CHAPTER ONE

Integration of bioinformatic and chemoproteomic tools for the study of enzyme conservation in closely related bacterial species

Laura J. Keller[a], Markus Lakemeyer[b], and Matthew Bogyo[b,c,*]

[a]Department of Chemical & Systems Biology, Stanford University, Stanford, CA, United States
[b]Department of Pathology, Stanford University, Stanford, CA, United States
[c]Department of Microbiology and Immunology, Stanford University, Stanford, CA, United States
*Corresponding author: e-mail address: mbogyo@stanford.edu

Contents

Abstract

Activity-based protein profiling (ABPP) is a commonly utilized technique to globally characterize the endogenous activity of multiple enzymes within a related family. While it has been used extensively to identify enzymes that are differentially active across various mammalian tissues, recent efforts have expanded this technique to studying bacteria. As ABPP is applied to diverse sets of bacterial strains found in microbial communities, there is also an increasing need for robust tools for assessing the conservation of enzymes across closely related bacterial species and strains. In this chapter, we detail the integration of gel-based ABPP with basic bioinformatic tools to enable the analysis of enzyme activity, distribution, and homology. We use as an example the family of serine hydrolases identified in the skin commensal bacterium *Staphylococcus epidermidis*.

Methods in Enzymology, Volume 664
ISSN 0076-6879
https://doi.org/10.1016/bs.mie.2021.11.017

1. Introduction

The rise of resistant and multi-resistant bacterial pathogens poses a serious threat to human health. To overcome the antibiotic resistance crises, discovery of new classes of antibacterial targets is urgently needed. One of the many challenges in this endeavor is that large fractions of many bacterial genomes remain functionally uncharacterized, thus requiring unbiased screening methods to select promising new drug targets that can be chemically modulated. Activity-based protein profiling (ABPP) is a powerful technique that enables rapid and direct identification and quantification of enzyme activities in complex biological samples. An activity-based probe, which is composed of an electrophilic warhead, a linker to confer specificity, and a reporter tag, covalently modifies an enzyme active site in an activity-dependent manner. The combination of the warhead and linker dictates the family of enzymes targeted by the probe. Once target enzymes are labeled, a fluorescent tag allows for visualization of active enzymes via in-gel fluorescence or microscopy, while an affinity tag (e.g. biotin) facilitates enrichment and subsequent mass spectrometry-based identification of the enzymes (Keller, Babin, Lakemeyer, & Bogyo, 2020). Thus, ABPP enables monitoring of the activity of multiple enzymes in parallel in the context of their native cellular environments (Fig. 1A). Specific activity-based probes containing diverse electrophiles such as fluorophosphonates, acyloxymethyl ketones, vinyl sulfones, and others have been developed to target enzyme families ranging from proteases to glucosidases. For overviews of recent advances in activity-based probes, we refer the reader to comprehensive reviews (Benns, Wincott, Tate, & Child, 2021; Fang et al., 2021; Galmozzi, Dominguez, Cravatt, & Saez, 2014; Keller, Babin, et al., 2020).

One of the most successful classes of ABPs are fluorophosphonates (FP, Fig. 1B), which were designed to target the large and diverse family of mammalian serine hydrolases (Fig. 1C) that includes proteases and lipases, among others (Jessani et al., 2005; Kidd, Liu, & Cravatt, 2001). More recently, FP probes have been applied to characterize serine hydrolases in bacterial pathogens such as *Vibrio cholerae* or *Mycobacterium tuberculosis* (Hatzios et al., 2016; Ortega et al., 2016). In addition to determining which relevant enzyme activities to prioritize for future drug development, ABPP inherently selects for "druggable" enzymes that can be covalently modified. Moreover, it enables characterization of the potency and selectivity of an inhibitor within

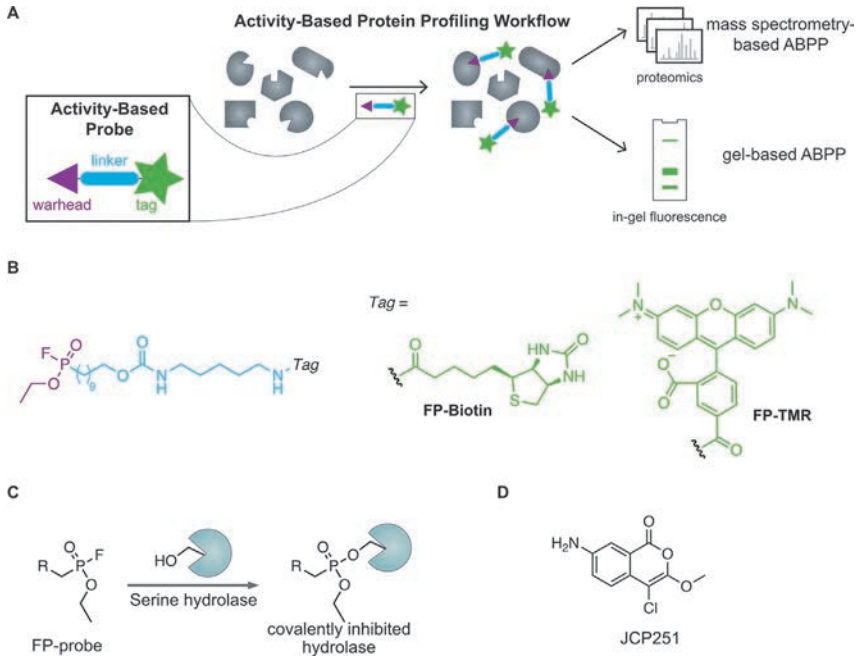


Fig. 1 Workflow for activity-based protein profiling (ABPP). (A) A sample of interest is treated with an activity-based probe, consisting of an electrophilic warhead, a linker than confers specificity, and a reporter tag. The probe labels enzymes in an activity-dependent manner, which can then be enriched with a biotin tag for identification by mass spectrometry (mass spectrometry-based ABPP) or visualized with a fluorescent tag via SDS-PAGE and in-gel fluorescence (gel-based ABPP). (B) Structure of fluorophosphonate (FP) probe with tetramethylrhodamine (TMR) and biotin tags, respectively. (C) Schematic for the covalent modification of serine hydrolases by FP probes. (D) Structure of SaFphB-targeted chloroisocoumarin JCP251 (Lentz et al., 2018).

a class of closely-related enzymes, as has been demonstrated in *M. tuberculosis* with compounds JCP276, EZ120, and AA692 (Babin et al., 2021; Lehmann et al., 2018; Li et al., 2021).

The human body harbors numerous defined communities of "beneficial" bacteria that comprise, for example, the gut, skin, or vaginal microbiota. The natural flora in these communities aids in immune development, contributes to xenobiotic metabolism, and competes with pathogens for ecological niches. Broad-spectrum antibiotics, which act on pathogen and commensal bacteria alike, can decrease microbial diversity, shift the gut metabolome, and even lead to recurring enteropathogen infections, for example by *Clostridioides difficile* (Avis, Wilson, Khan, Mason, & Powell, 2021; Ramirez et al., 2020). Thus, in the development of antibacterials, there has been a focus on targeting

pathogen-specific pathways. However, this approach is complicated in the case of pathobionts, which are normal residents of the healthy microbiome but have high pathogenic potential (Chow & Mazmanian, 2010; Jochum & Stecher, 2020; Mazmanian, Round, & Kasper, 2008). These pathobionts are particularly important therapeutic targets, as they are the leading causes of nosocomial infections and often acquire multi-drug resistance, yet they are difficult to selectively target because they have closely related species that are key members of the healthy microbiome (Kamada, Chen, Inohara, & Núñez, 2013; Otto, 2009; Wallace, Fishbein, & Dantas, 2020).

