Published in final edited form as:

Methods Mol Biol. 2009 ; 539: 93–114. doi:10.1007/978-1-60327-003-8_6.

High Throughput Substrate Phage Display for Protease Profiling

Boris Ratnikov, Piotr Cieplak, and Jeffrey W. Smith

Summary

The interplay between a protease and its substrates is controlled at many different levels, including coexpression, colocalization, binding driven by ancillary contacts, and the presence of natural inhibitors. Here we focus on the most basic parameter that guides substrate recognition by a protease, the recognition specificity at the catalytic cleft. An understanding of this substrate specificity can be used to predict the putative substrates of a protease, to design protease activated imaging agents, and to initiate the design of active site inhibitors. Our group has characterized protease specificities of several matrix metalloproteinases using substrate phage display. Recently, we have adapted this method to a semiautomated platform that includes several high-throughput steps. The semiautomated platform allows one to obtain an order of magnitude more data, thus permitting precise comparisons among related proteases to define their functional distinctions.

Keywords

Substrate phage display; Substrate; Protease; Specificity; Proteolysis; Filamentous phage; M13 coat protein; 3 gene protein

1. Introduction

Sequence comparisons from over 100 genomes shows that peptidases comprise about 2% of all gene products (1), and 50 years of research show that proteolysis plays an important role in most biological processes. Yet, we still have little knowledge on which protein substrates are cleaved by a particular protease, and we lack the technology for localizing and quantifying proteolytic events in cells and whole animals. One strategy for addressing these issues is to define the fine substrate recognition profile of individual proteases and then use this information to predict physiologic substrates and to design imaging agents and other probes of that particular protease.

Defining the substrate specificity of a protease involves assignment of amino acid preferences for each of the subsites within the catalytic cleft. These sites are normally referred to as primed (on the right side of the scissile bond) and unprimed (on the left side of the scissile bond) according to the definitions of Berger and Schechter (2). Different approaches have been developed for acquiring this information (for review see **ref.** 3). These include the use of peptide substrate libraries, as well as biological display using bacteriophage and *E. coli*. Filamentous phage display (4) was first used for protease specificity profiling by Matthews and Wells in 1993 (5). Since then, substrate phage display has been used to profile a large number of individual proteases (6–11).

Substrate phage display is typically carried out by displaying a randomized peptide substrate as a fusion protein with the gene 3 protein (g3p) of filamentous M13 bacteriophage. Since g3p is expressed in 3–5 copies per phage, there is a polyvalent display of the substrate peptide. This peptide is flanked on its C-terminal side by g3p and a “spacer” designed to keep the randomized sequence in a disordered conformation. An affinity tag is normally appended to the N-terminal side of the peptide, and this is used to separate cleaved from

uncleaved phages during the selection process (**Fig. 1**). Our group uses a library composed of randomized hexapeptides displayed on g3p and flanked by a FLAG epitope engineered at the NH₂ terminus of the g3p (12). The substrate phage library was generated using a modified version of the fUSE5 phagemid.

One of the major limitations of substrate phage display is the length of time required to obtain data that yield a comprehensive substrate recognition profile for a given protease. Recently we have incorporated several automated steps into the substrate phage method that allow substantially higher throughput, and that yield an order of magnitude more data. With this platform in place we believe it possible to define the substrate recognition specificity of every endopeptidase in the human genome within 5–10 years. Here we describe the detailed method linked to the semiautomated platform.

2. Materials

2.1. Phage Propagation and Purification

1. NZY Broth: 5 g NaCl, 2 g MgSO₄ · 7H₂O, 5 g Bacto-yeast extract, NZ amine (casein hydrolysate), ddH₂O to 1 L. Adjust pH to 7.5 with NaOH and autoclave.
2. 100 mg/mL Kanamycin Sulfate in water, store at –20°C.
3. 20 mg/mL Tetracycline in 50% glycerol store at –20°C.
4. PEG/NaCl: 100 g PEG 8000 (Sigma catalog # P4463, 116.9 g NaCl, 475 mL water; stir until solutes dissolve (may be necessary to heat to 65° briefly to dissolve the last crystals of PEG). Sterile filter through a 0.22 µm membrane. Store at 4°C (total volume 600 mL).
5. K91Kan *E. coli* cells.
6. Glycerol stock solution 2×: 25 mM Tris, pH 8.0, 0.1 M MgSO₄, 65% Glycerol.

2.2. Phage Substrate Selection

1. Substrate phage library based on the fUSE5 phagemid with a FLAG epitope engineered at the NH₂ terminus of the gene III protein.
2. Stock solution of the protease of interest preferably at high concentration.
3. High affinity inhibitor for active site titration and protease inhibition.
4. Monoclonal anti-FLAG antibodies M2 (Sigma catalog # F3165).
5. Monoclonal anti-M13 antibodies (GE Healthcare catalog # 27-9421-01).
6. Dynabeads M-450 epoxy (Invitrogen catalog # 14011).
7. Dynal MPC-S magnetic separator (Invitrogen catalog # 120-20D).
8. Monoclonal anti-M13 antibodies HRP conjugated (GE Healthcare catalog # 27-9420-01).
9. OPD HRP substrate (Sigma catalog # P3804-100TAB).
10. HRP substrate buffer: 2.43 mL of 0.1 M Citric Acid, 2.57 mL of 0.2 M Na₂HPO₄, 5 mL of DI water.
11. 30% H₂O₂.
12. 4 M H₂SO₄.
13. 3 M (NH₄)₂SO₄.

14. PBS: 50 mM sodium phosphate, pH 7.4, 150 mM NaCl.
15. 100 mM sodium phosphate, pH 7.0.
16. TBS: 50 mM Tris, pH 7.2, 150 mM NaCl.
17. TBS-T: TBS, 0.1% Tween 20.
18. Polystyrene 96-well microplates (Costar catalog # 9018).
19. 50 mL centrifuge tubes (Corning catalog # 430828 or similar).
20. 100 mg/mL BSA (bovine serum albumin, Sigma catalog # A7906) in TBS.
21. Microplate reader.
22. Multichannel micropipettor.
23. PC with Graph Pad Prism or other data fitting software.

