutics, but are more robust and able to perform in more extreme environments. Specific enrichment of peptide capture agents to a protein target of interest is enhanced using semi-automated sorting methods which improve binding and wash steps and therefore decrease the occurrence of false positive binders. A semi-automated sorting method is described herein for use with a commercial automated magnetic-activated cell sorting device with an unconstrained bacterial display sorting library expressing random 15-mer peptides. With slight modifications, these methods are extendable to other automated devices, other sorting libraries, and other organisms. A primary goal of this work is to provide a comprehensive methodology and expound the thought process applied in analyzing and minimizing the resulting pool of candidates. These techniques include analysis of on-cell binding using fluorescence-activated cell sorting (FACS), to assess affinity and specificity during sorting and in comparing individual candidates, and the analysis of peptide sequences to identify trends and consensus sequences for understanding and potentially improving the affinity to and specificity for the target of interest.

Video Link

The video component of this article can be found at https://www.jove.com/video/56061/

Introduction

Biopanning is a proven technique for affinity reagent discovery, with bacterial display libraries a convenient source for peptide affinity reagents [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24]. Similarly to phage display [25,26] and yeast display [27,28], the peptides are exposed on the cell surface and are available to bind to both biotic and abiotic targets, with these interactions driven by both the peptide backbone and the unique properties of the amino acid side-chains [29]. In contrast to phage display, the bacteria themselves are self-replicating and contain all genetic material required for display without elution of bound phage and reinfection of cells. In contrast to yeast display, bacterial display is faster due to the rapid (approximately 20 min [30]) doubling time of *Escherichia coli*. The resulting peptide affinity reagents can be produced off-cell for use in a variety of sensing platforms [16,17,26] and have potential for use as living materials when displayed on-cell [2,31,32]. As compared to monoclonal antibodies for recognition of specific targets, peptides are much smaller and more flexible, and peptide display libraries can be easily manipulated at the DNA level to avoid specific amino acids, such as cysteine [33]. Peptides are increasingly exploited for therapeutics [34,35,36] and other applications where antibodies would typically be utilized due to their ease of discovery, reproducible synthetic (off-cell) production, and superior maintenance of performance in extreme environments, providing a functional reagent even after exposure to high temperature or long-term storage without refrigeration [37,38]. The ability to overcome the cold chain for storage of reagents is of particular interest to the Department of Defense, for instance, which must operate in extreme environments. Thermal stability was therefore a desirable feature for the affinity reagents recognizing the chosen targets given as examples in this manuscript.

Recently, biopanning bacterial display libraries has become even more straightforward and reliable with the use of semi-automated magnetic-activated cell sorting (MACS or MCS) methods [9,14,39]. These methods have been demonstrated for biological targets using several similar sorting platforms with different advantages, including a commercially available automated magnetic-activated cell sorting (autoMACS or autoMCS) instrument (see table of materials) [1,14,15] and more specialized devices which enclose the sample in a disposable cartridge [7,9] or chip [39], a valuable specification for use with organisms or toxins that are Biosafety Level 2 (BSL-2) or higher [7]. These semi-automated methods could be extended to sort for peptides able to bind to abiotic materials if the materials are magnetic or have the ability to be coated onto a magnetic bead, and for use with different organisms (such as anaerobic bacteria) if the sorting and growth conditions are altered. In addition to sorting bacterial display libraries, the commercial autoMCS instrument used here has also been used in part for the initial steps of discovery of human single-chain antibody fragments displayed on the surface of yeast, followed by further isolation using Fluorescent Activated Cell Sorting (FACS) [40]. The primary advantages to semi-automation are decreased sorting time and effort, and more robust and reproducible elimination of unbound cells from the magnetic beads during wash steps, leading to reduced false positives, fewer required sorting rounds, and less complicated down-stream analysis [9,14]. In the case of the commercial autoMCS instrument, there are multiple pre-loaded programs which may be optimal for

different applications, such as negative pre-sorting (depletion), prioritizing purity, minimizing final sample volume, and increasing or decreasing sensitivity for isolation of rare cells [41].

The focus of this work is to demonstrate the methodology for biopanning a bacterial display library using the commercially available autoMCS instrument, and to expound the thought process applied in analyzing and minimizing the resulting pool of candidates in order to facilitate extension to other applications. The display library used here (see table of materials) [12,13] contains approximately 10^9- 10^{11} unique 15-mer peptides displayed via a modified version of the bacterial outer membrane protein OmpX in an unconstrained nature at the cell surface, which is expected to be representative of the free peptide in solution if produced synthetically. This protein scaffold additionally contains a positive control P2X peptide at the C-terminus for monitoring expression via binding to YPet Mona. However, a purchased or novel library with suitable diversity and other characteristics necessary for the specific application desired should work similarly with this biopanning procedure, though some minor changes may be required. A schematic of this biopanning protocol is illustrated in **Figure 1**. Additionally, the protocol could be adapted to other automated magnetic sorting platforms and other sorting strategies. Manual magnetic sorting with a benchtop magnet has been successfully demonstrated, albeit with more complicated and time consuming downstream analysis required due to increased false positives [14], and nonetheless provides an alternative option to the autoMCS steps if no automated magnetic cell sorting device is available.

Protocol

1. Biopanning Bacterial Display Libraries Using autoMCS

NOTE: This autoMCS-based sorting protocol has been previously described [14], and a more thorough "Expanded Protocol for New Users" is available in the supplementary materials for this manuscript for further detail, including an explanation of potential modifications. If an autoMCS device is unavailable for sorting, see the supplemental protocol from Sarkes *et al.* 2015 since a manual sorting protocol is also described in that work [14]. The autoMCS protocol provided here has been further adapted to include additional negative sorting steps for improved specificity [1,15]. A schematic of this biopanning protocol is illustrated in **Figure 1**.

1. **Negative Sorting to Remove Binders Likely to Cross-React with Magnetic Beads**
   1. Inoculate 500 mL of Luria Broth Miller (LB) containing appropriate antibiotic with approximately 1×10^{11} cells of a diverse bacterial display sorting library (see table of materials).
      NOTE: The bacterial display peptide library used here (see table of materials) contains approximately 10^9- 10^{11} unique members [8,12] and is grown in LB containing 25 μg/mL chloramphenicol (LB Cm$_{25}$). The remainder of this protocol will assume that this library is used.
   2. Incubate at 37 °C with shaking at 225 RPM until the culture reaches an OD$_{600}$ of 0.5 - 0.55. Induce peptide expression with 0.04% w/v L-arabinose by diluting a 4% stock 1:100, shaking at 225 RPM for 45 min at 37°C.
   3. Place induced culture on ice. Centrifuge approximately 2×10^{11} cells at 6000 x g for 20 min at 4°C. Remove supernatant and resuspend cells in 1.5 mL PBS by gently swirling.
      NOTE: An OD$_{600}$ of 1.0 approximately equals 1×10^9 cells/mL for *E. coli*.
      NOTE: <u>DO NOT VORTEX</u> as this may lyse the cells; resuspend by hand or by briefly shaking at ≤ 225 RPM, 4°C.
   4. Transfer cells to microcentrifuge tube(s) and spin at 6,000 x g for 5 min at RT or 4°C. Remove supernatant. Resuspend the cell pellet in 1 mL phosphate buffered saline (PBS) containing 0.5% bovine serum albumin (BSA; PBS-B).
   5. Wash 300 μL of streptavidin-coated beads (about 3×10^9 beads, see table of materials) in 1 mL PBS-B and centrifuge for 5 min at ≥5,000 x g (at RT). Place tube in a bench top magnetic particle separator. Carefully remove supernatant, avoiding the pellet.
   6. Resuspend the beads in the cell plus PBS-B mixture prepared above. Incubate at 4°C on a rotating platform for 45 min. Place samples on ice.
   7. Turn on autoMCS instrument to initialize the system. Prime lines using manufacturer's run and wash buffers (see table of materials) by selecting "Wash Now" at the bottom of the screen in the "Separation" menu, then "Rinse" and "Run".
   8. Prime the system with PBS-B, using PBS-B as both Wash Buffer and Running Buffer in the next steps. Select the "Separation" menu from the upper navigation bar and select "Wash Now". Select "Rinse" and "Run" to prime the system with fresh PBS-B.
   9. Transfer cell and bead mixture from the incubation step to a 15 mL conical tube. Place this tube in the sample slot, and empty 15 mL conical tubes in the positive and negative selection slots, of a pre-chilled (4°C) rack (see table of materials). Place rack on instrument platform.
   10. Assign a separation program for each sample on the rack (up to 5 samples can be sorted in one run). Add a "Rinse" step in between each sample and after the last sample. Select "Run" to start the cell separation. Select "OK" to confirm that enough buffer is available.
       NOTE: The "Posselds" separation program works well for negative sorting and positive sorting round 1 and "Posseld" is preferred for sorting rounds 2 - 4, but both of these programs have been successfully used for all negative and positive sorting rounds [14,15] and other programs may prove better for a particular application (described further in discussion).
   11. When the program is complete, remove the rack and retain the appropriate fraction. Here, for a negative sort, retain negative fractions (containing cells without beads). Place on ice.
   12. Use the negative sort fraction in its entirety to inoculate 1L of LB Cm$_{25}$ with 0.2% w/v D-glucose. Grow overnight at 37 °C, shaking at 225 RPM.
   13. Before turning off the autoMCS instrument, change the run and wash buffers to the manufacturer's run and wash buffers (see table of materials). Select "Wash Now", then "Rinse" and "Run". When finished, be sure there is 70% ethanol in the decontamination line, then press the "power icon" at the top right corner of the screen and select "Yes".
   14. When the system shutdown is complete (the bottles will be purple), turn off the machine.
   15. The next day, use the overnight culture to make freezer stocks with about 1×10^{11} cells per vial in LB with 15% glycerol. Additionally, or alternatively, use this culture in the next step: additional negative sorting (section 1.2) or the first round of positive sorting (section 1.3).

