Cell-Selective Chemoproteomics for Biological Discovery

Thesis by
Shannon Elizabeth Stone

In Partial Fulfillment of the Requirements for
the degree of
Doctor of Philosophy in Chemistry

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2018
(Defended July 20th, 2017)

Shannon Elizabeth Stone
ORCID: 0000-0002-6617-3874

ACKNOWLEDGEMENTS

The greatest part of my graduate school experience has been the privilege to learn from many brilliant people and be a part of the exceedingly collaborative nature of this scientific community. First, I thank my advisor and mentor in every sense of the word, Prof. Dave Tirrell, for his continued guidance and support. In addition to his immense scientific insight and encouragement throughout the years, I am continuously impressed by his degree of integrity and thoughtfulness, both of which I strive to emulate. Thanks also to the other members of my thesis committee: Prof. Sarkis Mazmanian for many illuminating discussions, both scientific and personal, Prof. Bob Grubbs for advice and guidance, and Prof. Long Cai for astute recommendations on my project.

I have been fortunate to share my time here with many members of the Tirrell lab, whose patience, brilliance, and humor made the time fly by. I will always cherish the days of the Best Office and summer softball with the Tools of Synthetic Biology.

I greatly enjoyed working with Judy Shon, a talented Caltech undergraduate, over the past years. The Proteome Exploration Laboratory is run by the most wonderful and knowledgeable team, who all contributed to most chapters of this thesis. Thank you also to members of both the Newman and Mazmanian labs for sharing knowledge and resources. The staff at the Beckman Imaging Center has also been exceedingly helpful. Thanks to Prof. Linda Hsieh-Wilson for her support, as well as former members of the Hsieh-Wilson lab for their advice and expertise, and especially Dr. Abby Pulsipher, who continues to be a source of scientific and personal inspiration. I was also fortunate to spend a month at the University of Chicago, learning from some of the world's top researchers in *Staphylococcus aureus*, for which I am very grateful.

Members of the Caltech community including the diversity center, health & counseling centers, and graduate office have all immensely enhanced experience: thank

you. I also thank the Chemistry division for its support, especially for help in starting the Women in Chemistry group. I greatly appreciate the enthusiastic lecturers I had during my time at Caltech, especially Justin Bois, who taught me the joys of data analysis.

To my friends who remind me that there is more to the world than science, and who drink wine with me to celebrate or drink wine with me to commiserate, thank you.

I am forever grateful to my family for their love and encouragement from the other side of the country. I would not be where I am today without each of you.

Finally, to Blake, Olive, and Slevin; thank you, I love you, and let the next great adventure begin.

ABSTRACT

Cellular protein synthesis changes rapidly in response to internal and external cues in ways that vary from cell to cell. Global proteomic analyses of microbial communities, tissues, and organisms have provided important insights into the behavior of such systems, but can obscure the diversity of responses characteristic of different cellular subpopulations. Recent advances in cell-specific proteomics—fueled in part by the development of bioorthogonal chemistries, more sensitive mass spectrometers and more advanced mining algorithms—have yielded unprecedented glimpses into how proteins are expressed in space and time. Whereas previous cell-specific proteomic analyses were confined to abundant cells in relatively simple systems, recent advances in chemoproteomics allow researchers to map the protein expression patterns of even rare cells in complex tissues and whole organisms.

Chapter 1 highlights recently developed strategies for cell-selective proteomics, including metabolic labeling strategies such as bioorthogonal noncanonical amino acid tagging (BONCAT). BONCAT is a chemoproteomic technique that enables temporal labeling of proteins using noncanonical amino acids. In the cell-selective version of BONCAT, expressing a mutant aminoacyl-tRNA synthetase under the control of cell-specific genetic elements affords cellular resolution; only cells of interest can selectively incorporate a noncanonical amino acid into proteins for subsequent detection and identification. Chapter 2 details protocols to set up a cell-selective BONCAT system.

While BONCAT had previously been applied to studies of microbial pathogenesis in tissue culture-based models of infection, we sought to further develop the method to identify the proteome of methicillin-resistant *Staphylococcus aureus* (MRSA) within a mouse model of infection, as detailed in Chapter 3. We used this technique to enrich for staphylococcal proteins made within the host and in addition to finding many factors known to be

important for infection, we also found many that had not previously been associated with infection. Screening several of these previously unknown factors *in vivo* led to the discovery of a novel protein important for MRSA infection. This unbiased approach to cell-selectively label pathogenic proteins during infection could be used as a global discovery tool for novel anti-infective strategies.