2.3. Substrate Identification

1. K-91Kan *E. coli* glycerol stock.
2. Terrific Broth: 12 g bacto-tryptone, 24 g yeast extract, 4 mL (5.04) g glycerol, DI water 900 mL. Autoclave 90 mL portions in 125 mL polypropylene bottles. When cooled, to each bottle add 10 mL of separately autoclaved potassium phosphate buffer (0.17 M KH_2PO_4 , 0.72 M K_2HPO_4).
3. LB Agar, 100 $\mu\text{g/mL}$ Kanamycin: 10 g Tryptone, 5 g yeast extract, 5 g NaCl, 200 μL NaOH, 15 g Bacto Agar DI water to 1 L. Autoclave, cool to 50°C, add 1 mL 100 mg/mL Kanamycin – pour 100 mm Petri dishes at 20 mL/dish. Store at 4°C after agar solidifies.
4. NZY Agar, 100 $\mu\text{g/mL}$ Kanamycin, 40 $\mu\text{g/mL}$ Tetracycline: 5 g NaCl, 2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 5 g Bacto-yeast extract, NZ amine (casein hydrolysate), 15 g Bacto Agar, DI H_2O to 1 L. Adjust pH to 7.5 with NaOH. Autoclave, cool to 50°C, add 1 mL 100 mg/mL Kanamycin – pour into Genetix QTray vented (Genetix catalog # 6023) at 200 mL/tray for automated colony picking using Genetix QPix colony picker or at 20 mL per 100 mm Petri dish. Store at 4°C after agar solidifies.
5. Multichannel pipettor.
6. Hamilton LabStar liquid handling robot for HTS phage display.
7. 96 deep well plate (Simport Bioblock deep well plate, catalog # T110-10) for HTS phage display or 15 mL sterile disposable round bottom culture tubes for low throughput phage display.
8. Genetix 96-well flat bottom 1/2 height plate (Genetix catalog # X6011) for HTS phage display or sterile 1.5 mL Eppendorf centrifuge tubes for low throughput phage display.
9. Genetix BreatheSeal Film (catalog # E1005).
10. Aluminum sealing tape, (ISC BioExpress catalog # T-2420-2).
11. Glycerol stock solution.
12. Monoclonal anti-M13 antibodies.
13. Monoclonal anti-FLAG M2 antibodies peroxidase conjugated (Sigma catalog # A8592).

14. TBS-T: TBS, 0.1% Tween 20.
15. Polystyrene 96-well microplates.
16. 100 mg/mL BSA in TBS.
17. HRP Substrate Buffer.
18. OPD HRP substrate.
19. 30% H₂ O₂.
20. Microplate reader.
21. ATR Multitron Incubator shaker.

2.4. Substrate Sequencing

1. FUSE5 forward primer: 5' –TAA TAC GAC TCA CTA TAG GGC AAG CTG ATA AAC CGA TAC AAT T–3' , 100 µM stock.
2. FUSE5 Super Reverse primer: 5' –CCG TAA CAC TGA GTT TCG TC–3' , 100 µM stock.
3. Platinum PCR Super Mix (Invitrogen catalog # 11306- 016).
4. 96-well PCR plate (Abgene catalog # AB1000).
5. PCR thermocycler.
6. 50× TAE Buffer.
 - a. 242 g Tris base
 - b. 57.1 mL acetic acid
 - c. 100 mL 0.5 M EDTA
 - d. Add ddH₂O to 1 L and adjust pH to 8.5
7. 10 mg/mL ethidium bromide (Sigma catalog # E1510).
8. 1% agarose in TAE buffer, 0.5 µg/mL ethidium bromide.

2.5. Identification of Scissile Bonds

1. NZY Broth.
2. 96 deep well plate (Simport Bioblock deep well plate, catalog # T110-10) for HTS phage display or 15 mL sterile disposable round bottom culture tubes for low throughput phage display.
3. Costar Assay Block 1 mL plate (catalog # 3958).
4. Costar Storage Mat III (catalog # 3080).
5. Monoclonal anti-FLAG antibodies M2.
6. Monoclonal anti-M13 antibodies.
7. Dynabeads M-450 epoxy.
8. Dynabeads M-280 Tosylactivated (Invitrogen catalog # 14204).
9. 2,5-Dihydroxybenzoic acid, 99% (Aldrich catalog # 14,935-7).
10. Trifluoroacetic acid (Sigma Aldrich catalog # T6508).

11. MALDI AnchorChip target (Bruker Daltonics catalog # 209515) or other MALDI target.
12. 3 M (NH₄)₂SO₄.
13. 100 mM sodium phosphate, pH 7.0.
14. TBS-T: TBS, 0.1% Tween 20.
15. V-bottom polypropylene plates (Costar catalog # 3357).

3. Methods

Prior to delving into procedures for substrate phage display, it is important to mention a few factors and pitfalls that have a significant impact on the success of the procedure. A major factor in the success of substrate phage display is the availability of an ample supply of the test protease, and a reliable method of quantifying its activity. Each time a protease is used for substrate phage its activity should be titrated with an active site inhibitor (13) whenever possible. When the test proteases are expressed and purified in our laboratories, more than 70% of the purified protein is active. With this level of activity, 500 µg to 1 mg of total protease are usually sufficient to complete the phage procedure. We are reluctant to move forward with phage selections for any protease in which less than 30–40% of the total protein has activity. With this quantitative information in hand, we have the basis for quantitative analysis and comparison of any of the data obtained in the course of the experiment.

Another factor that has a major impact on success is the quality of the K91Kan *E. coli* used for phage infection and propagation. These bacteria must be fresh; generated and used within 24 h. This is a very important requirement and should be followed strictly to obtain consistent results.

An important pitfall in this procedure, as it is with any phage display project, is the possibility of phage contamination. There are really two different manifestations of this problem. The first is the simple contamination of laboratory tools and instruments with a single phage, which then begins to dominate all successive phage amplifications and overrides any selection process. This can be prevented by highly stringent sterile techniques, the use of pipette tips containing filters, and frequent washing of instrumentation. All buffers and solutions should be sterilized either by autoclaving or 0.2 µm filtration.

In another manifestation this problem arises as the propagation of a phage lacking the affinity tag, which then appears as “cleaved” even when it is not (8). This is not a problem for the first round of selection, but can be an issue after the second round, when some phages have adapted to the selection pressure by eliminating the tag from their sequence. To overcome this issue, we include an affinity isolation of the tag-bearing phage prior to the next round of selection.