2. **Negative Sorting to Remove Binders Likely to Cross-React with Other Targets of Interest**
   NOTE: Section 1.2 can be skipped if no further negative sorting is needed or desired. In that case, proceed to section 1.3. Residual binding to any undesired targets can be assessed by FACS as in section 2: FACS Analysis of Sorting Rounds.

1. For negative sorting against a specific protein target, such as a similar protein with potential to also bind to the discovered peptides [1,15], repeat growth and induction steps 1.1.1 - 1.1.3, beginning with a frozen stock of streptavidin-depleted library from step 1.1.15 or the overnight culture from step 1.1.12 (using OD_{600} to estimate and inoculate with 1 x 10^{11} cells).
2. Transfer cells to microcentrifuge tube(s) and spin at 6,000 x g for 5 min at RT or 4°C. Remove supernatant. Resuspend cell pellet in 1 mL PBS containing 600 nM biotinylated cross-reactive protein target. Incubate at 4 °C for 45 min on a rotating platform.
   NOTE: 600 nM is a suggested starting concentration, which can be altered if a noted benefit is observed. Biotinylation of protein target can be achieved and quantified using the reagents suggested in the table of materials.
3. Meanwhile, wash 100 µL of streptavidin-coated beads (about 1 x 10^9 beads) in 1 mL PBS-B and centrifuge for 5 min at ≥ 5,000 x g (at RT). Place tube in a bench top magnetic particle separator and carefully remove supernatant, avoiding the pellet.
4. Centrifuge the target-bound cells at 6,000 x g for 5 min (RT or 4°C), remove supernatant, and resuspend the cell pellet in 1 mL PBS-B. Resuspend beads in entire volume of washed cells bound to cross-reactive protein target. Incubate at 4°C on a rotating platform for 30 min.
5. Place sample on ice and complete steps 1.1.7 - 1.1.15 for a negative sort, making appropriate freezer stocks. Continue to section 1.3 for a positive sort if all desired negative sorting has been achieved, or repeat section 1.2 to negative sort against another potential cross-reactive protein.

3. **Round 1 Positive Sort**
   NOTE: Round 1 positive sorting against a specific protein target is similar to a negative sort against a specific sorting target (see 1.2) except that the bead-containing, positive fraction is retained.
   1. Inoculate 500 mL of LB Cm$_{25}$ with approximately 1 x 10^{11} cells of a diverse bacterial display sorting library that has been streptavidin-depleted and grown overnight (from step 1.1.12) or frozen (from step 1.1.15) and thawed on ice, or that has been further depleted of binders to other cross-reactive protein targets (from step 1.2.5), as desired.
   2. Incubate at 37 °C with shaking at 225 RPM until the culture reaches an OD_{600} of 0.5 - 0.55. Induce peptide expression with 0.04% w/v L-arabinose by diluting a 4% stock 1:100, shaking at 225 RPM for 45 min at 37°C.
   3. Place induced culture on ice. Centrifuge approximately 2 x 10^{11} cells at 6,000 x g for 20 min at 4°C. Remove supernatant and resuspend cells in 1.5 mL PBS by gently swirling (by hand or by briefly shaking at ≤ 225 RPM, 4°C).
   4. Transfer cells to microcentrifuge tube(s) and spin at 6,000 x g for 5 min at RT or 4°C. Remove supernatant.
   5. Resuspend cell pellet in 1 mL PBS containing 600 nM biotinylated protein target of interest. Incubate at 4 °C for 45 min on a rotating platform.
      NOTE: 600 nM is a suggested starting concentration, which can be altered if a noted benefit is observed. Biotinylation of protein target can be achieved and quantified using the reagents suggested in the table of materials.
   6. Meanwhile, wash 100 µL of streptavidin-coated beads (about 1 x 10^9 beads) in 1 mL PBS-B and centrifuge for 5 min at ≥ 5,000 x g (at RT). Place tube in a bench top magnetic particle separator, and carefully remove the supernatant, avoiding the pellet.
   7. When incubation with target is complete, centrifuge the target-bound cells at 6,000 x g for 5 min (RT or 4°C) to remove any unbound target protein. Remove supernatant and resuspend the cell pellet in 1 mL PBS-B. Resuspend beads in entire volume of washed cells bound to protein target. Incubate at 4°C on a rotating platform for 30 min. Place on ice.
   8. Turn on autoMCS instrument to initialize the system. Prime lines using manufacturer's run and wash buffers (see table of materials) by selecting "Wash Now" at the bottom of the screen in the Separation menu, then "Rinse" and "Run".
   9. Prime the system with PBS-B, using PBS-B as both Wash Buffer and Running Buffer in the next steps. Select the Separation menu from the upper navigation bar and select Wash Now. Select Rinse and Run to prime the system with fresh PBS-B.
   10. Transfer cell and bead mixture from incubation step above to a 15 mL conical tube, rinsing the tube with an additional 500 µl of PBS-B and pooling the wash with the sample. Place this tube in the sample slot, and empty 15 mL conical tubes in the positive and negative selection slots, of a pre-chilled (4°C) rack (see table of materials). Place rack on instrument platform.
   11. Assign a separation program for each sample on the rack (up to 5 samples can be sorted in one run). Add a "Rinse" step in between each sample and after the last sample. Select "Run" to start the cell separation. Select "OK" or "CONTINUE" to confirm that enough buffer is available.
       NOTE: The "Posselds" separation program works well for negative sorting and positive sorting round 1 and "PosseId" is preferred for sorting rounds 2 - 4, but both of these programs have been successfully used for all negative and positive sorting rounds [14,15] and other programs may prove better for a particular application (described further in discussion).
   12. When the program is complete, remove the rack and retain the appropriate fraction. Here, for a positive sort, retain positive fractions (containing cells and beads). Place on ice.
   13. Before turning off the autoMCS instrument, change the run and wash buffers to the manufacturer's run and wash buffers (see table of materials). Select "Wash Now", then "Rinse" and "Run". When finished, be sure there is 70% ethanol in the decontamination line, then press the "power icon" at the top right corner of the screen and select "Yes".
   14. When the system shutdown is complete (the bottles will be purple), turn off the machine.
   15. Use the positive sort fraction in its entirety to inoculate 1L of LB Cm$_{25}$ with 0.2% w/v D-glucose. Grow overnight at 37 °C, shaking at 225 RPM.
   16. The next day, use the overnight culture to make freezer stocks in LB containing 15% glycerol, and/or continue to positive sorting round 2 in section 1.4.

4. **Subsequent Positive Sorting Rounds**
   NOTE: Typically, four sorting rounds are recommended, although three sorting rounds are generally sufficient [14,15]. When to stop sorting is assisted by the FACS Analysis of Sorting Rounds described in section 2. Concentrations of target and magnetic bead volume decrease with each subsequent positive sorting round.
   1. Using the overnight culture from the previous sorting round (or a frozen cell stock from that round), inoculate 5 mL LB Cm$_{25}$ with 100 µl cells (1:50 dilution). Incubate at 37°C with shaking at 225 RPM until the culture reaches an OD_{600} of 0.5 - 0.55.
   2. Induce peptide expression with 0.04% w/v L-arabinose, shaking at 225 RPM for 45 min at 37°C. Place induced cells on ice.
      NOTE: This induced culture can also be used to assess binding affinity and specificity for the sorting round it originated from, as described in section 2 below.

3.   Centrifuge 1×10^8 cells at 6,000 x g for 5 min at RT or 4 °C. Remove supernatant and resuspend cell pellet in 50 μL PBS containing half the concentration of biotinylated protein target of interest used for the previous round of sorting (therefore 300 nM for round 2, 150 nM for round 3, and 75 nM for round 4 from our suggested starting point). Incubate for 45 min on ice (or at 4 °C a rotating platform).

4.   Meanwhile, wash streptavidin-coated beads (15 μL for round 2, 8 μL for round 3, and 4 μL for round 4) in 1 mL PBS-B and centrifuge for 5 min at ≥ 5,000 x g (at RT). Place tube in a bench top magnetic particle separator and carefully remove supernatant, avoiding the pellet. Resuspend beads in 50 μL PBS-B.

5.   After the 45 min incubation with target in step 1.4.3, centrifuge the cells at 6,000 x g for 5 min at RT or 4°C, remove the supernatant, and resuspend the cells in the 50 μL washed beads from 1.4.4. Keep sample on ice and complete steps 1.3.7 - 1.3.13 for a positive sort.

6.   Inoculate a 5 mL culture of LB Cm_{25} supplemented with 0.2% w/v D-glucose with the entire positive, bead-containing fraction from the sort.

7.   The next day, use the overnight culture to make freezer stocks in LB containing 15% glycerol and/or continue to the next positive sorting round, returning to step 1.4.1.
     NOTE: Using the induced culture from 1.4.2 to assess binding affinity and specificity as described in section 2 below can help determine when to stop sorting. If two sorting rounds in a row show similar binding affinity, or a later round shows decreased binding affinity for the target of interest, this a desirable place to stop sorting. After the final round, return to step 1.4.1 but stop at 1.4.2.

2. FACS Analysis of Sorting Rounds

1.   Using the induced cells from step 1.4.2 for each sorting round, or identical cultures, add 5 μL induced cells to 25 μL of each of the following prepared solutions for binding assessment (see also table of materials): PBS alone; PBS with 150 nM YPet Mona (YPet [12,13]; positive control for expression), if available; 900 nM protein target conjugated to a fluorescent dye, such as amine-reactive dye with emission/excitation at 493 nm/518 nm (Target-488); 900 nM cross-reactive protein target(s) used for negative sorting labeled with the same fluorescent dye (Cross-Reactive Protein-488); and 900 nM streptavidin-R-phycoerythrin (SAPE). Incubate on ice for 45 min.
     NOTE: If continuing directly from step 1.4.2, this step will always be done the day after the biopanning for that round was completed since the down-selected library was grown overnight then subcultured and induced the following day. Cultures grown and induced similarly from frozen sorting round stocks can also be used; in that case, all rounds can be tested and compared in a single experiment.