In Chapter 4, we combine this cell-selective BONCAT strategy with microbial identification after passive clarity technique (MiPACT) to visualize both staphylococcal protein synthesis and ribosomal RNA within whole skin abscesses during infection. In Chapter 5, we continue developing cell-selective BONCAT to study microbial protein synthesis in the context of a living mouse by extending the system to *Bacteroides fragilis*, a common human gut commensal.

Finally, cell-selective BONCAT is wholly dependent on the bioorthogonal nature of the azide and its detection reagents. Fishing out an azide-tagged molecule from the rest of the cellular milieu requires optimization of enrichment-based strategies. In Chapter 6, we describe the development of a peptide to quantitate the gain of our enrichments.

While innovations in mass spectrometry and computational algorithms have facilitated the identification and quantification of thousands of proteins simultaneously from complex samples, this abundance of data does not necessarily lead to biological insight. Cell-specific proteomic techniques will play a key role in the identification of the mechanisms that govern cell specialization and that allow organisms to respond to changing environments. Overall, this work demonstrates the power of cell-selective chemoproteomics to ascertain biological insights in complex systems.

PUBLISHED CONTENT AND CONTRIBUTIONS

(1) Pulsipher, A., Griffin, M. E., **Stone, S. E.**, Brown, J. M., Hsieh-Wilson, L. C. "Directing neuronal signaling through cell-surface glycan engineering." J. Am. Chem. Soc. 2014, 136 (19), 6794-6797. DOI: 10.1021/ja5005174

S.E.S. participated in project conception, optimized the TrkA assay, and contributed to the writing of the manuscript. This article's figures are reproduced in part within Appendix B with permission according to the ACS AuthorChoice terms of use.

(2) Pulsipher, A., Griffin, M. E., **Stone, S. E.**, Hsieh-Wilson, L. C. "Long-lived engineering of glycans to direct stem cell fate." Angew. Chem. Int. Ed. **2015**, 54, 1466-1470. DOI: 10.1002/anie.201409258

S.E.S. participated in project conception, created the plasmids, determined conditions for cell labeling, and contributed to the writing of the manuscript. This article's figures are reproduced in part within Appendix C with permission from the publisher.

(3) Griffin, M. E., Jensen, E. H., Mason, D. E., Jenkins, C. L., **Stone, S. E.**, Peters, E. C., Hsieh-Wilson, L. C. "Comprehensive mapping of *O*-GlcNAc modification sites using a chemically cleavable tag." Mol. Biosys. **2016**, 12, 1756-1759. DOI: 10.1039/c6mb00138f

*S.E.S. synthesized and characterized compound **1** in Figure 1A. This article's figures are reproduced in part within Appendix D with permission according to the Creative Commons Attribution 3.0 Unported License terms of use.*

(4) **Stone, S. E.**, Glenn, W. S., Hamblin, G. D., Tirrell, D. A. "Cell-selective proteomics for biological discovery." Curr. Opin. Chem. Biol. **2017**, 36, 50-57. DOI: 10.1016/j.cbpa.2016.12.026

S.E.S. chose references, devised the outline, designed figures, created the table, and wrote the manuscript. This article including figures is reproduced in Chapter I with permission from the publisher.

(5) Glenn, W. S., **Stone, S. E.**, Ho, S. H., Sweredoski, M. J., Moradian, A., Hess, S., Bailey-Serres, J., Tirrell, D. A. "Bioorthogonal noncanonical amino acid tagging (BONCAT) enables time-resolved analysis of protein synthesis in native plant tissue." Plant Phys. **2017**, 173 (3), 1543-1553. DOI: 10.1104/pp.16.01762

S.E.S. performed data analysis, designed figures, and contributed to the writing of the manuscript. This article's figures are reproduced in part within Appendix E with permission from the publisher.

TABLE OF CONTENTS

LIST OF FIGURES

NOMENCLATURE

aaRS. Aminoacyl-tRNA synthetase
Aha. Azidohomoalanine
Anl. Azidonorleucine
APEX. Ascorbate peroxidase
Azf. 4-azido-L-phenylalanine
BONCAT. Bioorthogonal noncanonical amino acid tagging
CFU. Colony-forming units
Cm. Chloramphenicol
CTAP. Cell-type specific labeling with amino acid precursors
CuAAC. Copper catalyzed alkyne-azide cycloaddition
DAPI. 4,6-Diamidino-2-phenylindole
DBCO. Aza-dibenzocyclooctyne
Dde. N-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl)
EF. 4-ethynyl-L-phenylalanine
Erm. Erythromycin
FACS. Fluorescence activated cell sorting
FISH. Fluorescence in situ hybridization
Gent. Gentamicin
GF. Germ-free
GO. Gene ontology
H&E. Hematoxylin and eosin
HCR. Hybridization chain reaction
Hpg. Homopropargylglycine
IP. Intraperitoneal
LB. Luria-Bertani broth
LC. Liquid chromatography
LFQ. Label-free quantitation
LIMMA. Linear models for microarray data
MetRS. Methionyl-tRNA synthetase
MiPACT. Microbial identification after PACT
MRM. Multiple reaction monitoring