One such pathobiont is *Staphylococcus aureus* (*S. aureus*), which resides symbiotically in the nasal microbiome of 30% of adults but at the same time is a leading cause of bacteremia and infective endocarditis (Tong, Davis, Eichenberger, Holland, & Fowler, 2015). *S. aureus* is closely related to the commensal bacterium *Staphylococcus epidermidis* (*S. epidermidis*), which is frequently isolated from the skin of healthy humans and plays important roles in modulating the immune response during wound healing (Lai et al., 2009; Naik et al., 2012). Using activity-based protein profiling, we identified a serine hydrolase, FphB, in *S. aureus* that functions as a virulence factor and is important for colonization of the heart in a systemic infection mouse model. We identified a chloroisocoumarin-based compound, JCP251 (Fig. 1D), as a potent inhibitor of *S. aureus* FphB (SaFphB) by ABPP. To determine how selective the compound is for *S. aureus*, we performed gel-based ABPP against various other bacterial pathogens and the closely related *S. epidermidis* using a fluorescent derivative of JCP251. This analysis confirmed that while the other pathogen strains did not contain homologs of FphB that could be labeled by the JCP251 probe, we identified a probe-labeled protein in *S. epidermidis* (Lentz et al., 2018). Using a combination of mass spectrometry (MS)- and gel-based ABPP with a broad-spectrum FP probe, protein-protein BLAST, and cluster analysis, we were able to identify the set of serine hydrolases that are conserved between *S. aureus* and *S. epidermidis* as well as amongst clinical isolates of *S. epidermidis*. This allowed us to confirm the identity of the *S. epidermidis* FphB homolog (SeFphB) and its labeling by the SaFphB inhibitor JCP251. Despite the homology between SeFphB and SaFphB, JCP251 did not affect the skin colonization of *S. epidermidis* suggesting a possible alternate function for SeFphB that is not related to colonization (Keller et al., 2020). In this chapter, we will use our efforts in *S. epidermidis* to explain and outline basic bioinformatic tools and how to integrate them with gel-based ABPP to determine enzyme conservation and homology between strains and within closely related species of bacteria.

2. Integration of bioinformatics with ABPP

ABPP is a powerful tool to identify enzymes that are active in a disease-relevant state while also enabling screening for potent and selective inhibitors. ABPP has been successfully applied to pathogenic bacteria such as *Vibrio cholerae* and *Mycobacterium tuberculosis* (Babin et al., 2021; Hatzios et al., 2016; Lehmann et al., 2018; Li et al., 2021). Finding antibacterial targets that are selective for pathogens over beneficial commensal bacteria is relatively easy because close species relatives are not commonly found in the healthy flora of the human microbiota. On the other hand, pathobiont bacteria such as the Gram-positive species *Staphylococcus aureus*, *Clostridioides difficile*, and *Enterococcus faecalis* exist in the human body in complex communities with closely related beneficial bacteria (Kamada et al., 2013; Wallace et al., 2020). Combining bioinformatic approaches with ABPP can be an effective approach to identify promising drug targets in pathobiont species that are not conserved in closely related commensal bacterial species.

Genomic approaches, such as marker gene or whole genome sequencing, are state-of-the art approaches for the classification of clinical isolates of a single bacterial species (Conlan et al., 2012; Méric et al., 2018; Miragaia, Thomas, Couto, Enright, & de Lencastre, 2007). The combination of gel-based ABPP and quantitative image analysis represents a straight-forward and fast alternative to genomic-only techniques. Moreover, focusing on the dynamic regulation of the catalytic activity of an enzyme family, such as serine hydrolases, enables the generation of functional clusters that might better reflect biological conservation across bacterial strains compared to simple sequence homology.

2.1 Useful bioinformatic tools to integrate with ABPP

The most common method of identifying a protein homolog is assessing sequence similarity with the protein-protein Basic Local Alignment Search Tool (BLASTp), which rapidly compares a query protein sequence to a search set of sequences by measuring local similarity (Altschul, Gish, Miller, Myers, & Lipman, 1990). The set of sequences to search against can be composed of a single sequence, a set of sequences, the proteome of a single species, or even entire databases such as all non-redundant GenBank sequences. The BLASTp search returns the protein sequences most similar to the query sequence, their pair-wise alignments with the query sequence, and several metrics for each alignment. These include

the percent identity, which is the fraction of the aligned sequences that have the same amino acid in the same position, and the e-value, which is a statistical measure of the likelihood that another protein sequence in the search set would have a better similarity score by random chance. There is no standard threshold for what percent identity and e-value constitutes a homolog, but combining these metrics with searching the alignment for conservation of residues that are known to be important for protein function, such as active-site residues, can increase confidence in identifying a homolog. Additionally, if the two proteins are reciprocal best hits, meaning that each protein is the top result in a BLASTp search of the opposite proteome, this further confirms that two proteins are likely homologs (Ward & Moreno-Hagelsieb, 2014). All of these approaches can be applied to compare protein identifications in MS-based ABPP experiments across closely related bacterial species and are the starting point for more complex bioinformatic-based analyses.

On the other hand, gel-based ABPP (i.e. labeling samples of interest with fluorescent probes, separating the proteins on an SDS-PAGE gel, and measuring in-gel fluorescence) is an application of ABPP that does not as naturally lend itself to bioinformatic supplementation. However, with image quantification software such as ImageJ, the fluorescence intensity of individual protein bands on a gel can be measured and compared across conditions, allowing relative quantification of the activity of each labeled species. This method is commonly used in competitive gel-based ABPP, where a library of inhibitors is screened by pretreating the sample of interest (i.e. intact bacterial cells) with a compound followed by labelling with a fluorescent probe. The ability of the inhibitor to compete away the probe labeling of a given enzyme can be then quantified, allowing determination of IC_{50} values (Chen, Keller, Cordasco, Bogyo, & Lentz, 2019). Alternatively, the fluorescent intensities can be measured along an entire lane, thus generating a profile of enzymatic activity. This enables comparison of global enzyme activities for a given family across different conditions or within related strains. It also provides information on how each enzyme may be differentially regulated. These activity profiles can be visualized as a heatmap and clustered hierarchically with existing packages in R or Python. This allows any experimental sample (i.e. clinical isolates) to be rapidly stratified and compared to other strains in a functional manner.

2.2 Application of bioinformatic tools and ABPP techniques for profiling *Staphylococcus epidermidis* serine hydrolases

Using a fluorophosphonate activity-based probe, we have previously identified a serine hydrolase in the pathobiont *Staphylococcus aureus*, SaFphB, that

is important in a systemic infection mouse model. The protein is selectively targeted in *S. aureus* by the covalent inhibitor JCP251. However, this study also demonstrated that JCP251 labels an enzyme of similar molecular weight in the closely related skin commensal *S. epidermidis* (Lentz et al., 2018). We thus aimed to determine whether this enzyme was a homolog of SaFphB and how conserved the 11 additional fluorophosphonate-binding hydrolases (Fphs) identified in *S. aureus* are among different species of the *Staphylococcus* genus (Keller, Lentz, et al., 2020). By performing BLASTp searches on custom search sets of the reference proteome for each major staphylococcal strain, we were able to determine that most of the 12 *S. aureus* Fph enzymes were conserved in at least 60% of the almost 50 staphylococcal species tested (Fig. 2A). In particular, homologs of SaFphB were predicted in approximately two-thirds of the species (Fig. 2B). Interestingly, all but two of the Fphs from *S. aureus* have at least one putative homolog in *S. epidermidis*, and three have two homologs, suggesting that there may have been either multiple gene duplication or horizontal gene transfer events.