A protease phage display project involves four basic steps: (1) substrate selection, (2) substrate identification, (3) substrate sequencing, and (4) scissile bond assignment. Steps 2–4 have been automated by our group. The procedure for each step is described in detail later.

3.1. Phage Substrate Selection

These instructions assume that a FLAG-tagged phage library is available for substrate selection.

1. Before starting substrate selection, prepare M-450 M2 beads. Wash 5×10^8 M-450 epoxy magnetic beads three times by 1 mL 100 mM sodium phosphate buffer, pH

7.0 by magnetic separation. Resuspend the beads in 500 μ L 0.5 mg/mL M2 antibody and add 250 μ L 3 M $(\text{NH}_4)_2\text{SO}_4$. Incubate with end-over-end rotation or vortexing overnight at room temperature, making sure the beads are in suspension. In the morning collect the supernatant and measure light absorbance at 280 nm by spectrophotometry. Determine the efficiency of coupling by dividing by the absorbance of the starting antibody solution adjusted by dilution factor of 1.5. Usually 75–90% of the antibody is coupled under these conditions. Resuspend the beads in 1 mL of 1 mg/mL BSA in PBS and incubate with end-over-end rotation or vortexing for 3 h at room temperature. Store the beads at 4°C.

2. The day of substrate selection, perform active site titration of the protease or specific activity determination by some other method to be able to reproduce the conditions of the experiment.
3. For proteases requiring extreme pH for optimal activity, determine neutralization conditions.
4. Determine the concentration of the phage particles in the stock solution of the phage library. Dilute the phage in TBS to a concentration approximately 10^{12} particles/mL. Scan the diluted sample from 240 to 320 nm in a UV spectrophotometer. There should be a broad peak from 260 to 280 nm with a slight maximum at 269 nm. Measure A_{269} and A_{320} and calculate the net A_{269} by subtracting A_{320} from A_{269} . Calculate the virion concentration in particles per mL based on the following formula: phage particles/mL = net $A_{269} \times 6 \times 10^{16} / N$, where N = the number of nucleotides/phage genome. The genome size of a fUSE5 phage = 9,206 bases.
5. Add 1.2×10^{12} phage to the final volume of 1.2 mL in protease reaction buffer. Take out two 0.4 mL aliquots for negative control and protease treatment. To the first 0.4 mL aliquot add the protease in the protease buffer, to the final concentration of 0.2 μ M of active enzyme. To the second and third aliquots add the same volume of the protease buffer. Incubate the mixtures for 2 h at the optimal temperature for the protease. Upon completion of the incubation add protease inhibitor to all the samples and adjust the pH to 7.2–7.4. Bring the volumes up to 1 mL with TBST.
6. During the incubation wash 0.8 mL worth of M-450 M2 bead suspension in 1 mL TBST three times using magnetic separator. After the second wash, split the magnetic bead suspension in four 0.25 mL aliquots and remove the supernatant after magnetic separation from two of them. Resuspend the beads with protease treated and 1 aliquot of control phage in one tube each and incubate for 1 h at ambient temperature with rotation end-over-end. The second control aliquot of phage will be used as the before-depletion control. Do not use vortexing to keep the beads in suspension when doing phage separations. After the incubation is over, remove the supernatant from the remaining two tubes of M-450 M2 beads by magnetic separation. Collect the unbound phage from the control and the protease treated phage by magnetic separation and mix with the fresh M-450 M2 beads from the remaining tubes. Incubate at ambient temperature with end-over-end rotation for 3 h. By the end of incubation, collect the supernatants by magnetic separation and determine the efficiency of depletion the following day by ELISA. Store the phage samples at -70°C .
7. Coat a 96-well microtiter plate with 100 μ L/well of 5 μ g/mL anti-M13 antibody in TBS at 4°C overnight.

8. The same night streak an LB agar 100 µg/mL Kanamycin Petri dish with K91Kan *E. coli* cells and incubate at 37°C overnight. Remove the dish from the incubator in the morning of the following day and place at 4°C until use.
9. In the morning of the next day block the plate with 120 µL/ well of 100 mg/mL BSA in TBS for 1 h at 37°C.
10. Prepare seven dilutions of phage with known concentration (standard) from 10¹¹ phage/mL with a step of 1:2 by combining 0.35 mL of each dilution with 0.35 mL TBST. Dilute the before-depletion control, the protease treated and the control M2 depleted phage 1:20 by adding 35 µL of the solution to 665 µL TBST. Serially dilute the phage three times by combining 0.35 mL with 0.35 mL TBST. Add 100 µL/well from each dilution in triplicate. Incubate for 1 h at 37°C. When performing ELISA for subsequent rounds of selection, it is important to use phage from the starting stock used for this round of selection as standard, to obtain consistent results.
11. After incubation, wash the plate with 120 µL/well TBST three times. Add 100 µL/ well anti-M13 monoclonal antibody – HRP conjugate diluted in TBST as recommended by the manufacturer. Incubate for 1 h at 37°C.
12. Upon completion of the incubation prepare the HRP substrate solution in the HRP substrate buffer. Wash the plate three times with 120 µL/well TBST. Add 4 µL/10 mL of 30% H₂O₂ to the substrate solution and apply at 100 µL/well to the plate. Incubate with shaking until color develops and add 50 µL 4 M H₂SO₄ to stop the reaction. Read the plate in a microtiter plate reader at 490 nm.
13. Plot the A_{490} vs. phage concentration for the standard curve. Fit the data points with the two site binding equation: $A_{490} = B_{\max 1} C / (K_{d1} + C) + B_{\max 2} C / (K_{d2} + C)$, where $B_{\max 1}$ and $B_{\max 2}$ are the maximum binding per site, and K_{d1} and K_{d2} are respective dissociation constants, and C is phage concentration. Apparently, the anti-M13 antibody recognizes two different classes of binding sites on the M13 coat protein expressed by the phage. Determine the phage concentration for the samples from the standard curve. Typically 98–99% of phages are depleted in the untreated sample as observed for the first round of selection, while the difference in depletion between the protease treated and untreated phage represents the enrichment for protease substrates (**Fig. 2**).
14. Start an overnight culture of K91Kan *E. coli* by inoculating 3 mL of LB 100 µg/mL Kanamycin with a single colony of a K91Kan cells streaked the night before. Incubate at 37°C overnight with rotation at 250 rpm.
15. Take 100 µL of the overnight culture and add to 9.9 mL Terrific Broth in a sterile 50 mL centrifuge tube. Incubate at 37°C with shaking at 250 rpm for 2 h. After 2 h start monitoring A_{600} of the culture by taking out 0.1 mL and adding to 0.9 mL of DI water and measuring A_{600} . Multiply the obtained number by 10 to get the A_{600} for the culture.
16. When $A_{600} = 2$ AU (10¹⁰ cells/mL), slow down the shaking to 50 rpm for 5 min. Add an appropriate amount of the bacterial culture and the correspondingly appropriate volume of Terrific Broth so that after the addition of the substrate enriched phage, both the *E. coli* and the phage will be at 2×10^9 particles/mL. Shake the culture at 100 rpm at 37°C for 15 min. Transfer the infected cells into 10× volume of NZY broth containing 0.22 µg/mL tetracycline. Shake the culture for 35 min at 250 rpm, 37°C.