2.   Centrifuge cells at 6,000 x g for 5 min at RT or 4°C. Remove supernatant and store samples on ice.
     NOTE: Cell pellets can be stored on ice until all samples are ready for analysis.

3.   Turn on the FACS instrument, open the software, start-up the system, and calibrate instrument following manufacturer's instructions [43,44].

4.   Within the FACS software, click on "Administrator" and click on desired folder or create a "New Folder." Click on the "New Experiment" icon at the top left of the screen. Right click icon to "rename" experiment. Within that experiment, click on "Global Worksheets".

5.   Click on the "Dot Plot" icon, then click on the global worksheet spreadsheet to create this scatterplot. Click on the dot plot graph itself, then select the "Inspector" icon at the top left of the screen and adjust the axes to biexponential display by selecting the boxes next to "Y Axis" and "X Axis" in the open window. Close the biexponential display window by clicking the X.

6.   Create a "New Tube" by clicking the icon at the top left of the screen and rename the specimen with appropriate information by right clicking the specimen icon under "global worksheet" and selecting "rename". Expand the specimen by clicking the (+) sign, then double click on the tube icon, right click on it, and again select "rename" to describe the sample.

7.   Use this dot plot graph with default SSC-A vs FSC-A to run a negative control sample in PBS alone by double clicking on that tube or selecting the arrow next to it. Just before running the sample, resuspend the cell pellet in 500 μL ice cold FACS running buffer and transfer sample to a FACS tube (see table of materials), mixing well by pipetting and flicking the tube. Place the resuspended cells on the sample injection tube (SIT) of the instrument. Press "Acquire Data" with "Events to Display" at 30,000 - 50,000 and "Events to Record" at 10,000, flow rate at low (to start; increase if needed), and "SIT flush" selected.
     NOTE: Appropriate speed (low, medium, or high) is the speed at which the number of events/s is approximately between 200 and 2000. Use this range for all steps.

8.   While acquiring, adjust photomultiplier tube (PMT) voltage threshold if needed by selecting the "Threshold" tab. Click on and change the "Value". Adjust voltage for forward and side scatter (FSC and SSC) such that negative control cells in PBS alone fall slightly below the center.
     NOTE: Additional parameters (such as SSC) can be added in the "Threshold" tab using the "add" icon. Typically, PMT voltages of about 700 V to 1000V for both SSC and FSC work well for *E. coli*.

9.   If the sample is reading properly, adjust the flow rate to display 200-2000 events/s and press "Record Data". When finished recording, remove tube from SIT and place on ice. Choose the "Polygon Gate" icon and use the mouse to draw a gate around the majority of the cell population in the scatterplot. Right click on dot plot and select "Show Population Hierarchy".

10. Use this "P1" gate as a parent by clicking on "P1" in the population hierarchy. Create a new dot plot following step 2.5. Right click each axis and change the Y-axis to FITC-A and the X-axis to FSC-A. Gate the negative control (PBS alone) using the "polygon gate" icon, with the gate as tight as possible around the top and left side of the population.
     NOTE: Double check the population hierarchy to ensure that the P1 gate is a parent of the P2 gate. If it is not, it will appear in line with the P1 gate and "All Events" will be the parent; delete the gate and re-create it following the step 2.10.

11. Select "P2" in the population hierarchy, right click, and select "invert gate". Right click on the dot plot with P2 gate and select "Show Populations". Select the populations "P2" and "Not P2" for display (less than 1% of the population should fall outside the gate).

12. Right click the dot plot with P2 gate and select "Create Statistics View". Right click the statistics view window and select "Edit Statistics View". Click the "Populations" tab and add or remove populations, as desired, including the parent population, P1 (P2, and Not P2 should already be shown). Click on the "Statistics" tab and add "FITC-A Median" and any other statistics of interest. Close the window by pressing OK.

13. Create a similar dot plot graph for PE-A vs FSC-A and gate using PBS alone sample (previous plots can be duplicated and altered). Use this spreadsheet to run all samples (including negative control cells incubated with each fluorophore-labeled target) in this manner, recording about 10,000 events for each.
     NOTE: The cells that fall outside the gate after incubation with the Target-488 protein, the YPet positive control, etc. are binders to that protein, and the value from "Not PX" (where X is the number of the gated population for that graph) should be recorded as % bound. When using SAPE as a negative control, use the PE-A vs FSC-A plot. The median fluorescence intensity (MFI) of P1 should also be recorded as

this gives information beyond % binding, to show the extent of binding in relative terms, and is more robust in the presence of outliers [45]. Median fluorescence intensity can be normalized (nMFI) by dividing MFI of each peptide by the MFI of a scaffold only negative control (with no N-terminal peptide), or a clone containing a peptide sequence that does not bind the target of interest, after incubation with the same target conjugated to the same fluorophore [14]. If these negative control cells are unavailable, uninduced cells incubated with the same target and labeled with the same fluorophore can also be used for normalization.

3. Sequence Analysis of Potential Candidates and Assessment of Binding Affinity

1. **Peptide Sequence Determination**
    1. Select and sequence tens or hundreds of bacterial colonies from the final round(s) of biopanning (typically rounds 3 and/or 4 are sufficient [14]).
    2. Analyze the sequence data using the macro file that we have developed (see supporting information for code, "Sub eCPX_Sequencing") to specifically analyze the sequences generated from biopanning the 15mer bacterial display library listed in the table of materials. Use the spreadsheet software listed in the table of materials, or other preferred compatible software.
        NOTE: A number of tools are available online that can also aid in the DNA sequence analysis [47]. If preferred, use a different established method for sequence analysis and skip to step 3.1.3.
        1. Download the set of sequence files (.seq files, text documents also work) and extract them into a new folder. If necessary, move or copy the folder to the computer's hard drive (as opposed to a network folder) for improved speed.
        2. Open a new spreadsheet window. Make sure to enable macros and enable all features, if those messages pop up.
            NOTE: Steps 3-6 below should only need to be completed the first time the macro is used. For subsequent analysis, simply "Run the Macro" starting at step 7.
        3. Copy the entire contents of the macro found in the "Sub eCPX_Sequencing" File.
            NOTE: The current version is set up with flanking sequences for the 15-mer library listed in the table of materials and will translate the amino acids between them. Additional sequences can be inputted by copying 5' flanking sequences in column A and 3' flanking sequences in column B of the spreadsheet, up to 10 sequences for each, before running the macro.
        4. In the blank spreadsheet file, select the "View" tab, then double click on "Macros." Select the personal.xlb folder. If this is unavailable, record a macro first to make this appear.
        5. Click "create" or "step into", depending on what can be highlighted. Paste the entire macro into the module.
        6. Click the save icon or go to file, save. Exit the module. If a window pops up, press OK.
        7. Run the Macro: Go to the folder where the desired.seq files are located. Click on the bar next to the icon of the folder at the top to view the folder location. Copy it.
        8. Go to the new spreadsheet window. Select the view tab, then double click on macros. Select "eCPX_Sequencing" from the list and click "run."
        9. In the box that pops up, paste the file location copied in step 7 and hit enter. The macro should start going through the sequences and organizing them into tables on various sheets.
        10. Use the "Summary Table" sheet to determine if it will be necessary to check trace files for errors and make corrections manually (if sequences contain "X" or could not be translated).
            NOTE: The "Summary Table" displays the translated peptide sequences, sorted by amino acid sequence (from A to Z), unless a sequencing error prevented translation. An "X" denotes an individual amino acid that could not be determined.
    3. Once the sequences are translated, organized, and any sequencing errors corrected, check the sorted peptide sequence list on the "Summary Table" sheet for any repeating sequences.
    4. Grow overnight cultures for sequenced colonies of interest (including the repeats and/or peptides with noticeable trends at a minimum) in 5 mL LB Cm$_{25}$ at 37°C, shaking at 225 RPM, and save freezer stocks the next day in LB containing 15% glycerol, as in step 1.4.7. Test the peptides with repeating sequences for binding affinity and specificity using the FACS methods described in section 3.3 and use the FASTA format of the sequence list (full list and repeats only) to align the sequences using preferred alignment and analysis programs, as in section 3.2.

2. **Sequence Alignment Using Clustal Omega, Kalign, and Similar Programs**
    1. Copy sequences in FASTA format from the FASTA spreadsheet of the macro, or otherwise create a list of FASTA files from the sequences to be aligned (such as those from step 3.1.3).
        NOTE: Column A contains the FASTA files in numerical order by "Seq #" while column B displays them sorted by "AA sequence" from the "Summary Table" sheet. The list of peptide sequences sorted by amino acid sequence will display unusable sequences at the top, such as empty vector (as "# Empty") and those sequences in which neither the 5' or 3' search criteria were found (as "# Value"). This feature allows for easy update or removal of those sequences from the analysis.
    2. Open Clustal Omega[48] or Kalign[49] software by going to the website [50,51], or use other preferred software for sequence alignment. Copy and paste the FASTA sequence list and input it into the box below "sequences in any supported format". Change "gap open penalty" to 30 under "more options" in Kalign (this is not possible in Clustal Omega), but otherwise keep default settings before clicking "submit".
    3. Analyze sequence alignment using Jalview [52,53] software directly within Clustal Omega and Kalign online software by selecting the "result summary" tab above the alignment and then clicking the "Jalview" icon.
        NOTE: If preferred, or for analysis of sequence alignments generated by alternative programs, download [54] the software separately and copy and paste the alignment into the analysis window, which is opened by clicking "File", "Input Alignment", and "from Textbox". This provides a means to easily determine if a consensus sequence is present in the input sequence alignment.

3. **Comparing Binding Affinity and Specificity Using FACS**
    1. Inoculate 5 mL LB Cm$_{25}$ supplemented with 0.2% w/v D-glucose with each individual isolate of interest (from step 3.1.4 and/or colonies of interest from further sequence analysis in section 3.2, *etc.*) and an appropriate negative control (display scaffold only or a peptide that does not bind target, as described in the note for step 2.13). Grow overnight at 37 °C, shaking at 225 RPM.