Fphs are a subset of the serine hydrolase enzyme superfamily, which use a catalytic serine residue to hydrolyze amide, ester, and thioester bonds in metabolites that are important for cell-cell signaling and metabolism (Simon & Cravatt, 2010). Because serine hydrolases are defined by their catalytic mechanism, ABPP techniques with a probe containing a broad serine-reactive warhead such as the fluorophosphonate is an effective method to identify novel serine hydrolases. Importantly, identification of probe-labeled enzymes helps to confirm bioinformatically-predicted annotations. Thus, we also followed up our BLASTp-based Fph homolog predictions with MS-based ABPP of 17 strains of *S. epidermidis*, including reference strains and clinical isolates. In total, we identified 18 serine hydrolases, most of which were the BLASTp-predicted *S. aureus* Fph homologs. Interestingly, four serine hydrolases had not been identified in the profiling of *S. aureus* and one did not have a clear homolog in *S. aureus*, demonstrating how combining bioinformatics and ABPP across multiple related species can enhance the identification of novel enzymes and uncover their conservation within related families of bacteria. In particular, the homolog of SaFphB in *S. epidermidis*, SeFphB, was predicted with BLASTp, was measured by MS-based ABPP, and its similar catalytic function was confirmed by active site conservation and substrate specificity.

To compare the relative activities of the Fphs across the various *S. epidermidis* strains, we performed gel-based ABPP with a fluorescent fluorophosphonate probe (Fig. 3A). We identified substantial variabilities in the activity levels of multiple Fphs in these 17 *S. epidermidis* strains, which

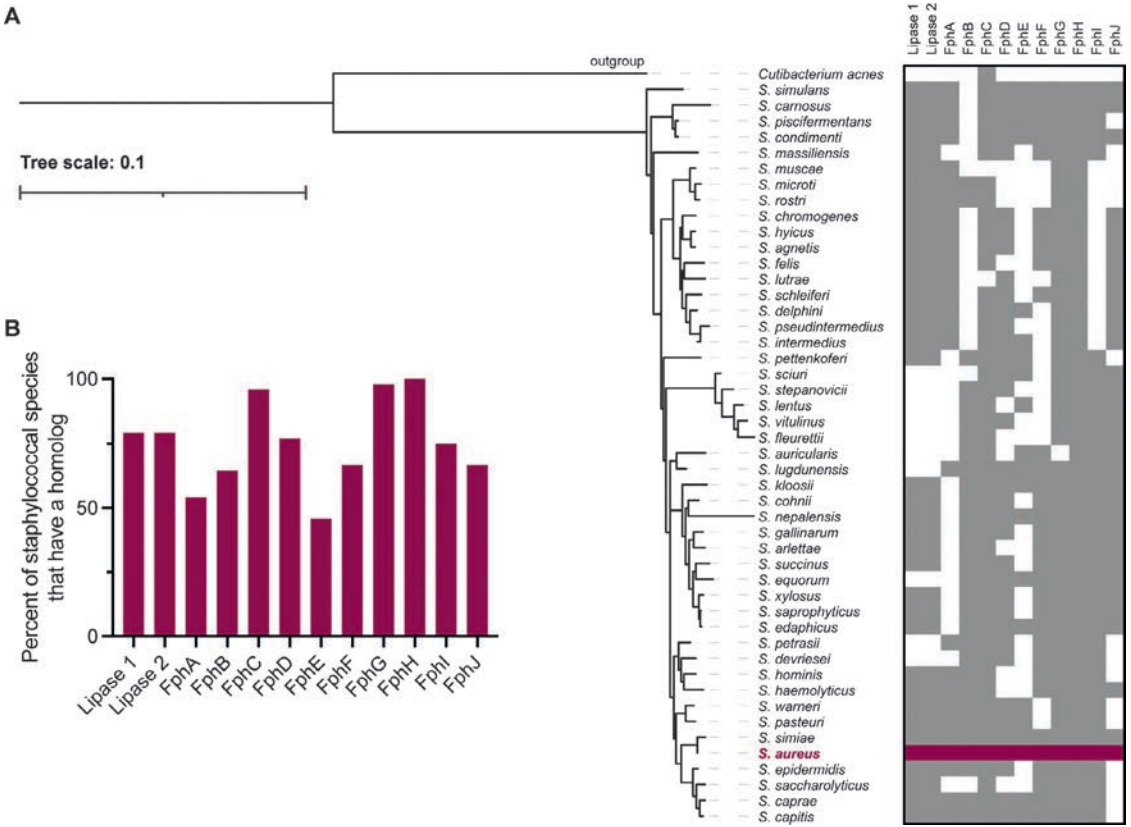


Fig. 2 Conservation of Fph enzymes throughout the genus *Staphylococcus*. (A) Phylogenetic tree of 47 staphylococcal species, with the skin bacterium *Cutibacterium acnes* as the outgroup, by 16S rRNA sequences. Each filled grey cell in the heatmap indicates that species has a homolog of the *S. aureus* Fph, as determined by a threshold e-value of 1×10^{-40}. *Staphylococcus aureus* and its Fph enzymes, which were the query sequences for each BLASTp search, are highlighted in maroon. (B) Percentage of the 47 staphylococcal species that contain a homolog of the indicated *S. aureus* hydrolase by BLASTp search. *Part B: Adapted from Keller, L. J., Lentz, C. S., Chen, Y. E., Metivier, R. J., Weerapana, E., & Fischbach, M. A. (2020). Characterization of serine hydrolases across clinical isolates of commensal skin bacteria Staphylococcus epidermidis using activity-based protein profiling.* ACS Infectious Diseases, 6*(5), 930–938. doi: 10.1021/acsinfecdis.0c00095*.

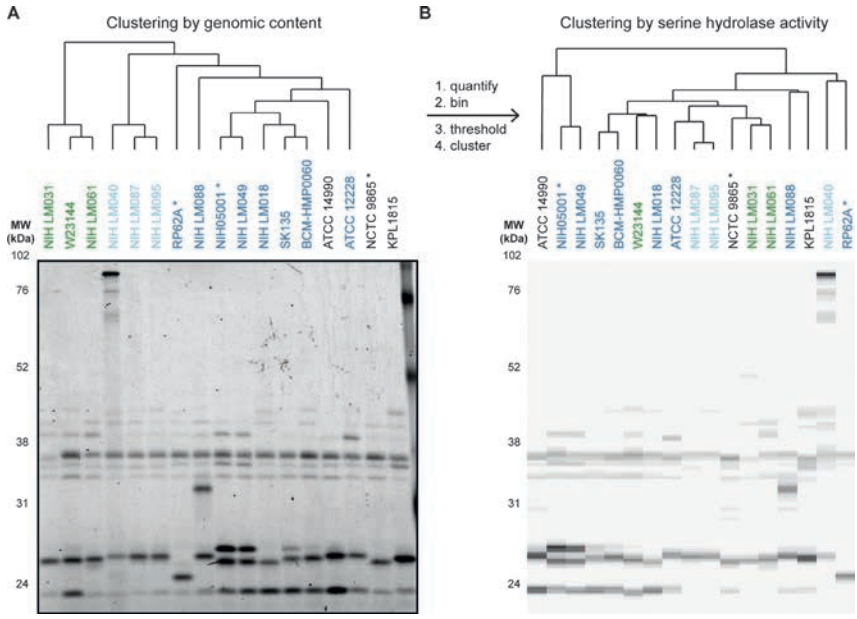


Fig. 3 Method of clustering bacterial strains by their activity profiles. (A) Clinical isolates of *Staphylococcus epidermidis* were labeled with 1 μM fluorophosphonate-tetramethylrhodamine before being lysed and separated by SDS-PAGE. The strains are clustered by a phylogenetic analysis of their genomic content (Chen et al., 2019; Gardner, Slezak, & Hall, 2015). The genomes of isolates NCTC 9685 and KPL1815 have not been sequenced and assembled and thus are not clustered. (B) The Fph activity profile of each clinical isolate was quantified, visualized in the heatmap, and used for hierarchical clustering. Clinical isolates are colored by their clonal lineages (group A, dark blue; group B, green; group C, light blue; unknown, black (Méric et al., 2018)) and strains isolated from infected patients are denoted with an asterisk. *Part B: Adapted from Keller, L. J., Lentz, C. S., Chen, Y. E., Metivier, R. J., Weerapana, E., & Fischbach, M. A. (2020). Characterization of serine hydrolases across clinical isolates of commensal skin bacteria Staphylococcus epidermidis using activity-based protein profiling.* ACS Infectious Diseases, 6*(5), 930–938. doi: 10.1021/acsinfecdis.0c00095.*