17. Add $0.001 \times$ volume of 20 mg/mL Tetracycline to the culture and incubate overnight at 250 rpm 37°C.
18. Determine the number of infected cells. Add 7 μ L of culture to 63 μ L LB medium. Perform serial dilutions three times by adding 20 μ L of bacteria from the previous dilution to 180 μ L LB. Plate 100 μ L of each dilution onto 100 μ g/ mL Kanamycin, 40 μ g/mL Tetracycline NZY agar 10 cm Petri dish. Incubate overnight at 37°C. In the morning of the following day, count the colonies on each dish. Multiply by the appropriate dilution factor and then by 10 to get the number of infected cells per mL. Knowing the total number of the infected cells is useful in estimating how many individual phage clones have been obtained during infection.
19. The following morning centrifuge the culture in a sterile centrifuge tube or bottle at $2,400 \times g$ for 10 min at 4°C. Without disturbing the cell pellet, transfer the phage-containing super-natant into a clean sterile centrifuge container and centrifuge again at $6,200 \times g$ for 10 min at 4°C. Carefully pour the supernatant into a clean sterile container with a screw cap.
20. Add $0.15 \times$ volume of PEG/NaCl solution to the phage, close the container and mix thoroughly by inverting the container. Incubate on ice for at least 4 h or overnight at 4°C. The medium should turn cloudy.
21. Pellet the phage at $6,200 \times g$ for 40 min at 4°C. Carefully discard the supernatant without disturbing the pellet. Remove the residual supernatant by briefly centrifuging the pellets and carefully aspirating the rest of the liquid.
22. Add 5 mL of sterile TBS and shake the pellets at 150 rpm at 37°C for approximately 30 min to resuspend the pellet.
23. Centrifuge the phage solution at $10,100\text{--}22,700 \times g$ for 10 min at 4°C. Transfer the supernatant into a fresh sterile centrifuge container.
24. Add $0.15 \times$ volume of PEG/NaCl solution to the container and mix thoroughly by inversion. Allow the phage to precipitate for 1 h on ice. A heavy precipitate should appear.
25. Centrifuge the phage at $10,100 \times g$ for 40 min at 4°C. Remove the supernatant.
26. Add 2 mL of sterile TBS to the pellet and allow the pellet to soften for 1 h. Vortex the pellet again to bring the phage into solution
27. Clear the supernatant by centrifuging at $10,100\text{--}22,700 \times g$ for 10 min at room temperature or 4°C.
28. Collect the supernatant. Determine the phage concentration as described in **step 4** above. Store the phage short term at 4°C, long term at -70°C . The first round of selection is complete.
29. Repeat the above steps for the second round. If more than two rounds of selection are required, as can be the case if poor enrichment for substrate phage is observed after the second round, we recommend enriching the phage from the second round for high FLAG expressers. The reason for this is that after each amplification, there is a loss of tag expressing phage and it can be quite substantial after the second round. We routinely enrich phage for FLAG expression before proceeding to substrate identification.
30. Take 200 μ L M2 M450 beads (10^8) and wash three times with TBST by magnetic separation. Add 2.5×10^{12} phage to the beads and bring the volume up to 1 mL with TBST. Incubate overnight with end-over-end rotation.

31. The following day, make 10 mL of 0.005% Brij 35 solution in DI water and sterilize it by filtration through a 0.2 μ m filter.
32. Prepare 0.05% TFA solution in 0.005% Brij 35 solution.
33. Wash the beads three times with 1 mL TBST by magnetic separation after discarding the supernatant followed by two-time wash with 0.005% Brij solution. Remove any traces of liquid from the beads and add 100 μ L of 0.05% TFA to the beads. Shake for 10 min at room temperature making sure the beads stay in suspension. Collect the supernatant by magnetic separation and mix with 100 μ L TBS. Determine the phage concentration as described in **step 4** above. Multiply the obtained number by the total volume to get the total number of phage. Typically we get around 5×10^{11} phage after enrichment.

3.2. Verification of Substrate Phage

After the final round of selection is complete, one needs to verify that individual phage clones are bearing substrates prior to nucleotide sequencing. This is done by substrate phage ELISA of the media from bacterial cultures grown from single colonies of K91Kan cells bearing a single phage clone. To ensure that we select phage clones with adequate FLAG epitope expression, we use a positive control phage with high FLAG expression level (same or higher than the starting library) and select only those substrate phage clones, whose FLAG expression level is above 0.5 of that of the control. This is done to minimize the error of measurement of the degree of hydrolysis of substrate phage, which is determined by comparing the amount of FLAG with and without exposure to protease. In addition, the level of expression of the FLAG epitope reflects the number of copies of the substrate peptide per phage, thus allowing us to select high expressers for MALDI TOF analysis used for identification of the scissile bond in the substrate peptide.