2. Use the overnight cultures to inoculate 3 mL LB Cm$_{25}$ with 60 μl cells (1:50 dilution). Incubate at 37°C with shaking at 225 RPM until the culture reaches an OD$_{600}$ of 0.5 - 0.55. Induce peptide expression with 0.04% w/v L-arabinose plus 2 mM EDTA (for facilitation of peptide display [42]), shaking at 225 RPM for 45 min at 37°C.

3. Place induced cultures on ice. For each isolate, add 5 μL cells to 25 μL PBS alone or PBS containing: 150 nM YPet [12,13] (positive control for expression), if available; 250 nM Target-488; 250 nM Cross-Reactive Protein(s)-488; and 250 nM SAPE (see 2.1 for more detail). Incubate on ice for 45 min.

4. Centrifuge cells at 6,000 x g for 5 min at RT or 4°C. Carefully remove supernatant by drawing liquid from the opposite side of the pellet. Store cell pellets on ice until all samples and are ready for analysis.

5. Resuspend each cell pellet in 500 μL ice cold FACS running buffer, mixing well by pipetting and flicking the tube, just before reading using a flow cytometer (see table of materials).
   NOTE: See step 2.13 notes for further explanation of gating and calculation of nMFI to compare binding affinity and specificity. Non-binders or non-specific binders can now be removed from further analysis, and the highest affinity binders should be re-assessed by sequence alignment following section 3.2 to look for further trends that may have been missed in the initial sequenced population. Sections 3.2 and 3.3 may be repeated as needed as new trends and consensus sequences are noted. This analysis can include more random sequences if trends are not seen.

Representative Results

With the use of automated magnetic-activated cell sorting (autoMCS) rather than manual magnetic cell sorting, false negative candidates are significantly reduced during biopanning of bacterial display libraries [9,14]. Negative sorting steps can be also be added as needed to reduce non-specific binders to the target of interest, and it is suggested at a minimum to add a negative sort against the magnetic beads themselves up front. This negative selection against the magnetic beads themselves was completed before four rounds of biopanning the chosen bacterial display library for protective antigen (PA) binders, as shown in **Figure 2**, and before additional negative selection against rivax and four rounds of positive selection for abrax, as shown in **Figure 3**. The beads used here are conjugated to streptavidin for capture of biotinylated proteins, so unintended isolation of peptides binding to streptavidin itself is a concern and is monitored using FACS with streptavidin-conjugated to Phycoerythrin (**Figure 3, SAPE**). At a minimum, binding to streptavidin should be tested for any promising individual candidates when assessing binding to the target protein. Both the percent binding and nMFI for SAPE should be low values in all sorting rounds. The level of binding to streptavidin can increase somewhat with each sorting round, as occurred in **Figure 3**, because the population is repeatedly exposed to the streptavidin-coated magnetic beads after each sorting round. If the level of background streptavidin binding becomes problematic to downstream analysis, further negative sorting against the magnetic beads could be performed following the positive sorting round in which the streptavidin binding population increased in order to reduce downstream screening of false positives. When proteins or materials are available which could specifically interfere with downstream use of the peptide for a specific application, or which are expected to cross-react with affinity reagents for the target due to structural and/or sequence similarity, further negative sorting against that protein or material is recommended. An example is negative sorting against rivax before isolation of abrax binding peptides, due to the structural similarity and sequence homology of these two proteins [1,15]. Again, cross-reactive binding to proteins used for negative sorting steps can be monitored during the analysis of sorting rounds using FACS (**Figure 3, Rivax-488**). In the best case scenario, percent binding and nMFI are low, as in this example.

When available, a fixed positive control peptide, such as the P2X peptide at the C-terminus of the display scaffold produced by the bacterial display library used in the representative results here, can help determine if lack of binding affinity in the FACS assay is due to lack of expression of the display scaffold itself, rather than lack of affinity of the peptide(s) for the target. In **Figure 2** and **Figure 3**, YPet Mona binding to this P2X peptide is monitored and demonstrates that the early rounds of sorting express the scaffold poorly. This is likely predominantly due to the high frequency of stop codons in N-terminal peptides in the random library, which means that the scaffold itself was not produced. Other effects may additionally contribute, such as the presence of peptides with unexpected toxic effects to the *E. coli*, or decreased growth or peptide display rates for other reasons (such as mutations in the bacterial genome). Addition of EDTA during induction may improve expression during these early rounds of sorting [42], but has not yet been tested. Binding to YPet Mona in each biopanning population is improved significantly by round 3. Comparing **Figure 2** to **Figure 3**, it is also apparent that expression level can be improved by increasing induction time to 90 min (as in **Figure 2, YPet Mona**) rather than 45 min (as in **Figure 3, YPet Mona**), although 45 min is typically sufficient. Note that the nMFI for YPet Mona is normalized to the YPet Mona binding level of the negative control cells, so values near 1.0 are typical if expression level is acceptable. Normalizing YPet Mona MFI to PBS alone MFI from the same sample is also a valid way to compare relative expression levels, but gives much higher values (compare **Figure 2** and **Figure 3** with results from Sarkes *et al.* 2015 [15]).

Enrichment of the library for peptides binding to the specific target of interest is typically achieved within three positive sorting rounds, but continuing to a fourth round of sorting can be beneficial, as shown in **Figure 2** and **Figure 3**. Here, percent bound cells and nMFI continue to increase from round 3 to round 4 for both example targets, PA and abrax, which are boxed in red in **Figure 2** and **Figure 3**, respectively. The highest affinity peptide sequences may already be present in round 3 but are likely to further enrich in round 4 as well, which aids in down-selection of potential candidates [14]. Analysis of sequences from round 4 tends to reveal repeating sequences, and these repeating sequences are an excellent starting point for affinity and specificity analysis and consensus sequence determination. The eCPX_Sequencing macro provided as a supplement to this manuscript helps to streamline this initial analysis, as shown in **Figure 4**. Beginning with a folder of .seq or text files, the DNA sequence which lies between chosen 5' and 3' sequences is translated, organized by amino acid sequence, and analyzed for number of individual amino acid residues in each peptide. The sorted list of isolated peptide sequences, as shown on the Summary Table sheet screenshot in **Figure 4**, can be used to easily determine which sequences repeat and at what frequency. The cells in this spreadsheet containing repeating sequences are outlined in different colors for easier analysis. It is recommended to check these repeating sequences for consensus sequences and other trends. This is aided by sequence alignment software such as Kalign and Clustal Omega, and analysis can be performed on the entire sequence output (including both repeating and non-repeating sequences at the frequency they appeared) and on the down-selected list of repeating sequences (with or without their relative frequency). Representative results for PA and abrax repeating sequences, without taking frequency into account, are shown in **Figure 5A** and **5B**, respectively. Note that in **Figure 5A** for the PA target, several of the individual repeating sequences themselves contain the consensus sequence WFCFTC (or a similar sequence), underlined in red, which was determined by applying Jalview software analysis to a Kalign sequence alignment of the repeating sequences. This consensus, as WXCFTC, was previously determined to be a PA binding consensus, which provides confidence in the sorting method [9]. In **Figure 5B**, where repeating sequences for the abrax target were analyzed in the same manner, the result was quite different. First, there were only five sequences that repeated in round 4 of biopanning for abrax binders, as opposed to the fifteen repeating sequences isolated from round 4 of biopanning for PA binders (although 44% more colonies were also sequenced for PA). Second, none of the five repeating sequences contained the most promising "consensus" sequence (FWAWF, underlined in purple), although candidate AX-A15 contained the sequence that best matched (FWDTWF). Further analysis, including binding affinity and specificity determination, helped to provide the consensus of FWDTWF determined from the top five candidates for abrax binding in **Figure 5C**, as explained below.

A representative colony from each repeating sequence can be further analyzed for on-cell affinity and specificity using FACS, as in **Figure 6A** for the repeating sequences from biopanning for abrax binding peptides. Here it is clear that all five repeating sequences have higher affinity for abrax, the intended target, than rivax, the structurally similar protein used for negative sorting, or streptavidin, which is present during the sort due to the magnetic beads used for biopanning. This is consistent with the already promising results for affinity and specificity of the sorting rounds shown in **Figure 3**, where it is clear that enrichment for binding to the intended target, abrax, is increasing with each round of biopanning while affinity for the similar protein, rivax, is not. However, a satisfactory consensus sequence was not determined for the abrax sort using the repeating sequences from round 4 alone (**Figure 5B** and above discussion). Returning to the 100 sequenced colonies from round 4, however, it was noted by eye when searching for sequences similar to the best match for the predicted consensus that one additional candidate, AX-A12, contained the FWDTWF sequence that was noted in isolate AX-A15, and that another candidate, AX-A14, contained a similar sequence, DWNTWF. These and other sequences that either contained similarities to the repeating sequences, demonstrated similarities to other non-repeating sequences, or were chosen randomly, were analyzed by FACS for binding affinity and specificity. Those tested are ranked in Sarkes *et al.* 2016 [1] and the top 5 binders from this analysis are shown in **Figure 6B**, with the consensus sequence determined to be FWDTWF, shown in **Figure 5C**.