may be due to the fact that these strains are isolated from different microenvironments in the human body (i.e. the nares of a healthy volunteer versus the bloodstream of an infected patient). There have been efforts to categorize *S. epidermidis* clinical isolates into clonal lineages via marker gene and genomic-based techniques, as well as attempts to correlate this clustering with functional assays (Conlan et al., 2012; Espadinha et al., 2019; Méric et al., 2018; Miragaia et al., 2007). Previous work from the Bertozzi group utilized an in-gel assay with a sulfatase-activatable fluorophore to distinguish different mycobacterial species based on their activity "fingerprint" (Beatty et al., 2013). We thus wanted to determine if the Fph activity "fingerprint" or

profile could be used to stratify these 17 *S. epidermidis* strains. Clustering the strains based on their Fph activity profile maintained the close relationship between several pairs of strains, such as SK135 and BCM-HMP0060, but did not preserve the global genomic-based dendrogram nor distinguish the clonal lineages (Fig. 3B). These results were not surprising considering almost all the strains have the same Fph enzymes encoded in their genomes; however, they do suggest that there are interesting strain-specific regulatory pathways of these enzymes that should be further explored.

3. Protocols

There are many useful protocols for the design of activity-based probes and their implementation in various ABPP techniques (Galmozzi et al., 2014; Prothiwa & Böttcher, 2020; Sharifzadeh et al., 2020; Steiger, Fansler, Whidbey, Miller, & Wright, 2020). This section will detail general ABPP sample preparation, gel-based analysis to visualize differences in enzyme activity, and basic bioinformatic tools that can be integrated with ABPP.

3.1 Intact bacteria labeling with an activity-based probe

3.1.1 Materials

- Bacterial cultures
- Phosphate-buffered saline (PBS)
- Fluorophosphonate-tetramethylrhodamine (FP-TMR) activity-based probe at a concentration of 100 μM in DMSO
- 10% sodium-dodecyl sulfate (SDS) in water
- Non-sterile 1.5 mL Eppendorf tubes or equivalent
- 0.5 mL screw cap microcentrifuge tubes with O-ring (Thermo Fisher #02-707-350) or equivalent
- 0.1 mm glass beads (BioSpec #11079101) or equivalent
- Spectrophotometer
- Bench top incubator
- Bench top microcentrifuge
- Mini-beadbeater-96 (BioSpec) or equivalent
- Aluminum vial rack (BioSpec #702ALU) or equivalent

3.1.2 Procedure

1. Prechill aluminum vial rack by storing at -20 °C.
2. Measure the optical density of the bacterial cultures at 600 nm (OD_{600}) using the spectrophotometer.

3. Centrifuge the bacterial cultures on the bench top microcentrifuge at 3,000 x *g* for 5 min, remove the culture medium, and resuspend the cell pellets in a volume of PBS such that the final OD_{600} is 16.
4. Add 2 μL FP-TMR probe (100 μM stock) to each 200 μL bacterial culture (1 μM final concentration) and incubate at 37 °C for 30 min.
5. Centrifuge the bacterial cultures on the bench top microcentrifuge at 3,000 x *g* for 5 min, remove the supernatant, and resuspend the labeled cell pellets in the same volume of PBS.
6. Add 10% SDS at a 1:100 dilution (0.1% final concentration).
7. Transfer the samples to 0.5 mL O-ringed tubes filled halfway with 0.1 mm glass beads.
8. Put samples in prechilled aluminum vial rack and lyse by bead-beating, three times for 45 s with 2 min breaks on ice in between each burst.
9. Centrifuge the bacterial lysate at 8,000 x g for 10 min to pellet insoluble cell debris and beads. Transfer the lysate to a new 1.5 mL tube.
10. Quantify protein concentration using the Pierce BCA Protein Assay Kit (Thermo Fisher), following manufacturer's protocol.

3.1.3 Notes

This protocol is written specifically for labeling with a fluorescent FP probe to label serine hydrolases, but the concentrations and incubation times should be optimized for each probe. This can be achieved by first labeling with a range of probe concentrations for a set time and then selecting optimal probe concentrations and performing labeling for varying time periods. Furthermore, for each bacterial species, we recommend optimizing with several different lysis conditions prior to undergoing any full-scale experiment. In our experience, lysis by probe sonication (i.e. three 20-s bursts with the Sonic Dismembrator Model 100 (Fisher Scientific)) is preferred with smaller volume samples; however, lysis by bead-beating is often necessary with Gram-positive bacterial species. Furthermore, various detergent additives and concentrations should be screened to optimize solubilization of membrane-associated proteins.

To identify enzymes that are more likely to be accessible to small molecule inhibitors and thus potential drug targets, we label intact bacterial samples, which, depending on the cell permeability of the activity-based probe used, may bias the labeling toward enzymes that are membrane-associated or found on the external surface of the bacteria cells. Bacterial cells can also be lysed prior to labeling if intracellular targets are of interest—in

this case, normalize the protein amount by BCA quantification prior to labeling with the activity-based probe.

If the activity-based probe contains an alkyne handle instead of the fluorescent tag, after sample labeling, it can be derivatized with copper-catalyzed azide alkyne cycloaddition (CuAAC) to append a fluorophore (Rostovtsev, Green, Fokin, & Sharpless, 2002; Steiger et al., 2020).

Similar labeling methods can be performed with biotinylated probes to perform mass spectrometry-based identification of the active enzymes. Because methods vary based on instrumentation, a detailed protocol is not provided in this chapter. The general workflow involves (1) labeling the bacteria with a biotinylated activity-based probe, (2) lysing the bacteria, (3) removing excess probe, (4) enriching for labeled enzymes on streptavidin resin, (5) on-bead reducing, alkylating, and digesting the enriched proteins, (6) eluting from the resin, (7) desalting, and (8) MS-measurement. For more detailed descriptions, we refer the reader to previous methods sections from the following papers (Chen et al., 2017; Galmozzi et al., 2014; Kuljanin et al., 2021).