1. Perform enrichment of the phage from the last round of selection for FLAG epitope expression as described in **Subheading 3.1, steps 30–33**.
2. Prepare NZY agar QPix trays or 10 cm Petri dishes by pouring 200 or 20 mL of NZY agar with 100 μ g/mL Kanamycin, 40 μ g/mL Tetracycline per tray/dish for HTS or LTS versions of the procedure respectively. Store the NZY agar trays/dishes at 4°C in a plastic bag.
3. Start an overnight 3 mL culture in LB 100 μ g/mL Kanamycin from a freshly streaked colony. Incubate overnight at 37°C, shaking at 250 rpm.
4. Next morning add 100 μ L of the overnight culture to 10 mL of Terrific Broth. Incubate at 37°C, shaking at 250 rpm.
5. After 2 h of incubation start monitoring A_{600} of 1:10 diluted culture in DI water until it reaches 0.2 AU. Slow the shaking down to 50 rpm for 5 min.
6. Place the Qtrays or Petri Dishes in a 37°C incubator.
7. Add 1 mL of the K91Kan culture (10^{10} cells/mL at $A_{600} = 2.0$) to 9 mL Terrific Broth and supplement with 10^9 FLAG enriched phage from the last round of selection. Continue incubation at slow shaking for 15 min.
8. Prepare 10 mL NZY broth with 0.22 μ g/mL tetracycline from 20 mg/mL stock. Add 1 mL of infected culture to 10 mL of NZY broth 0.22 μ g/mL Tetracycline. Shake 35 min at 250 rpm, 37°C.
9. Add 10 μ L of 20 mg/mL stock of Tetracycline to the culture.

10. Dilute the culture 50-fold in NZY Broth and plate 0.75 mL per QPix tray or 75 μ L per 10 cm Petri dishes. Place the trays in a 37°C incubator with 70% humidity. The Petri dishes can go in a 37°C incubator without humidification.
11. The next morning pick colonies using the QPix robot in the case of QPis trays or manually from the Petri dishes into 96-well deep-well 2 mL plates filled with 1 mL per well of NZY broth supplemented with 20 μ g/mL Tetracycline. After picking, put the BreatheSeal film on the plates and incubate in the ATR Multitron incubator shaker at 750 rpm, 37°C, 70% humidity. For a small scale project, a regular shaking incubator can be used at 250 rpm, 37°C.
12. The same day, coat 96 well ELISA plates with 100 μ L/well of 5 μ g/mL anti-M13 antibody in PBS at 4°C overnight.
13. In the morning, block the plates with 120 μ L of 100 mg/mL BSA in PBS for 1 h at 37°C.
14. Remove the BSA and add 100 μ L/well of the bacterial cultures from the deep well plates as shown in **Fig. 3**: To the BLNK wells add 100 μ L NZY broth. To the + CTRL wells add 100 μ L of 10^{12} phage/mL of the positive control clone in NZY broth.
15. Incubate the plates for 1 h at 37°C.
16. Wash the plates with 120 μ L/well TBST four times and remove the residual buffer from the plate by gently tapping the plates face down against a paper towel.
17. Add 100 μ L of (50 nM active enzyme) protease solution per well to the left half of the plate as shown in **Fig. 3**. To the right half add 100 μ L/well of the protease buffer. Incubate for 2 h under the optimal conditions for the protease.
18. Wash the plates four times with 120 μ L/well TBST and add 100 μ L/well of HRP conjugated M2 antibody appropriately diluted in TBST. Incubate for 1 h at 37°C. Dissolve 1 tablet of OPD per 10 mL of HRP substrate buffer right before the end of the incubation period.
19. Wash the plates four times with 120 μ L/well TBST, remove the residual buffer by gently tapping the plates face down against a paper towel. Add 4 μ L of 30% H_2O_2 per 10 mL of OPD solution, mix thoroughly and dispense 100 μ L/well. Shake the plates gently and observe color development. Usually the plates can be read 5 min after addition of the substrate, but one should exercise proper judgment to stop the reaction at an appropriate time, so as not to over or under develop the reaction. A target A_{490} value of 1 is optimal. Add 50 μ L/well of 4 M H_2SO_4 to stop the reaction. Read the plates at 490 nm in a microtiter plate reader.
20. Analyze the data by first eliminating from consideration any clones with A_{490} less than 0.5 of that of the positive control. For the remaining clones calculate the degree of hydrolysis by dividing the A_{490} of the + protease well by that of the – protease well and subtracting the result from 1. Disregard any clones with the degree of hydrolysis of less than 0.2. The substrate identification step is complete.
21. Prepare glycerol stocks of the substrate phage by mixing 100 μ L of culture with 100 μ L of glycerol stock solution in a Genetix 96-well flat bottom 1/2 height plate. Store the glycerol stocks at –70°C. In some cases there is a pH requirement for proteolytic activity that is incompatible with phage capture by anti-M13 antibodies. In these cases, we perform protease treatment of PEG precipitated phage in the protease buffer with a pH necessary for optimal protease activity. After that we adjust the pH to neutral and perform the ELISA.

22. Prepare glycerol stocks of the substrate phage by mixing 100 μ L of culture with 100 μ L of glycerol stock solution in a Genetix 96-well flat bottom 1/2 height plate. Seal the plates with aluminum sealing tape. Store the glycerol stocks at -70°C .
23. Centrifuge the remaining overnight cultures at $4,000 \times g$ for 10 min at 4°C . Collect the supernatants into fresh deep 96-well plates. Add 0.15 $\times$ volume of PEG/NaCl solution, cover the plate with a storage mat, mix by inversion end-over-end about 100 times, spin briefly at $300 \times g$ and put at 4°C overnight. Centrifuge at $4,600 \times g$ for 90 min at 4°C . Carefully discard the supernatant and remove residual liquid by gently tapping the plates face down against a paper towel.
24. Add 120 μ L/well of the protease buffer and shake the plates at 750 rpm in a ATR Multitron shaker incubator or similar for 30 min.
25. Transfer 50 μ L of the phage solution to 50 μ L of protease buffer per well on the right side of a 96-well uncoated ELISA plate (**Fig. 3**). Transfer another 50 μ L of the phage solution to 50 μ L of 100 nM protease in protease buffer in the counterpart wells on the left side of the plate. Incubate at appropriate temperature for 2 h.
26. Neutralize the samples by addition of an appropriate amount of neutralization solution (determined earlier). Transfer the samples to an anti-M13 coated BSA blocked plate. Incubate for 1 h at 37°C .
27. Proceed as described in **steps 17–19** above.