A similar analysis of binding affinity for repeating peptide sequences in round 4 of biopanning for peptides that recognize the PA target revealed that the best binders, as ranked by the ratio of PA-488 nMFI:SAPE nMFI, contained the WXCFTC consensus or a similar sequence (**Table 1**). Using this ratio, the sequences self-organized into those that contained the consensus and those that did not. The top candidates, all containing sequences related to the consensus, were analyzed similarly to the methods described above for abrax in **Figure 6**, as presented in Sarkes *et al.* 2015 [14]. These results demonstrate that for some targets, analysis of peptides with repeating sequences may be adequate to obtain binders with significant affinity and specificity for a chosen target, and to determine a consensus sequence that may be, in itself, sufficient for binding. Note, however, that the number of repeats does not necessarily reflect the relative binding affinity for the target of interest (**Table 1**). When analysis of repeating sequences alone is not sufficient to determine a pattern, it is helpful to alternate between sequence analysis and binding analysis to narrow down the pool of candidates and reexamine the trends among those candidates with higher affinity. Even when a trend is observed from analyzing the repeating sequences alone, additional non-repeating sequences may demonstrate the same trends, or other trends, upon further investigation.
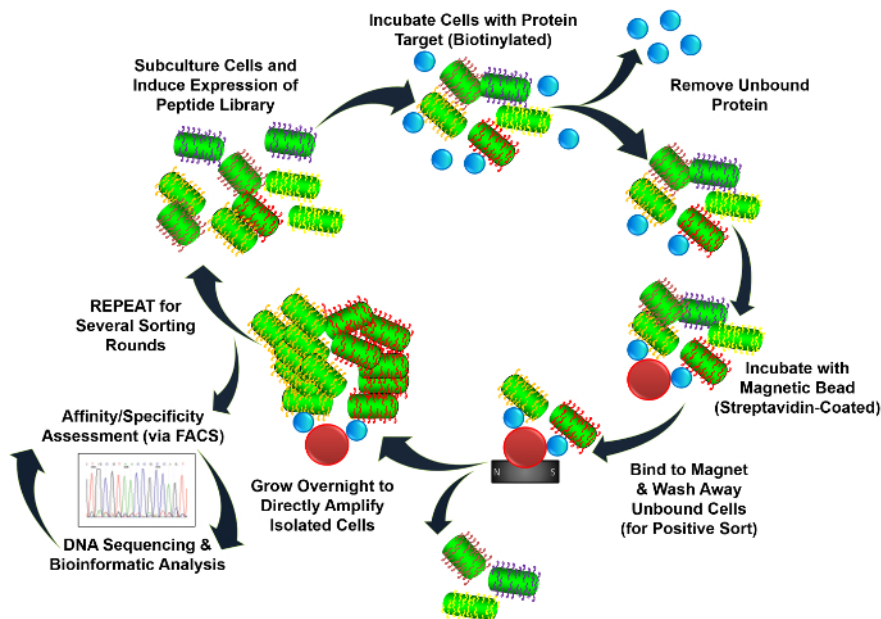
Figure 1: Schematic of biopanning protocol for bacterial display libraries. As illustrated here, biopanning bacterial display libraries is a cyclical process. Each cell in the bacterial display library (see table of materials) contains plasmid DNA which is retained as long as the cells are grown in the presence of the antibiotic for which the plasmid contains a resistance gene (choloramphenicol in this case). Each plasmid also encodes the display scaffold protein which exposes a random peptide to the exterior of the cell. Since each cell should only contain a single plasmid DNA sequence, each cell should only display a single peptide, in multiple locations throughout the cell membrane. Depending on level of library diversity at the start of biopanning and after each sorting round, the DNA sequence may or may not be unique from cell to cell. Biopanning begins with growth and induction of the cell library to display individual peptides on their outer membrane. These peptides can then interact with a target protein (which is biotinylated or otherwise tagged for capture). For magnetic-activated cell sorting (MCS), unbound protein is removed and the cells are incubated with magnetic beads that are coated with a capture protein (streptavidin in this example, which has a strong interaction with biotin). The entire culture is then exposed to a magnet to separate bound cells from unbound cells. Using an automated magnetic sorting device is preferred for cell separation to reduce false positives. Illustrated here is a positive sort, in which the cells that are bound to and co-elute with the magnetic beads are retained. For a negative sort, which we typically perform up front, the process is the same except that the cells that are washed off of the magnetic beads are retained instead. The library fraction of interest, bound (positive sort) or unbound (negative sort), is grown overnight and used in the next round of sorting. Each subsequent round of positive biopanning requires a decrease in target concentration and bead volume to improve stringency. After each round of biopanning, affinity and specificity of the peptides for the protein target are assessed using fluorescence-activated cell sorting (FACS) to help determine when to stop biopanning and start investigating individual candidates. As needed, DNA sequencing and peptide sequence analysis are performed. These analysis steps may be in themselves a cyclical process, particularly for revealing trends in the individual isolates after the final round of biopanning. At this point, peptide sequence trends and/ or consensus are correlated with binding affinity and specificity. Please click here to view a larger version of this figure.
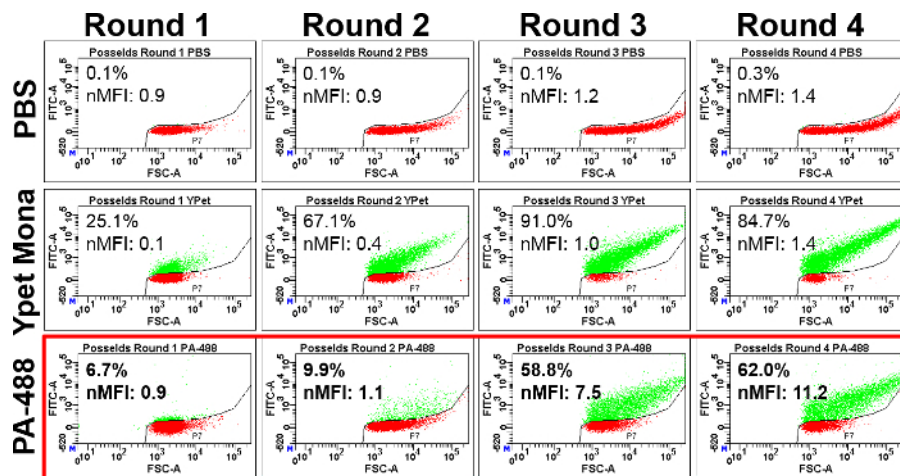
Figure 2: Example data for FACS analysis of sorting rounds: PA target. Binding affinity as determined by FACS is shown after 4 rounds of biopanning (positive sort) the chosen bacterial display library for peptides binding to the target, protective antigen (PA). This was performed after first performing a negative sort against the streptavidin-coated magnetic beads. For binding assessment, cells were induced with 0.4% L-arabinose for 90 min before incubation with target and control solutions and analyzing by FACS. Binding affinity analysis for the intended target is boxed in red. Shown here are scatter plots of FITC-A vs FSC-A and values for percent cells bound (as compared to the gated negative control incubated with PBS alone) and normalized median fluorescence intensity (nMFI, as compared to peptide-free negative control incubated with the same fluorophore-labeled protein) of induced cells that were incubated for 45 min with: PBS buffer alone, 150 nM YPet Mona (positive control for peptide expression), or 250 nM PA-488 (labeled target). Note the increasing enrichment in binding affinity for the target of interest after each round of biopanning. In contrast to **Figure 3**, all autoMCS cell separations were completed using program Posselds. Part of this data can also be visualized in scatter plots of FITC-H vs FSC-H in Sarkes *et al.* 2015 [14] for comparison to the FITC-A vs FSC-A plots shown here. Please click here to view a larger version of this figure.
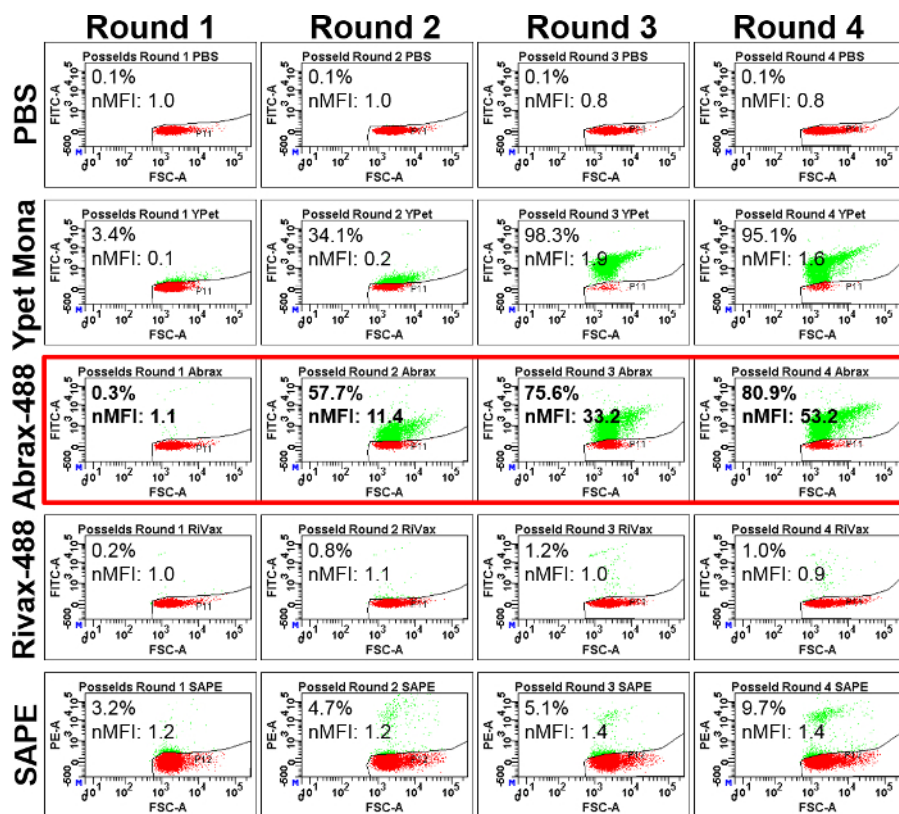
Figure 3: Example data for FACS analysis of sorting rounds: abrax target. Similarly to **Figure 2**, shown here is comparative binding affinity for a complete biopanning experiment, as determined by FACS. The bacterial display library was subjected to negative sorting against the streptavidin-coated magnetic beads, as in **Figure 2**, but was then subjected to an additional round of negative sorting against a homologous protein with similar structure and function, rivax, before performing 4 rounds of positive biopanning for peptides binding to the intended target, abrax. For binding assessment, cells were induced with 0.4% L-arabinose for 45 min before binding to target, positive control, and negative controls and reading on FACS. Binding affinity for the intended target is boxed in red. Shown here are scatter plots of FITC-A vs FSC-A (or PE-A vs FSC-A in the case of SAPE) and values for percent cells bound (as compared to the gated negative control incubated with PBS alone) and nMFI of induced cells that were incubated for 45 min with: PBS buffer alone, 150 nM YPet Mona (positive control for peptide expression), 250 nM Abrax-488 (labeled protein target), 250 nM Rivax-488 (labeled potential cross-reactive protein target), or 250 nM SAPE (negative control for direct binding to streptavidin-coated magnetic beads). Note the increasing enrichment in binding affinity for the target of interest, abrax, after each round of biopanning, and minimal binding affinity for the structurally similar protein, rivax. In contrast to **Figure 2**, positive biopanning round 1 was performed using the autoMCS Posselds program while subsequent sorting rounds were completed using program Posseld, as suggested for biopanning in this manuscript. This data can also be visualized in scatter plots of FITC-H vs FSC-H in Sarkes *et al.* 2016 [15] for comparison to the FITC-A vs FSC-A plots shown here. Please click here to view a larger version of this figure.