3.2 Gel-based ABPP

3.2.1 Materials

- Fluorescent probe-labeled bacterial samples (from Section 3.1)
- PBS
- 4x Laemmli Sample Buffer (1 mL 1.0 M Tris pH 6.8, 400 mg sodium dodecyl-sulfate, 2 mL glycerol, 1 mL β-mercaptoethanol, 10 mg bromophenol blue, dH_2O to a final volume of 5 mL)
- Precision Plus Protein All Blue Prestained Protein Standards (Bio Rad #1610373) or equivalent
- Isopropanol
- 12% acrylamide Resolving Gel (2.4 mL ProtoGel (30% (w/v) acrylamide/bisacrylamide solution (National Diagnostics #EC-890)), 2 mL dH_2O, 1.5 mL 1.5 M Tris pH 8.8, 60 µL 10% SDS in dH_2O, 60 µL 10% ammonium persulfate (APS) in dH_2O, 6 µL TEMED)
- Stacking Gel (750 µL ProtoGel, 2.25 mL dH_2O, 1.75 mL 1 M Tris pH 6.8, 50 µL 10% SDS in dH_2O, 50 µL 10% ammonium persulfate (APS) in dH_2O, 5 µL TEMED)
- 10x Tris-glycine Running Buffer (30 g Tris base, 144 g glycine, 10 g SDS, dH_2O to a final volume of 1 L)
- Standard Dry Block Heater (VWR) or equivalent
- Mini PROTEAN Tetra Cell system (Bio Rad #1658003) or equivalent

- Mini Cell Buffer Dam (Bio Rad #1653130) or equivalent
- Power supply (Thermo Fisher #EC105) or equivalent
- Typhoon FLA 9500 gel scanner (GE) or equivalent

3.2.2 Procedure

1. Dilute each fluorescent probe-labeled sample to 25 μg total protein in 15 μL using PBS.
2. Add 5 μL 4x Laemmli Sample Buffer to each sample, pipet to mix, and boil in a preheated Block Heater for 10 min at 95 °C.
3. Assemble the short and 1.0 mm spacer plate in the casting frame, ensuring the short plate is facing out and that the bottoms of the plates are flush with the bench, and transfer the assembly to the casting stand.
4. Combine the reagents for the 12% acrylamide Resolving Gel and pipet between the glass plates to roughly 75% of the height of the short plate, checking for leaks.
5. Add a thin layer of isopropanol on top of the Resolving Gel layer to ensure there are no bubbles and the top of the gel is even, and allow to polymerize for at least 15 min.
6. Pour off isopropanol into a receptacle for organic waste and rinse with dH_2O before proceeding.
7. Combine the reagents for the Stacking Gel and pipet between the glass plates to the top. Carefully place the 15-well comb in between the glass plates without introducing bubbles and allow the Stacking Gel to polymerize for at least 15 min.
8. Dilute 100 mL 10x Tris-glycine Running Buffer in 900 mL dH_2O and add 400 mL of 1x Running Buffer to the Mini-PROTEAN Tetra Cell Buffer Tank.
9. Remove the comb from the acrylamide gel.
10. Sandwich the acrylamide gel with the spacer plate facing out and the Mini Cell Buffer Dam in the Mini-PROTEAN Tetra Electrode Assembly, ensuring the seal with the rubber gasket. If running two gels simultaneously, the Buffer Dam can be swapped for a second acrylamide gel.
11. Fill the sandwiched assembly with 1x Running Buffer, rinse the wells with 1x Running Buffer, and place in the Buffer Tank.
12. Load 20 μL of labeled sample into each well in the desired order. Add 3 μL of Precision Plus Protein All Blue Prestained Protein Standards to either the first or last lane. When loading each well, pipet slowly and

smoothly, ensuring the sample collects at the bottom of the well and there are no bubbles to disrupt the sample.

13. Put the Lid on the Buffer Tank, attach the leads to the Power Supply, and set the voltage at approximately 85 V. Once the sample has passed through the Stacking Gel and entered the Resolving Gel (as monitored by the blue "front" of the samples), the voltage can be increased up to 120 V if desired.
14. Once the blue "front" has run off the bottom of the gel, remove the gel from the sandwiched assembly, rinse with water, and gently dry off.
15. Image the gel using the Typhoon gel scanner at the appropriate wavelength (for the TMR fluorophore, excitation/emission 544 nm/570 nm) for the fluorescent dye in the activity-based probe to generate an in-gel fluorescence image. The photomultipler tube (PMT) voltage and resolution can be adjusted to increase image quality.

3.2.3 Notes

When casting the acrylamide gel, add the acrylamide just prior to casting the gel as polymerization will begin right away. Avoid any spills as acrylamide in solution is toxic. The percent acrylamide in the resolving gel can be tuned depending on the molecular weight of the protein of interest, with lower percentages such as 10% being better for resolving higher molecular weight proteins and higher percentages such as 15% being better for resolving lower molecular weight proteins (the volume of water is thus increased or decreased accordingly to maintain the same final volume). Additionally, proteins that are difficult to resolve can be better separated by running the gel at lower voltages, i.e. 65 V. The recipes above make one gel but can be scaled up to cast multiple gels. Acrylamide gels can be stored inside the glass plates for up to a week at 4 °C, wrapped in a damp paper towel and cling wrap, prior to use.

3.3 Protein-protein BLAST analysis

3.3.1 Workflow

1. Visit the web-based BLAST tool at https://blast.ncbi.nlm.nih.gov/Blast.cgi and navigate to the Protein BLAST functionality (BLASTp).
2. Enter query sequence(s) of interest. The sequences can be formatted as a list of GenBank accession numbers or as FASTA sequences, or an entire . FASTA file can be uploaded. If entering multiple queries, each accession number or sequence must be on a new line. The single letter amino acid code must be used for BLASTp.

3. Choose a search space. The default is the non-redundant protein sequence (nr) database, which contains all GenBank protein translations. This space can be narrowed at various taxonomic levels to all proteins associated with a single phylum, genus, species, or strain (Firmicutes, *Staphylococcus*, *Staphylococcus epidermidis*, *Staphylococcus epidermidis* ATCC 14990, respectively, for example).
4. Select the blastp algorithm and press BLAST.
5. The results include a list of sequences with significant alignments with a hyperlink to its entry in GenBank, metrics like e-value and percent identity, and a visualization of the alignment, including which residues are conserved.

3.3.2 Notes

While the standard algorithm parameters on the web-based BLASTp tool are often sufficient, they can be fine-tuned if needed, i.e. the e-value threshold or the matrix used to calculate the score for a given alignment. The web-based tool can be utilized for many BLASTp searches, but BLAST+ applications can also be downloaded locally onto a server and searches can be performed using command-line tools. Local BLAST+ searches are especially useful for high-volume searches or developing custom target search sets to search against, such as proteomes from the Human Microbiome Project (Consortium et al., 2012). For more detailed instructions, visit the BLAST Command Line Applications User Manual (https://www.ncbi.nlm.nih.gov/books/NBK279690/).

In brief, proteins of interest to be searched against (e.g., an entire proteome) can be downloaded into a .FASTA file from NCBI or other similar databases to create a search set. After installing the standalone BLAST+ programs, the *makeblastdb* function converts the search set .FASTA file into a BLAST database with the Unix command "*makeblastdb -in searchDB.fasta -out blastDB -dbtype prot -parse_seqids*." Then, the query sequence(s), also saved in a .FASTA file, can be used in a BLASTp search and the results saved to an output file with the Unix command "*blastp -query query.fasta -db blastDB -outfmt 6 > output.txt*." The output file will be saved in a tabular form containing the query sequence ID, the search sequence ID, percent identity, alignment length, number of mismatches, number of gap openings, residue number for the start of the alignment in the query sequence, residue number for the end of the alignment in the query sequence, residue number for the start of the alignment in the search sequence, residue number for the end of the alignment in the search sequence, e-value, and bitscore, which is

another measurement of the confidence of the alignment but independent of the size of the search set. Additional modifications of the BLASTp search can include setting a maximum number of resulting sequences for each query sequence (for one sequence per query, *"-max_target_seqs 1"*) or a threshold of the minimal e-value (for an e-value threshold of 1e-40, *"-evalue 1e-40"*).