3.3. Sequencing Phage Substrates

Once the protease substrate phage clones have been identified, the nucleotide sequence of the random insert is determined. This is done by PCR amplification of the region of the g3p bearing the randomized insert followed by sequencing of the PCR amplicons. The PCR procedure is performed in-house in a 96-well format and the sequencing of the resulting amplicons is outsourced.

1. Thaw the glycerol stocks of the substrate phage at room temperature. Cherry pick the substrate clones identified by substrate phage ELISA and inoculate 1,250 μ L of LB 20 $\mu\text{g/mL}$ Tetracycline per well of 2 mL 96-deep well plates, using QPix colony picking robot for an HTS phage display or a sterile Pasteur loop for an LTS one. Seal the plates with BreatheSeal film. Grow the inoculates overnight at 750 rpm, 37°C , 70% humidity in an ATR Multitron shaker incubator. For a small scale project, a regular shaking incubator can be used at 250 rpm, 37°C .
2. The following morning prepare the PCR master mix:
 - a. 2 μ L/sample 5 μM FUSE5 forward primer
 - b. 2 μ L/sample 5 μM FUSE5 Super Reverse primer
 - c. 45 μ L Platinum PCR Super Mix
3. Dispense 49 μ L of the master mix per well of a 96-well PCR plate. Add 2 μ L per well of the overnight culture. Run the PCR as follows:
 - a. 72°C 10 min
 - b. 94°C 3 min
 - c. $34 \times (94^{\circ}\text{C}$ 50 s, 50°C 1 min, 72°C 1 min)
 - d. 72°C 6 min

e. Hold at 4°C.

4. To determine the quality of the PCR, run 10 PCR products per plate at 10 µL/lane on 1% agarose gel.
5. Freeze the PCR products and send out for sequencing.

3.4. Identification of Scissile Bonds

Once the sequences of phage substrates are known from nucleotide sequencing, the position of the scissile bond within the substrate peptide is determined directly from purified phage using MALDI TOF MS. We take advantage of the fact that the phage express sufficient levels of g3p, so that the peptide hydrolyzed from the phage by protease is present in adequate quantity to analyze by MS. Once the mass of the cleaved portion of the peptide substrate is known, this information can be combined with sequence information to pinpoint the position of the scissile bond.

1. Prepare the necessary amount of M-450 anti-M13 beads at 10^8 beads per sample by following the protocol described in **Subheading 3.1**. Use anti-M13 instead of M2 antibodies.
2. Prepare M2 coupled M-280 beads. Wash 1.29×10^9 beads three times with 1 mL of 100 mM sodium phosphate pH 7.0 by magnetic separation. Resuspend the beads in 1 mL of 250 µg/mL M2 antibody in 100 mM sodium phosphate, pH 7.0.
3. Incubate with end-over-end rotation overnight at 37°C.
4. In the morning of the following day measure the A_{280} of the supernatant after magnetic separation. Determine the amount of antibody coupled. Typically 80–90% of antibody gets coupled to M-280 beads.
5. Remove the supernatant by magnetic separation and resuspend the beads in 1 mL of sterile 1 mg/mL BSA in TBS. Incubate with end-over-end rotation for 4 h at room temperature. Store at 4°C.
6. Thaw the glycerol stocks of substrate phage and inoculate 1.2 mL cultures of substrate phage in 2 mL 96-deep well plates and grow overnight as described in **Subheading 3.3** of **step 1**.
7. The following morning wash the M-450 anti-M13 beads three times with TBST. Resuspend in the original volume of TBST and dispense into an Assay Block 1 mL 96-well plate at 200 µL/well.
8. Spin down the cultures at $4,000 \times g$, 4°C for 10 min.
9. Remove TBST from the M-450 anti-M13 beads by magnetic separation and resuspend the beads in 1 mL of culture super-natant. Seal the plates with the storage mats and incubate with rotation end-over-end for 3 h at room temperature.
10. Wash the beads with 1 mL of the protease incubation buffer three times by magnetic separation and resuspend in 100 µL/well of 50–200 nM active protease. Incubate with vortex shaking overnight at the optimal temperature.
11. Alternatively, one can use PEG precipitation for buffer exchange, especially if the protease buffer has acidic or alkaline pH incompatible with the antibody capture of the phage. This approach provides good results as well, although there will be residual PEG left over. One has to test if PEG has an effect on the protease activity before using this method.

- a. Add 0.15× volume of PEG/NaCl solution, cover the plate with a storage mat, mix by inversion end-over-end about 100 times and incubate at 4°C overnight. Centrifuge at $4,600 \times g$ for 90 min at 4°C. Carefully discard the supernatant and remove residual liquid by gently tapping the plates face down against a paper towel.
 - b. Add 100 µL/well of the protease solution and shake the plates at 750 rpm in a ATR Multitron shaker incubator or similar overnight.
 - c. The next morning, add the appropriate amount of neutralization buffer.
12. Wash an appropriate amount of M2 M-280 magnetic beads three times in TBST by magnetic separation. Resuspend in the original volume of TBST and dispense into an Assay Block 1 mL 96-well plate at 200 µL/well. Remove the TBST after magnetic separation. Resuspend the beads in the protease treated phage and incubate with vortex mixing for 3 h. Wash three times with 100 µL/well of 5 mM Tris, 0.1 mg/mL BSA by magnetic separation. Wash three times with 100 µL/well of 5 mM Tris by magnetic separation. Remove the residual liquid.
 13. Elute the bound peptides by addition of 10 µL/well of 0.1% TFA and vortex mixing for 5 min at room temperature.
 14. Prepare the matrix solution by mixing 10 mg DHB with 0.333 mL acetonitrile and 0.667 mL 0.1% TFA. Add 5 µL of the matrix solution to a v-bottom polypropylene plate.
 15. Remove 5 µL of the eluted peptides after magnetic separation and mix with the matrix solution in the v-bottom polypropylene plate.
 16. Place 1 µL of the peptide/matrix solution per spot of an AnchorChip 400 target. Let dry at room temperature.
 17. Determine the masses of the released peptides by MALDI TOF.
 18. Determine the scissile bond position by matching the mass of the released peptide to the amino acid sequence obtained by translation of the DNA sequencing data of the PCR products (**Fig. 4**).