Figure 4: Example sequence analysis output using "eCPX_Sequencing" macro. Shown is a screenshot of 48 sequenced colonies from round 4 of biopanning the chosen bacterial display library for PA binders using the Posselds method. The "Summary Table" sheet is shown here. Note that the colonies were numbered in the alphabetical order of their given file name during the sequencing process and that the boxes surrounding sequences that repeat at least once are outlined in different colors for easier data analysis. The eCPX_Sequencing macro is provided as a supplemental code file. Please click here to view a larger version of this figure.
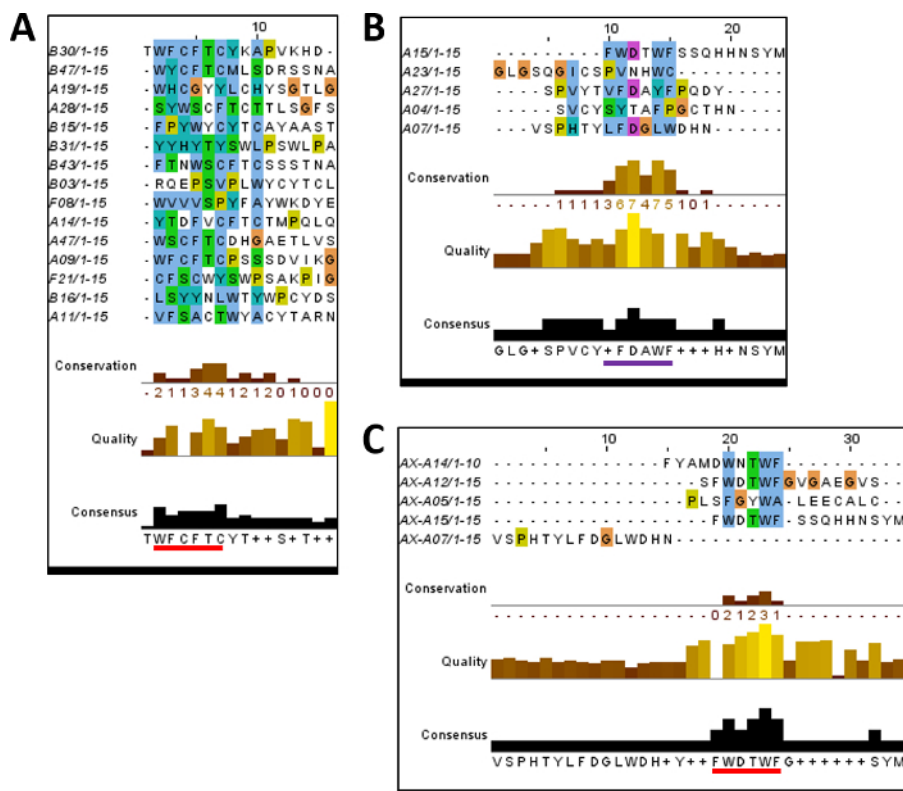
Figure 5: Sequence alignment and consensus determination for two distinct protein targets. All images shown here were generated using Jalview software with Clustal_X coloring after aligning sequences using Kalign online analysis software [51]. A) Alignment of repeating sequences from PA biopanning round 4 (corresponds with **Table 1**). B) Alignment of repeating sequences from abrax biopanning round 4 (corresponds with **Figure 6A**). C) Alignment of top 5 candidates from abrax biopanning round 4, as determined by FACS analysis of individual candidates (corresponds with **Figure 6B**). The red line underlines the consensus sequence in A and C, while the purple line in B underlines the precursor to the consensus sequence determined for abrax binding when examining repeating sequences only, highlighting the potential need to test binding affinity using FACS before a consensus can be determined in some cases. Similar alignments generated with different analysis software can be seen in Sarkes *et al.* 2015 [14] and Sarkes *et al.* 2016 [1]. Please click here to view a larger version of this figure.
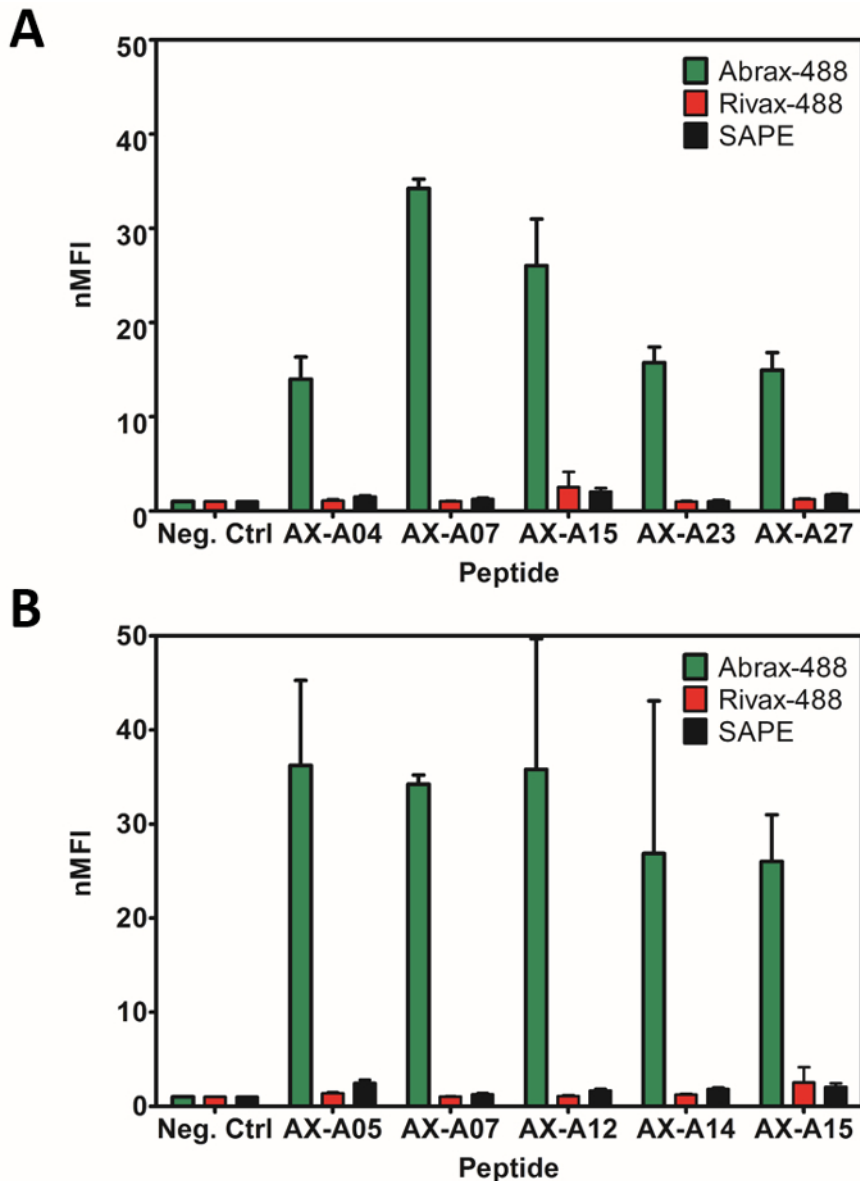
Figure 6: Example binding affinity and specificity for individual candidates analyzed using FACS. Shown here is the normalized median fluorescence intensity (nMFI) determined using FACS for A) repeating sequences and B) the top 5 candidates, as determined by comparative binding affinity for the target, abrax, from round 4 of biopanning. The data represents the average (bars) and standard deviation (error bars) of three independent replicate experiments. Note that all candidates shown are quite specific for abrax (abrax-488, green bars) over a structurally similar protein, rivax (Rivax-488, red bars), and the streptavidin negative control (SAPE, black bars). Although two of the repeating sequences in A (AX-07 and AX-15) were also among the top 5 candidates in B, comparison of results in A and B demonstrates that analysis of repeating sequences alone may not be enough to isolate the highest affinity peptides. Data shown here was adapted from Sarkes *et al*. 2016 [1]. Please click here to view a larger version of this figure.

Peptide Name	Peptide Sequence	# Repeats	PA nMFI: SAPE nMFI
A14	YTD**FVCFTC**TMPQLQ	2	21.1
B30	T**WFCFTC**YKAPVKHD	10	15.9
A28	SY**WSCFTC**TTLSGFS	4	11.9
B43	FTN**WSCFTC**SSSTNA	3	10.7
A47	**WSCFTC**DHGAETLVS	2	10.5
A09	**WFCFTC**PSSSDVIKG	2	9.2
F21	**CFSCWY**SWPSAKPIG	2	8.8
B47	**WYCFTC**MLSDRSSNA	5	7.3
B03	RQEPSVPL**WYCYTC**L	3	5.1
B15	FPY**WYCYTC**AYAAST	4	4.1
A19	WHCGYYLCHYSGTLG	5	3.6
F08	WVVVSPYFAYWKDYE	3	3.0
B16	LSYYNLWTYWPCYDS	2	1.4
A11	VFSACTWYACYTARN	2	1.0
B31	YYHYTYSWLPSWLPA	4	0.03

Table 1: Ratio of specific binding to non-specific binding for individual repeat candidates. In this example, the number of repeat sequences and the ratio of PA (target) nMFI to SAPE (negative control) nMFI is shown for the repeating candidate sequences from PA biopanning round 4. This data was sorted by relative PA nMFI:SAPE nMFI ratio and analyzed for the presence of the known WXCFTC consensus, or a similar sequence, which is shown in bold font and underlined. Notice that the sequences self-organize such that those candidates with the consensus have the highest PA nMFI:SAPE nMFI ratio, and therefore interact more specifically with the target. The results shown are from a single FACS experiment. Peptides with lowest affinity for the target were excluded from the additional replicates and analysis that were published in Sarkes *et al.* 2015 [14].