3.4 Activity profile clustering

3.4.1 Workflow

1. Download an ImageJ image processing package from https://imagej.net/software/ and install (Schneider, Rasband, & Eliceiri, 2012). There are many iterations of the original software, ImageJ, most of which are sufficient for standard image analyses. For the purposes of this chapter, exact details will be based on the Fiji package (Schindelin et al., 2012).
2. Open the in-gel fluorescence image (e.g. in the .tiff or .gel format) from gel-based ABPP in Fiji. Ensure that the gel is perfectly horizontally aligned for the cleanest analysis. The image can be rotated using the *Image* > *Transform* > *Rotate* function.
3. If the image has dark bands on a light background, invert using the *Edit* > *Invert* function (Fig. 4, Step i).
4. Using the *Straight Line* tool, draw a vertical line through a lane (Fig. 4, Step ii).
5. Using the *Analyze* > *Plot Profile* function, calculate the fluorescent intensity along the vertical line (Fig. 4, Step iii) and save the data with the *Data* > *Copy All Data* function (Fig. 4, Step iv). The copied data can be pasted into an Excel file or similar.
6. Move the vertical line to another lane by using the arrow keys, preserving the y-value of the top point of the line and the length of the line, and measure the fluorescent profile along that line. Continue for every lane of interest. Additionally, measure the fluorescent profile along a region outside of the lanes to serve as a background profile. The resulting Excel file should contain the "Gray Value," or intensity, at each pixel for every lane, including the background, and be saved as a .CSV file.
7. To smooth the effects of any blemish in the image, the fluorescence profile can be divided into roughly 100 bins by summing across the fluorescent intensities in each bin. As highlighted in the sample code (see Fig. 5, ##BIN), every 15 intensity values of the activity profile

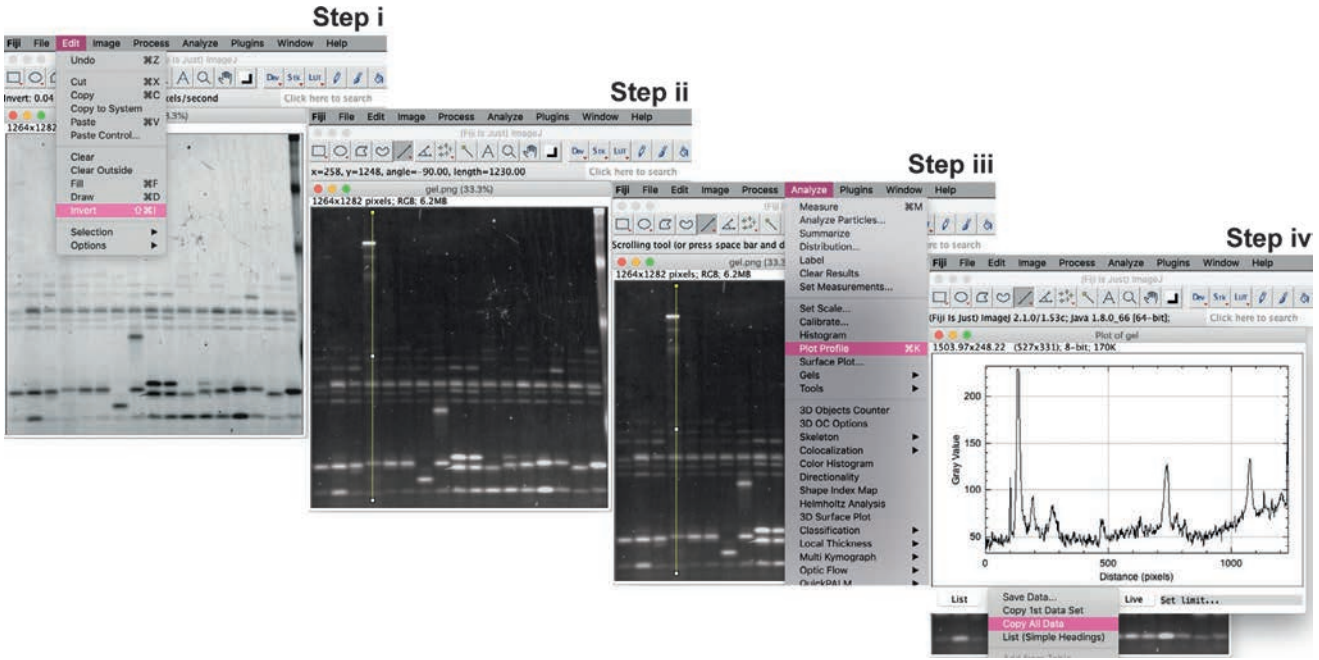
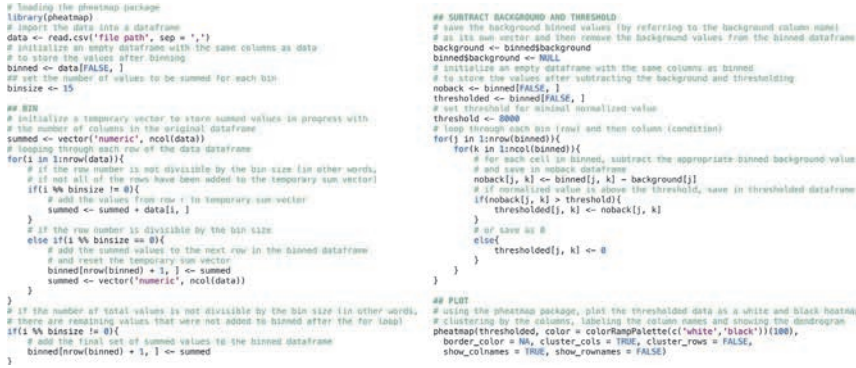


Fig. 4 Quantification of activity profiles from gel-based ABPP using the ImageJ software, Fiji. Step i, *Invert*, if the image has dark bands on a light background; Step ii, draw a *Straight Line* vertically through a lane; Step iii, *Plot Profile* along the vertical line; Step iv, *Copy All Data* to paste in Excel and save the intensity values for the profile.



```
# loading the pheatmap package
library(pheatmap)
# import the data into a dataframe
data <- read.csv('file path', sep = ',')
# initialize an empty dataframe with the same columns as data
# to store the values after binning
binned <- data[FALSE, ]
## set the number of values to be summed for each bin
binsize <- 15

## BIN
# initialize a temporary vector to store summed values in progress with
# the number of columns in the original dataframe
summed <- vector('numeric', ncol(data))
# looping through each row of the data dataframe
for(i in 1:nrow(data)){
    # if the row number is not divisible by the bin size (in other words,
    # if not all of the rows have been added to the temporary sum vector)
    if(i %% binsize != 0){
        # add the values from row i to temporary sum vector
        summed <- summed + data[i, ]
    }
    # if the row number is divisible by the bin size
    else if(i %% binsize == 0){
        # add the summed values to the next row in the binned dataframe
        # and reset the temporary sum vector
        binned[nrow(binned) + 1, ] <- summed
        summed <- vector('numeric', ncol(data))
    }
}
# if the number of total values is not divisible by the bin size (in other words,
# there are remaining values that were not added to binned after the for loop)
if(i %% binsize != 0){
    # add the final set of summed values to the binned dataframe
    binned[nrow(binned) + 1, ] <- summed
}
```

```
## SUBTRACT BACKGROUND AND THRESHOLD
# save the background binned values (by referring to the background column name)
# as its own vector and then remove the background values from the binned dataframe
background <- binned$background
binned$background <- NULL
# initialize an empty dataframe with the same columns as binned
# to store the values after subtracting the background and thresholding
noback <- binned[FALSE, ]
thresholded <- binned[FALSE, ]
# set threshold for minimal normalized value
threshold <- 8000
# loop through each bin (row) and then column (condition)
for(j in 1:nrow(binned)){
    for(k in 1:ncol(binned)){
        # for each cell in binned, subtract the appropriate binned background value
        # and save in noback dataframe
        noback[j, k] <- binned[j, k] - background[j]
        # if normalized value is above the threshold, save in thresholded dataframe
        if(noback[j, k] > threshold){
            thresholded[j, k] <- noback[j, k]
        }
        # or save as 0
        else{
            thresholded[j, k] <- 0
        }
    }
}

## PLOT
# using the pheatmap package, plot the thresholded data as a white and black heatmap,
# clustering by the columns, labeling the column names and showing the dendrogram
pheatmap(thresholded, color = colorRampPalette(c('white','black'))(100),
  border_color = NA, cluster_cols = TRUE, cluster_rows = FALSE,
  show_colnames = TRUE, show_rownames = FALSE)
```

Fig. 5 Sample R code for processing and visualization of in-gel fluorescence imaging as a clustered heatmap.

were summed for each lane in the gel; however, the exact number and size of each bin should be tailored for each gel.