3.5. Data Analysis

The simplest and at the same time the most informative way of representing positional preferences of a given amino acid in a binding subsite is to use graphical sequence logos. This method was developed by Schneider and Stephens (14) and was implemented as a Web service, called WebLogo, by Crooks et al. (15). It is available at: <http://weblogo.berkeley.edu/>. It displays the nucleic acid bases or amino acids patterns as a result of multiple sequence alignment. The patterns consist of stacks of characters representing the sequence. Each stack represents a single position in the sequence. The characters that occur most commonly are located at the top of the stack. The height of each character is proportional to its frequency at that position. The overall height of the stack reflects the sequence conservation at that position. Instead of using frequencies as a way of scaling the height of the entire stack, that height can be adjusted according to the information content (measured in bits) of the sequence at that position.

According to Schneider and Stephens (14), the information content is calculated in the following way. First, the sequence conservation at a specific position in the alignment: R_{seq} is calculated as the difference between the maximum possible entropy and the entropy of the observed symbol distribution:

$$R_{\text{seq}} = S_{\text{max}} - S_{\text{obs}} = \log_2 \left(N - \left(- \sum_{n=1}^N p_n \log_2 p_n \right) \right)$$

where p_n is the observed frequency of symbol n at a specific position, N is a number of distinct symbols used in a sequence type, which is 4 for nucleic acids and 20 for proteins, yielding the maximum value of entropy ($\log_2 N$) equal to 2 or 4.32, for nucleic acids and proteins, respectively. The sum of sequence conservation (R_{seq}) at each position of the logo measures the information content of the logo. This is true under the assumption that there is no inter-positional correlation and the background distribution of symbols is uniform.

In our calculation for determining patterns in the set of substrates for matrix metalloproteinases studied in the phage display experiments, we use the frequency-based instead of information-based logos. To prepare input for the WebLogo program, e.g., the set of aligned sequences, we use an approach that is based on Position Weight Matrix (PWM). A PWM assumes independence between positions in the pattern and allows one to take into account the background distribution. An element in a PWM is calculated as $m_{i,j} = \log(p_{i,j}/b_j)$, where $p_{i,j}$ is the probability of observing symbol (amino acid) i at position j of the motif, and b_j is the probability of observing the symbol i in a background model. The value of particular element of the PWM is the log-odds of a given symbol being found in the current pattern vs. being found in the background distribution. From the PWM matrix one can rederive the set of sequences that correspond to appropriate log-odds values, which in turn can be used as input to WebLogo program.

The specific procedure of deriving PWM for substrates of metalloproteinases is as follows. First, we aligned the sequences of all substrates for a given enzyme along the hydrolyzed peptide bond. Each sequence of the substrate in the phage display experiment consists of six amino acids variable regions and the N-terminal and C-terminal constant tags. Since a hydrolysis site can occur at any position in the six amino acids variable region, we included some portion of the constant tags in the sequence alignment. All aligned sequences have been uniformly trimmed to the length of eight amino acids; only those amino acids that span P5-P3 positions have been left for further consideration. In the next step, we derived the position weight matrices (PWMs), specific for each metalloproteinase, using the frequencies of each individual amino acid at each position in the alignment. The PWMs have been normalized by distribution of amino acids in the background sequences. In deriving distribution of the background sequences we included also short, three and two amino acid long fragments of the constant tags at the N- and C-terminals, respectively. In this way, we took into account biasing of the background and substrate distributions by constant fragments of the sequences. The background sequences are chosen randomly and are predicted to be hydrolyzed with less than 5% probability. In the last step, each position of the PWM is renormalized to the values in the range between 0 and 1 over all amino acids. Then each value for a given amino acid at a given position is multiplied by 100. This is the final number each amino acid would appear in the set of sequences, which are used as input for the WebLogo program. An example of the PWM for the 234 substrates of MMP-16 metalloproteinase and the corresponding sequence logo are shown in **Figs. 5 and 6**.

3.6. Concluding Remarks

Protease substrate phage display is a powerful tool for delineation of protease substrate specificity at the level of amino acid sequence. Our group adapted this approach for high throughput data acquisition and analysis, thereby making it capable of studying the evolution of substrate specificity in closely related families of proteases. These studies may

shed light on the biological processes affected by different members of the family as a result of differential substrate specificities. As a bonus, one gets information on unique and selective peptide substrates that can be used for such applications as activity-based probes, selective inhibitors, and so on. We feel that the ambitious goal of characterizing substrate specificities of all human proteases is now well within reach. This work was supported by Grant Number RR020843 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

References

1. Barrett, AJ.; Rawlings, Neil D.; Woessner, J. Fred Handbook of Proteolytic Enzymes. Elsevier Academic Press; San Diego: 2004.
2. Berger A, Schechter I. Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Philos Trans R Soc Lond B Biol Sci.* 1970; 257(813):249–64. [PubMed: 4399049]
3. Diamond SL. Methods for mapping protease specificity. *Curr Opin Chem Biol.* 2007; 11(1):46–51. [PubMed: 17157549]
4. Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science.* 1985; 228(4705):1315–7. [PubMed: 4001944]
5. Matthews DJ, Wells JA. Substrate phage: selection of protease substrates by monovalent phage display. *Science.* 1993; 260(5111):1113–7. [PubMed: 8493554]
6. Ohkubo S, Miyadera K, Sugimoto Y, Matsuo K, Wierzbka K, Yamada Y. Identification of substrate sequences for membrane type-1 matrix metalloproteinase using bacteriophage peptide display library. *Biochem Biophys Res Commun.* 1999; 266(2):308–13. [PubMed: 10600499]
7. Cloutier SM, Chagas JR, Mach JP, Gygi CM, Leisinger HJ, Deperthes D. Substrate specificity of human kallikrein 2 (hK2) as determined by phage display technology. *Eur J Biochem.* 2002; 269(11):2747–54. [PubMed: 12047384]
8. Smith MM, Shi L, Navre M. Rapid identification of highly active and selective substrates for stromelysin and matrilysin using bacteriophage peptide display libraries. *J Biol Chem.* 1995; 270(12):6440–9. [PubMed: 7896777]
9. Hills R, Mazzarella R, Fok K, et al. Identification of an ADAMTS-4 cleavage motif using phage display leads to the development of fluorogenic peptide substrates and reveals matrilin-3 as a novel substrate. *J Biol Chem.* 2007; 282(15):11101–9. [PubMed: 17311924]
10. Pan W, Arnone M, Kendall M, et al. Identification of peptide substrates for human MMP-11 (stromelysin-3) using phage display. *J Biol Chem.* 2003; 278(30):27820–7. [PubMed: 12738779]
11. Deng SJ, Bickett DM, Mitchell JL, et al. Substrate specificity of human collagenase 3 assessed using a phage-displayed peptide library. *J Biol Chem.* 2000; 275(40):31422–7. [PubMed: 10906330]
12. Kridel SJ, Chen E, Kotra LP, Howard EW, Mobashery S, Smith JW. Substrate hydrolysis by matrix metalloproteinase-9. *J Biol Chem.* 2001; 276(23):20572–8. [PubMed: 11279151]
13. Bieth JG. Theoretical and practical aspects of proteinase inhibition kinetics. *Methods Enzymol.* 1995; 248:59–84. [PubMed: 7674947]
14. Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 1990; 18(20):6097–100. [PubMed: 2172928]
15. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res.* 2004; 14(6):1188–90. [PubMed: 15173120]