Suppplemental File 1: Please click here to download this file.

Suppplemental File 2: Please click here to download this file.

Discussion

Biopanning bacterial display libraries has been a successful approach to the isolation of affinity reagents, and peptides offer a useful alternative to antibodies for recognition when specific criteria are required, such as stability in extreme environments. Biopanning of these libraries has been streamlined using the autoMCS methods described here, and a commercially available autoMCS platform extends this technology to a much wider audience. In comparison to traditional alternating MCS/FACS sorting methods for bacterial display libraries, autoMCS methods are less expensive because they do not require investment in a more expensive FACS instrument capable of sorting and isolating the bound cells, rather than the type of flow cytometer that is used here, which is only capable of analysis [14]. The critical steps for successful biopanning by autoMCS include separation program selection, negative sorting against potential cross-reactants to reduce false positives, monitoring library enrichment using FACS to determine when to stop biopanning, and smart analysis of the candidate pool to avoid cumbersome screening of hundreds of candidates.

The examples described herein demonstrate successful biopanning of the chosen bacterial display library for two separate targets, PA and abrax, although with slightly different analysis strategies due to differences in the peptide sequences for each pool of candidates after four rounds of biopanning. PA was the more straightforward case, with sequence analysis demonstrating clear trends from the peptide sequences that repeated at least once in 144 colonies. This alone was enough to determine a consensus sequence for the PA target, and those colonies that contained the consensus or a similar sequence were the best candidates in terms of the ratio of binding to PA versus binding to the streptavidin negative control. However, this will not always be the case. For the abrax target, sequencing 100 colonies led to five repeating sequences but the trend in these sequences was less clear, other than a tendency to contain aromatic amino acid residues and aspartic acid or asparagine. The predicted "consensus" from the repeating sequences only could have been produced and tested for affinity to abrax, but since no parent sequences contained that exact consensus sequence, we returned to the list of sequenced colonies and observed that other candidates that had more trends in common with each other. In fact, two colonies contained a similar sequence to the "consensus" from the repeats alone, (FWDTWF instead of FWAWF), which had the same trend of tryptophan, phenylalanine, and aspartic acid residues but with different spacing. One of these peptides was a repeating sequence, one was not. These two sequences, along with a third sequence containing a similar variant, were among the top 5 binders tested in terms of affinity for abrax. The top binder overall, AX-A05, also displayed similar trends. The results for abrax demonstrate that an approach which alternates several times between sequence analysis and affinity and specificity analysis will sometimes be required, depending on the chosen target. The results for the abrax target also demonstrate the level of specificity that can be achieved over a very similar protein (rivax [1,15]).

The autoMCS programs selected for our studies were Posselds and Posseld, which are both for positive selection of labeled target cells with low frequency, less than 5% of the initial population. In both cases, cells pass through two magnetic columns. In the case of the Posselds program, however, cells pass more slowly through the first column to increase exposure time [41]. In addition to the program descriptions given by the manufacturer of the commercial autoMCS device (see **Materials Table**) [41], these programs were chosen after an initial comparison of several potential programs. Each program's ability to avoid pulling out false positives from a homogenous culture of negative control cells, and

to successfully isolate a known binder to a target of interest from a mixed culture containing negative control cells spiked with a decreasing percentage of the known binder, were compared. If significantly deviating from the applications described here, a similar comparison could be helpful in program selection for the new application. However, previously published results comparing these two programs (Posseld and Posselds) for isolation of peptide affinity reagents for PA showed that using either of these programs for all four rounds of biopanning led to isolation of peptides with similar affinity and specificity for PA and determination of the WXCFTC consensus. The main differences were that Posseld, with its faster flow rate, sorted more rapidly and binding affinity for the target was therefore higher after only two rounds of biopanning, while using Posselds led to more unique sequences after four rounds of biopanning [14]. Learning from this, the strategy was changed slightly when biopanning for abrax binding peptides to incorporate both benefits: Posselds was used for Round 1 to allow more exposure time during this critical step where diversity is highest, but then the program was switched to Posseld for faster isolation of the highest affinity binders during Rounds 2-4 [1,15]. This appeared to be ideal for abrax, especially since the nMFI and percent binding were both higher for abrax sorting round 4 as compared to PA sorting round 4 with either program (**Figure 2** and **Figure 3** and Sarkes *et al.* 2015 [14]), but a direct comparison using one strategy versus the other with the same target would be required for a more conclusive comparison. Which program is best for a particular application may depend on the individual target, the sorting library used, and the level of peptide expression that is achieved with any change to the conditions described here.

Not discussed here are potential results from biopanning with abiotic materials, but we have also used the same bacterial display library for biopanning against bulk aluminum [2]. This study was not performed using autoMCS, but similar studies could be accomplished using autoMCS if the material is available on, or could be coated on or conjugated to, a magnetic bead. In the bulk aluminum study, individual amino acid analysis and secondary structure modeling was helpful in understanding what was driving the binding affinity of the isolated peptides, since no consensus sequence was determined [2]. Even with biological targets, this may be required to understand what is driving the binding affinity for some targets, which is why the spreadsheet generated from the "eCPX_Sequencing" macro includes a tab with individual residue frequency analysis. If no repeating sequences are seen in addition to the lack of a consensus, and residue frequency shows no pattern, however, other types of analysis may be required. Additionally, further sorting rounds may be necessary to avoid screening a much larger number of candidates for down-selection of those with sufficient affinity to the desired target.

With the increasing availability of next generation sequencing, a more thorough study could be performed to determine the most promising candidates before testing affinity and to determine the extent of enrichment after each round of biopanning. However, sequences moving on to the binding analysis step may need to be re-created by cloning, if they are not of high enough frequency to easily isolate with routine colony sequencing of tens or hundreds of colonies. As this technology advances, becoming less expensive and more routine, and especially with an option for colony isolation, it will likely be the best way to analyze biopanning data. It may lead to elucidation of the way individual sequences, and the entire library, evolve. Additionally, the bacteria in the sorting library may mutate over time, which could bias sorting towards cells with faster growth rates or bacteria that overexpress genomic proteins able to bind a material even without display scaffold and peptide expression, for instance. For that reason, once promising candidates emerge, it is worth isolating the plasmid DNA, retransforming fresh *E. coli,* and confirming peptide binding via FACS. If planning to use the peptide in a cell-free format, affinity should also be assessed for the free peptide. For example, peptide affinity reagents discovered through bacterial display for the target SEB were synthetically produced off-cell and analyzed by more traditional methods such as enzyme-linked immunosorbant assay (ELISA), and on-cell (via FACS) and off-cell (via ELISA) affinity were compared using the K_ds determined by both methods [7].

Although the strength of the biotin-streptavidin interaction makes it an ideal strategy for capture of protein target and bound peptides, this procedure can be modified for other capture strategies. Direct coupling of protein to magnetic beads, such as epoxy and carboxylic acid beads, has been successful and removes a binding step. However, direct attachment may hinder potential binding sites in a specific or non-specific manner, depending on attachment strategy. An additional advantage to using streptavidin beads is that less stable protein targets can be freshly biotinylated in 30 min or stored frozen, if needed, while the beads themselves tend to require longer incubation with the protein for cross-linking and are generally stored at 2 - 8°C. The suggested starting concentration of biotinylated protein target here, 600 nM, has been successful for multiple targets, but it may be possible to isolate peptide capture reagents with increased affinity if this concentration is reduced. Additionally, this method can be extended to and modified for other bacterial display libraries and other organisms. For example, these steps can be performed in the absence of oxygen for use with anaerobes, or in other environments for use with extremophiles to isolate peptides with unique properties. The peptides and/or consensus determined for a particular target of interest can be potentially used on-cell as a living material, or produced synthetically off-cell. Synthetically produced peptides can be further matured for the development of even higher affinity and more robust Protein Catalyzed Capture (PCC) agents for use in sensors, or for other applications where antibodies would typically be used for binding or recognition [38,55]. Molecular modeling can also be used to help determine binding location of the peptide with its target, as was recently performed for the abrax binding peptides and consensus listed in **Figure 5C** [1]. Overall, biopanning bacterial display libraries for peptide affinity reagents is a fast, straightforward, and powerful alternative to the production of antibodies for recognition and detection of protein targets, and this semi-automated biopanning approach has produced reliable results with far reaching applications.