8. Subtract the binned background fluorescent profile from every other profile (see Fig. 5, ##SUBTRACT BACKGROUND AND THRESHOLD).
9. As minor changes in the gel background do not provide valuable information for the purposes of clustering, any values below a background threshold are reassigned a value of 0 (see Fig. 5, ##SUBTRACT BACKGROUND AND THRESHOLD). The appropriate background threshold to use will vary from gel to gel, but should correspond to the background subtracted binned value from a portion of the gel that has no visual bands.
10. Visualize the binned and thresholded quantification of the image utilizing the pheatmap package in R, with the rows being each bin along the gel and the columns being each condition in the gel. Use hierarchical clustering within the pheatmap package on the columns and show the corresponding dendrogram (see Fig. 5, ## PLOT).

3.4.2 Notes

The data processing can be performed manually in Excel, but use of computational platforms such as RStudio (https://www.rstudio.com/products/rstudio/download/) allow more easy manipulation and modification of parameters including bin size and background threshold. In either case, the heatmap and dendrogram generation should be performed with R, Python, or a similar bioinformatically-inclined coding language. If the data

processing is performed computationally, the conversion of the in-gel fluorescence image to a heatmap can be optimized for each gel by altering the size/number of bins and the background threshold by monitoring the changes in the outputted heatmap, ensuring that each distinct band visualized in the gel is represented in the heatmap. Included is some sample code in R for the data processing and visualizing that should be edited and embellished as needed (Fig. 5).

4. Summary

Activity-based protein profiling is a powerful approach for identifying relevant and druggable bacterial targets. In the context of studying pathobiont bacteria that exist with closely related species in the healthy microbiome, this technique can be supplemented with bioinformatic tools such as BLASTp, which can assess enzyme conservation across species, to understand which commensal bacteria might be impacted by targeting a given enzyme. Furthermore, quantification of activity profiles generated with gel-based ABPP and image analysis software enables comparison and stratification of strains of a single bacterial species, thus enabling rapid and functional clustering of clinical isolates. In this chapter, we have detailed the protocols necessary to perform BLASTp searching, gel-based ABPP, image analysis, and clustering to identify related enzymes and functional conservation across related bacterial species. As our reference example, we have used our data generated by the characterization of serine hydrolase conservation in the skin commensal bacterium *Staphylococcus epidermidis*. We are confident that this integration of approaches has great potential to be applied to study additional classes of enzymes and their activity profiles in a wide variety of clinically relevant bacteria.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Avis, T., Wilson, F. X., Khan, N., Mason, C. S., & Powell, D. J. (2021). Targeted microbiome-sparing antibiotics. *Drug Discovery Today*. https://doi.org/10.1016/j.drudis.2021.07.016.

Babin, B. M., Keller, L. J., Pinto, Y., Li, V. L., Eneim, A. S., Vance, S. E., et al. (2021). Identification of covalent inhibitors that disrupt M. tuberculosis growth by targeting multiple serine hydrolases involved in lipid metabolism. *Cell Chemical Biology*. https://doi.org/10.1016/j.chembiol.2021.08.013.

Beatty, K. E., Williams, M., Carlson, B. L., Swarts, B. M., Warren, R. M., Van Helden, P. D., et al. (2013). Sulfatase-activated fluorophores for rapid discrimination

of mycobacterial species and strains. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(32), 12911–12916. https://doi.org/10.1073/pnas.1222041110.
Benns, H. J., Wincott, C. J., Tate, E. W., & Child, M. A. (2021). Activity-and reactivity-based proteomics: Recent technological advances and applications in drug discovery. *Current Opinion in Chemical Biology*. https://doi.org/10.1016/j.cbpa.2020.06.011.
Chen, L., Keller, L. J., Cordasco, E., Bogyo, M., & Lentz, C. S. (2019). Fluorescent triazole urea activity-based probes for the single-cell phenotypic characterization of *Staphylococcus aureus*. *Angewandte Chemie International Edition*, *58*(17), 5643–5647. https://doi.org/10.1002/anie.201900511.
Chen, X., Wong, Y. K., Wang, J., Zhang, J., Lee, Y.-M., Shen, H.-M., et al. (2017). Target identification with quantitative activity based protein profiling (ABPP). *Proteomics*, *17*(3–4), 1600212. https://doi.org/10.1002/pmic.201600212.
Chen, Y. E., Bouladoux, N., Hurabielle, C., Mattke, A. M., Belkaid, Y., & Fischbach, M. A. (2019). Decoding commensal-host communication through genetic engineering of Staphylococcus epidermidis. *BioRxiv*, *664656*. https://doi.org/10.1101/664656.
Chow, J., & Mazmanian, S. K. (2010). A pathobiont of the microbiota balances host colonization and intestinal inflammation. *Cell Host & Microbe*, 7(4), 265–276. https://doi.org/10.1016/j.chom.2010.03.004.
Conlan, S., Mijares, L. A., NISC Comparative Sequencing Program, Becker, J., Blakesley, R. W., Bouffard, G. G., et al. (2012). Staphylococcus epidermidis pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. *Genome Biology*, *13*(7), R64. https://doi.org/10.1186/gb-2012-13-7-r64.
Consortium, T. H. M. P, Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*(7402), 207–214. https://doi.org/10.1038/nature11234.
Espadinha, D., Sobral, R. G., Mendes, C. I., Méric, G., Sheppard, S. K., Carriço, J. A., et al. (2019). Distinct phenotypic and genomic signatures underlie contrasting pathogenic potential of staphylococcus epidermidis clonal lineages. *Frontiers in Microbiology*, *10*(1971). https://doi.org/10.3389/fmicb.2019.01971.
Fang, H., Peng, B., Ong, S. Y., Wu, Q., Li, L., & Yao, S. Q. (2021). Recent advances in activity-based probes (ABPs) and affinity-based probes (AfBPs) for profiling of enzymes. *Chemical Science*. https://doi.org/10.1039/d1sc01359a.
Galmozzi, A., Dominguez, E., Cravatt, B. F., & Saez, E. (2014). *Application of activity-based protein profiling to study enzyme function in adipocytes*. *Methods in Enzymology*. *Vol. 538* (1st ed.). Elsevier Inc. https://doi.org/10.1016/B978-0-12-800280-3.00009-8.
Gardner, S. N., Slezak, T., & Hall, B. G. (2015). kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics*, *31*(17), 2877–2878.
Hatzios, S. K., Abel, S., Martell, J., Hubbard, T., Sasabe, J., Munera, D., et al. (2016). Chemoproteomic profiling of host and pathogen enzymes active in cholera. *Nature Chemical Biology*, *12*(4), 268–274. https://doi.org/10.1038/nchembio.2025.
Jessani, N., Niessen, S., Wei, B. Q., Nicolau, M., Humphrey, M., Ji, Y., et al. (2005). A streamlined platform for high-content functional proteomics of primary human specimens. *Nature Methods*, *2*(9), 691–697. https://doi.org/10.1038/nmeth778.
Jochum, L., & Stecher, B. (2020). Label or concept—What is a pathobiont? *Trends in Microbiology*, *28*(10), 789–792. https://doi.org/10.1016/j.tim.2020.04.011.
Kamada, N., Chen, G. Y., Inohara, N., & Núñez, G. (2013). Control of pathogens and pathobionts by the gut microbiota. *Nature Immunology*. https://doi.org/10.1038/ni.2608.
Keller, L. J., Babin, B. M., Lakemeyer, M., & Bogyo, M. (2020). Activity-based protein profiling in bacteria: Applications for identification of therapeutic targets and characterization of microbial communities. *Current Opinion in Chemical Biology*. https://doi.org/10.1016/j.cbpa.2019.10.007.