Fig. 1.
The affinity tagged randomized peptide substrate fusion protein with the gene 3 protein (g3p) of filamentous M13 bacteriophage bound to anti-FLAG antibody . (*See Color Plates*)

Sample	Phage/mL x 10 ¹¹	A 490 nm
Start 1:20	0.214	0.428
Start 1:40	0.105	0.303
Start 1:80	0.050	0.202
Start 1:160	0.023	0.121
Start M2 1:20	0.004	0.025
Start M2 1:40	0.002	0.012
MMP-2 M2 1:20	0.013	0.076
MMP-2 M2 1:40	0.005	0.035
MMP-2 M2 1:80	0.002	0.010
MMP-2 M2 1:160	0.001	0.003

Phage/mL:

Start: 4.040 +/- 0.232 x 10¹¹/mL

Start M2: 0.080 +/- 0.000 x 10¹¹/mL (98% depletion)

MMP-2 M2 0.195 +/- 0.040 x 10¹¹/mL (95% depletion 2.4 x enrichment)

Fig. 2.

A typical standard curve for an M13 phage ELISA and the results of the first round of selection of MMP-2 substrates.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BLNK	6	14	22	30	38	+	6	14	22	30	38
							CTRL					
B	BLNK	7	15	23	31	39	+	7	15	23	31	39
							CTRL					
C	BLNK	8	16	24	32	40	+	8	16	24	32	40
							CTRL					
D	1	9	17	25	33	41	1	9	17	25	33	41
			+ PROTEASE						- PROTEASE			
E	2	10	18	26	34	42	2	10	18	26	34	42
F	3	11	19	27	35	43	3	11	19	27	35	43
G	4	12	20	28	36	44	4	12	20	28	36	44
H	5	13	21	29	37	45	5	13	21	29	37	45

Fig. 3.

Template for substrate phage ELISA. Bacterial cultures are applied to anti-M13 antibody coated 96-well plate at 100 μ L/well. Each culture is applied to the plate twice: on the left half of the plate and symmetrically on the right half. The phage captured on the left half of the plate is treated with protease (*shaded area*), while their counterparts on the right are not. The degree of hydrolysis is determined by subtracting the ratios of A_{490} between the protease treated and nontreated samples from 1.

Fig. 4.

Mapping of the position of the scissile bond by MALDI TOF MS. The amino acid sequence was determined by DNA sequencing of a PCR amplicons from an MMP-2 substrate phage infected bacterial culture followed by translation of the DNA sequence. The mass of the peptide released from 10^{12} MMP-2 substrate phage by MMP-2 was determined by MALDI TOF MS performed using Bruker Daltonics UltraFlex II mass spectrometer. (*See Color Plates*)

	P5	P4	P3	P2	P1	P1'	P2'	P3'
G	1.0443	1.0953	0.6594	0.7880	1.2216	0.0000	0.0000	0.9380
A	1.1629	1.1686	2.5989	2.4216	3.0451	0.0000	0.1562	2.2099
P	1.7153	0.6433	6.2323	0.0000	1.7865	0.0000	0.0000	0.0000
C	0.0000	0.7773	0.0000	0.0000	0.5263	0.0000	0.0000	0.0000
T	0.7203	0.7897	0.2410	0.2045	0.0000	0.0000	2.3218	0.9794
S	1.2667	0.6583	0.7563	0.8359	1.8858	0.0000	0.4945	0.6868
D	0.4686	0.0000	0.0000	0.1799	0.1799	0.0000	0.0000	0.2349
N	1.3567	0.9567	0.0000	0.4922	2.8299	0.0000	0.0000	0.5255
E	0.9691	1.8290	0.0000	0.5800	0.0000	0.0000	0.2858	0.6043
Q	0.5922	1.7768	0.0000	1.4764	0.3667	0.3667	0.9638	0.5907
K	0.3676	1.5645	0.0000	0.3790	0.1263	0.0000	1.2163	0.8699
H	0.9910	1.1307	0.4276	0.8938	5.5413	0.0000	0.0000	0.9001
R	1.2300	1.5049	0.0831	2.2120	0.2011	0.1340	3.1541	0.5378
V	0.9008	1.1364	2.6148	0.1858	0.0000	0.5572	0.5783	0.5983
I	0.4180	0.9013	0.8124	0.5337	0.0000	2.2787	0.8477	0.3116
L	0.5670	1.3711	2.4054	1.3978	0.0667	8.2521	0.4679	0.8476
M	0.6396	0.5876	1.3449	1.6718	2.3406	6.6864	2.5044	1.3941
F	0.1904	0.6862	2.1288	1.8252	1.3037	0.5214	0.7187	0.5800
Y	1.4578	1.3267	0.7444	1.2860	1.4147	0.9001	0.8065	0.4230
W	0.2632	0.4564	1.5908	2.3544	1.6817	2.3541	1.4311	0.0000

Fig. 5.
Position weight matrix of 234 substrates of MMP-16 .

Fig. 6.
WebLogo representation of the position weight matrix depicted in **Fig. 5**. (*See Color Plates*)