Disclosures

The authors have nothing to disclose

Acknowledgements

References

1. Sarkes, D. A., Hurley, M. M., & Stratis-Cullum, D. N. Unraveling the Roots of Selectivity of Peptide Affinity Reagents for Structurally Similar Ribosomal Inactivating Protein Derivatives. *Molecules.* **21** (11), 1504 (2016).
2. Adams, B. L., Finch, A. S., Hurley, M. M., Sarkes, D. A., & Stratis-Cullum, D. N. Genetically Engineered Peptides for Inorganics: Study of an Unconstrained Bacterial Display Technology and Bulk Aluminum Alloy. *Adv. Mater.* **25** (33), 4585-4591 (2013).
3. Bessette, P. H., Rice, J. J., & Daugherty, P. S. Rapid isolation of high-affinity protein binding peptides using bacterial display. *Protein Eng. Des. Sel.* **17** (10), 731-739 (2004).
4. Daugherty, P. S. Protein engineering with bacterial display. *Curr. Opin. Struct. Biol.* **17** (4), 474-480 (2007).
5. Getz, J. A., Schoep, T. D., & Daugherty, P. S. Peptide Discovery Using Bacterial Display and Flow Cytometry. *Methods Enzymol.* **503** 75 (2012).
6. Kenrick, S. A., & Daugherty, P. S. Bacterial display enables efficient and quantitative peptide affinity maturation. *Protein Eng. Des. Sel.* **23** 9-17 (2010).
7. Kogot, J. M. *et al.* Screening and characterization of anti-SEB peptides using a bacterial display library and microfluidic magnetic sorting. *J. Mol. Recognit.* **27** (12), 739-745 (2014).
8. Kogot, J. M., Pennington, J. M., Sarkes, D. A., Stratis-Cullum, D. N., & Pellegrino, P. M. *Population Enrichment and Isolation with Magnetic Sorting, DTIC Document No. ARL-TN-0452.* Report No. ARL-TN-0452, US Army Research Laboratory, Sensors and Electron Devices Directorate, Adelphi, MD., (2011).
9. Kogot, J. M. *et al.* Screening of peptide libraries against protective antigen of Bacillus anthracis in a disposable microfluidic cartridge. *PLOS ONE.* **6** (11), e26925 (2011).
10. Little, L. E., Dane, K. Y., Daugherty, P. S., Healy, & K. E., Schaffer, D. V. Exploiting bacterial peptide display technology to engineer biomaterials for neural stem cell culture. *Biomaterials.* **32** (6), 1484-1494 (2011).
11. Pennington, J. M., Kogot, J. M., Sarkes, D. A., Pellegrino, P. M., & Stratis-Cullum, D. N. Isolation and characterization of anti-SEB peptides using magnetic sorting and bacterial peptide display library technology. in *SPIE Defense, Security, and Sensing.* 83581Z-1-83581Z-6 International Society for Optics and Photonics, (2012).
12. Rice, J. J., & Daugherty, P. S. Directed evolution of a biterminal bacterial display scaffold enhances the display of diverse peptides. *Protein Eng. Des. Sel.* **21** (7), 435-442 (2008).
13. Rice, J. J., Schohn, A., Bessette, P. H., Boulware, K. T., & Daugherty, P. S. Bacterial display using circularly permuted outer membrane protein OmpX yields high affinity peptide ligands. *Protein Sci.* **15** (4), 825-836 (2006).
14. Sarkes, D. A., Dorsey, B. L., Finch, A. S., & Stratis-Cullum, D. N. Method for Discovery of Peptide Reagents Using a Commercial Magnetic Separation Platform and Bacterial Cell Surface Display Technology. *J. Anal. Bioanal. Tech.* **6** (4), 1 (2015).
15. Sarkes, D. A. *et al.* Rapid discovery of peptide capture candidates with demonstrated specificity for structurally similar toxins. in *SPIE Commercial+ Scientific Sensing and Imaging.* 986305-1-986305-10 International Society for Optics and Photonics, (2016).
16. Stratis-Cullum, D., Kogot, J. M., Sarkes, D. A., Val-Addo, I., & Pellegrino, P. M. Bacterial display peptides for use in biosensing applications. in *On Biomimetics.* Pramatarova, L., ed., InTech, Ch. 30, 629-642 (2011).
17. Stratis-Cullum, D. N., & Finch, A. S. Current Trends in Ubiquitous Biosensing. *J. Anal. Bioanal. Tech.* **S7** (009) (2013).
18. Rockberg, J., Löfblom, J., Hjelm, B., Uhlén, M., & Ståhl, S. Epitope mapping of antibodies using bacterial surface display. *Nature Methods.* **5** (12), 1039-1045 (2008).
19. Francisco, J. A., Campbell, R., Iverson, B. L., & Georgiou, G. Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. *Proceedings of the National Academy of Sciences.* **90** (22), 10444-10448 (1993).
20. Georgiou, G. *et al.* Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat. Biotechnol.* **15** (1), 29-34 (1997).
21. Löfblom, J., Wernérus, H., & Ståhl, S. Fine affinity discrimination by normalized fluorescence activated cell sorting in staphylococcal surface display. *FEMS Microbiol. Lett.* **248** (2), 189-198 (2005).
22. Kjærgaard, K., Sørensen, J. K., Schembri, M. A., & Klemm, P. Sequestration of zinc oxide by fimbrial designer chelators. *Appl. Environ. Microbiol.* **66** (1), 10-14 (2000).
23. Kishimoto, J., Fukuma, Y., Mizuno, A., & Nemoto, T. K. Identification of the pentapeptide constituting a dominant epitope common to all eukaryotic heat shock protein 90 molecular chaperones. *Cell stress & chaperones.* **10** (4), 296-311 (2005).
24. Westerlund-Wikström, B. Peptide display on bacterial flagella: principles and applications. *International journal of medical microbiology.* **290** (3), 223-230 (2000).
25. Gaskin, D. J., Starck, K., Turner, N. A., & Vulfson, E. N. Phage display combinatorial libraries of short peptides: ligand selection for protein purification. *Enzyme Microb. Technol.* **28** (9), 766-772 (2001).
26. Goldman, E. R. *et al.* Phage-displayed peptides as biosensor reagents. *J. Mol. Recognit.* **13** (6), 382-387 (2000).
27. Boder, E. T., & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15** (6), 553-557 (1997).
28. Cherf, G. M., & Cochran, J. R. Applications of Yeast Surface Display for Protein Engineering. *Yeast Surface Display: Methods, Protocols, and Applications.* 155-175 (2015).
29. London, N., Movshovitz-Attias, D., Schueler-Furman, O. The structural basis of peptide-protein binding strategies. *Structure.* **18** (2), 188-199 (2010).
30. Sezonov, G., Joseleau-Petit, D., & D'Ari, R. Escherichia coli physiology in Luria-Bertani broth. *J. Bacteriol.* **189** (23), 8746-8749 (2007).
31. Jahnke, J. P., Terrell, J. L., Smith, A. M., Cheng, X., & Stratis-Cullum, D. N. Influences of Adhesion Variability on the "Living" Dynamics of Filamentous Bacteria in Microfluidic Channels. *Molecules.* **21** (8), 985 (2016).
32. Chen, A. Y. *et al.* Synthesis and patterning of tunable multiscale materials with engineered cells. *Nature materials.* **13** (5), 515-523 (2014).
33. Löfblom, J., Frejd, F. Y., & Ståhl, S. Non-immunoglobulin based protein scaffolds. *Current opinion in biotechnology.* **22** (6), 843-848 (2011).

34. Ladner, R. C., Sato, A. K., Gorzelany, J., & de Souza, M. Phage display-derived peptides as therapeutic alternatives to antibodies. *Drug Discov. Today.* **9** (12), 525-529 (2004).

35. Pini, A., Falciani, C., & Bracci, L. Branched peptides as therapeutics. *Current Protein and Peptide Science.* **9** (5), 468-477 (2008).

36. Gongora-Benitez, M., Tulla-Puche, J., & Albericio, F. Multifaceted roles of disulfide bonds. Peptides as therapeutics. *Chemical reviews.* **114** (2), 901-926 (2013).

37. Coppock, M. B. *et al.* Peptide-based protein capture agents with high affinity, selectivity, and stability as antibody replacements in biodetection assays. in *SPIE Sensing Technology+ Applications.* 910711-1-910711-6 International Society for Optics and Photonics, (2014).

38. Coppock, M. B. *et al.* Protein Catalyzed Capture Agents with Tailored Performance for In Vitro and In Vivo Applications. *Peptide Science.* (2016).

39. Bessette, P. H., Hu, X., Soh, H. T., & Daugherty, P. S. Microfluidic library screening for mapping antibody epitopes. *Anal. Chem.* **79** (5), 2174-2178 (2007).

40. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1** (2), 755-768 (2006).

41. Miltenyi Biotec GmbH. *autoMACS™ Pro Separator User Manual, Version 1.1.* Miltenyi Biotec GmbH, Bergisch Gladbach, Germany, (2007).

42. Henriques, S. T. *et al.* A novel quantitative kinase assay using bacterial surface display and flow cytometry. *PLOS ONE.* **8** (11), e80474 (2013).

43. *BD FACSCanto II Operator Course Workbook.* Becton, Dickinson and Company, (2007).

44. *BD FACSCanto II Instructions For Use.* Becton, Dickinson and Company, (2007).

45. Chan, L. Y., Yim, E. K., & Choo, A. B. Normalized Median Fluorescence: An Alternative Flow Cytometry Analysis Method for Tracking Human Embryonic Stem Cell States During Differentiation. *Tissue Eng. Part C Methods.* **19** (2), 156-165 (2012).

46. Genewiz. *DNA Sequencing Services.* <https://www.genewiz.com/> (2017).

47. Honegger, A., & PluÈckthun, A. Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *Journal of molecular biology.* **309** (3), 657-670 (2001).

48. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7** (1), 539 (2011).

49. Lassmann, T., & Sonnhammer, E. L. Kalign-an accurate and fast multiple sequence alignment algorithm. *BMC bioinformatics.* **6** (1), 298 (2005).

50. EMBL-EBI. *Clutal Omega.* <http://www.ebi.ac.uk/Tools/msa/clustalo/> (2017).

51. EMBL-EBI. *Kalign.* <http://www.ebi.ac.uk/Tools/msa/kalign/> (2017).

52. Clamp, M., Cuff, J., Searle, S. M., & Barton, G. J. The jalview java alignment editor. *Bioinformatics.* **20** (3), 426-427 (2004).

53. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* **25** (9), 1189-1191 (2009).

54. *Jalview.* <http://www.jalview.org/Download> (2017).

55. Farrow, B. *et al.* A chemically synthesized capture agent enables the selective, sensitive, and robust electrochemical detection of anthrax protective antigen. *ACS Nano.* **7** (10), 9452-9460 (2013).