Keller, L. J., Lentz, C. S., Chen, Y. E., Metivier, R. J., Weerapana, E., Fischbach, M. A., et al. (2020). Characterization of serine hydrolases across clinical isolates of commensal skin bacteria Staphylococcus epidermidis using activity-based protein profiling. *ACS Infectious Diseases*, *6*(5), 930–938. https://doi.org/10.1021/acsinfecdis.0c00095.

Kidd, D., Liu, Y., & Cravatt, B. F. (2001). Profiling serine hydrolase activities in complex proteomes. *Biochemistry*, *40*(13), 4005–4015. https://doi.org/10.1021/BI002579J.

Kuljanin, M., Mitchell, D. C., Schweppe, D. K., Gikandi, A. S., Nusinow, D. P., Bulloch, N. J., et al. (2021). Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nature Biotechnology*, *39*(5), 630–641. https://doi.org/10.1038/s41587-020-00778-3.

Lai, Y., Di Nardo, A., Nakatsuji, T., Leichtle, A., Yang, Y., Cogen, A. L., et al. (2009). Commensal bacteria regulate toll-like receptor 3-dependent inflammation after skin injury. *Nature Medicine*, *15*(12), 1377–1382. https://doi.org/10.1038/nm.2062.

Lehmann, J., Cheng, T.-Y., Aggarwal, A., Park, A. S., Zeiler, E., Raju, R. M., et al. (2018). An antibacterial β-lactone kills Mycobacterium tuberculosis by disrupting mycolic acid biosynthesis. *Angewandte Chemie International Edition*, *57*(1), 348–353. https://doi.org/10.1002/anie.201709365.

Lentz, C. S., Sheldon, J. R., Crawford, L. A., Cooper, R., Garland, M., Amieva, M. R., et al. (2018). Identification of a S. aureus virulence factor by activity-based protein profiling (ABPP). *Nature Chemical Biology*, *14*(6), 609–617. https://doi.org/10.1038/s41589-018-0060-1.

Li, M., Patel, H. V., Cognetta, A. B., Smith, T. C., Mallick, I., Cavalier, J.-F., et al. (2021). Identification of cell wall synthesis inhibitors active against Mycobacterium tuberculosis by competitive activity-based protein profiling. *Cell Chemical Biology*. https://doi.org/10.1016/j.chembiol.2021.09.002.

Mazmanian, S. K., Round, J. L., & Kasper, D. L. (2008). A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature*, *453*(7195), 620–625. https://doi.org/10.1038/nature07008.

Méric, G., Mageiros, L., Pensar, J., Laabei, M., Yahara, K., Pascoe, B., et al. (2018). Disease-associated genotypes of the commensal skin bacterium Staphylococcus epidermidis. *Nature Communications*, *9*(1), 5034. https://doi.org/10.1038/s41467-018-07368-7.

Miragaia, M., Thomas, J. C., Couto, I., Enright, M. C., & de Lencastre, H. (2007). Inferring a population structure for Staphylococcus epidermidis from multilocus sequence typing data. *Journal of Bacteriology*, *189*(6), 2540–2552. https://doi.org/10.1128/JB.01484-06.

Naik, S., Bouladoux, N., Wilhelm, C., Molloy, M. J., Salcedo, R., Kastenmuller, W., et al. (2012). Compartmentalized control of skin immunity by resident commensals. *Science*, *337*(6098), 1115–1119. https://doi.org/10.1126/science.1225152.

Ortega, C., Anderson, L. N., Frando, A., Sadler, N. C., Brown, R. W., Smith, R. D., et al. (2016). Systematic survey of serine hydrolase activity in Mycobacterium tuberculosis defines changes associated with persistence. *Cell Chemical Biology*, *23*(2), 290–298. https://doi.org/10.1016/J.CHEMBIOL.2016.01.003.

Otto, M. (2009). Staphylococcus epidermidis—The "accidental" pathogen. *Nature Reviews. Microbiology*, *7*(8), 555–567. https://doi.org/10.1038/nrmicro2182.

Prothiwa, M., & Böttcher, T. (2020). Competitive profiling for enzyme inhibitors using chemical probes. *Methods in Enzymology*, *633*, 49–69. https://doi.org/10.1016/bs.mie.2019.10.031.

Ramirez, J., Guarner, F., Bustos Fernandez, L., Maruy, A., Sdepanian, V. L., & Cohen, H. (2020). Antibiotics as major disruptors of gut microbiota. *Frontiers in Cellular and Infection Microbiology*. https://doi.org/10.3389/fcimb.2020.572912.

Rostovtsev, V. V., Green, L. G., Fokin, V. V., & Sharpless, K. B. (2002). A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie International Edition*, *41*(14), 2596–2599. https://doi.org/10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO;2-4.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*. https://doi.org/10.1038/nmeth.2019.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*. https://doi.org/10.1038/nmeth.2089.

Sharifzadeh, S., Brown, N. W., Shirley, J. D., Bruce, K. E., Winkler, M. E., & Carlson, E. E. (2020). Chemical tools for selective activity profiling of bacterial penicillin-binding proteins. *Methods in Enzymology*, *638*, 27–55. Academic Press Inc. https://doi.org/10.1016/bs.mie.2020.02.015.

Simon, G. M., & Cravatt, B. F. (2010). Activity-based proteomics of enzyme superfamilies: Serine hydrolases as a case study. *The Journal of Biological Chemistry*, *285*(15), 11051–11055. https://doi.org/10.1074/jbc.R109.097600.

Steiger, A. K., Fansler, S. J., Whidbey, C., Miller, C. J., & Wright, A. T. (2020). Probe-enabled approaches for function-dependent cell sorting and characterization of microbiome subpopulations. *Methods in Enzymology*, *638*, 89–107. Academic Press Inc. https://doi.org/10.1016/bs.mie.2020.03.014.

Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L., & Fowler, V. G. (2015). Staphylococcus aureus infections: Epidemiology, pathophysiology, clinical manifestations, and management. *Clinical Microbiology Reviews*, *28*(3), 603–661. https://doi.org/10.1128/CMR.00134-14.

Wallace, M. J., Fishbein, S. R. S., & Dantas, G. (2020). Antimicrobial resistance in enteric bacteria: Current state and next-generation solutions. *Gut Microbes*. https://doi.org/10.1080/19490976.2020.1799654.

Ward, N., & Moreno-Hagelsieb, G. (2014). Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: How much do we miss? *PLoS One*, *9*(7). https://doi.org/10.1371/journal.pone.0101850.