MRSA. Methicillin-resistant *Staphylococcus aureus*
MS. Mass spectrometry
ncAA. Noncanonical amino acid
OCT. Optimum cutting temperature
PACT. Passive CLARITY technique
PBS. Phosphate-buffered saline
PCA. Principal coordinate analysis
PCR. Polymerase chain reaction
PEG. Polyethylene glycol
PFA. Paraformaldehyde
Phe. Phenylalanine
PheRS. Phenylalanyl-tRNA synthetase
PMN. Polymorphonuclear leukocyte
PO. Per os (Oral Gavage)
Pra. Propargylglycine
Rcf. Relative centrifugal force
RIMS. Refractive index matching solution
S/C. Sub-cutaneous
SILAC. Stable isotope labeling using amino acids in culture
SORT. Stochastic orthogonal recoding of translation
SPAAC. Strain-promoted azide-alkyne cycloaddition
Spec. Spectinomycin
SPF. Specific-pathogen free
TAMRA. Tetramethylrhodamine
Tet. Tetracycline
TRAP. Translating ribosome affinity purification
TSB. Tryptic soy broth

Chapter 1

CELL-SELECTIVE PROTEOMICS

Published as:

Stone, S. E., Glenn, W. S., Hamblin, G. D., Tirrell, D. A. "Cell-selective proteomics for biological discovery." Curr. Opin. Chem. Biol. **2017**, 36, 50-57. DOI: 10.1016/j.cbpa.2016.12.026

1.1 Abstract

Proteomic plasticity is a hallmark of development and adaptation. Organisms rely upon translational regulation to respond rapidly to both internal and external cues. Convenience, and in some cases necessity, drove early systems-level studies of translational control to adopt transcript-based proxies instead of direct protein measurements. However, discordance between steady-state transcript and protein levels argues for the development of methods that more accurately quantify expression. Recent advances in cell-specific translatomics and proteomics—fueled in part by the development of bioorthogonal chemistries, more sensitive mass spectrometers and more advanced mining algorithms—have yielded unprecedented glimpses into how proteins are expressed in space and time. Whereas previous cell-specific proteomic analyses were confined to abundant cells in relatively simple matrices, recent advances allow researchers to map the protein expression patterns of even rare cells in complex tissues and whole organisms.

1.2 Introduction

Cellular protein synthesis changes rapidly in response to internal and external cues in ways that vary from cell to cell. Global proteomic analyses of microbial communities, tissues, and organisms have provided important insights into the behavior of such systems, but can obscure the diversity of responses characteristic of different cellular subpopulations (Figure 1). Cell-selective methods for the analysis of protein synthesis are being developed to resolve proteomic changes in space and time.

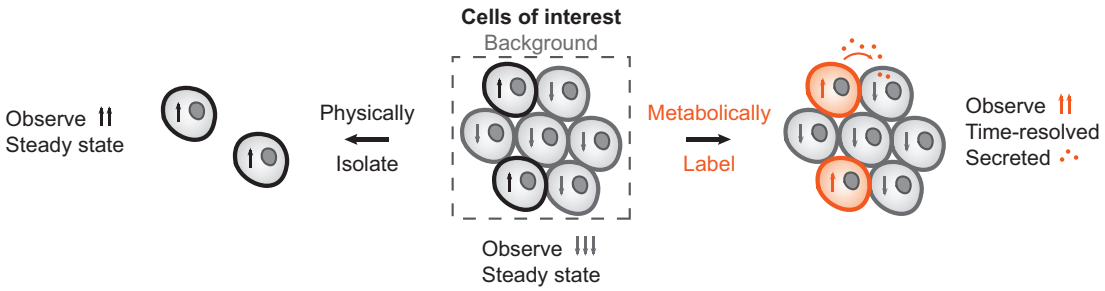


Figure 1.1: The importance of cell-type-specific proteomics. Bulk measurements of complex tissues can obscure proteomic changes that occur in specific sub-populations of cells. A protein that is highly expressed (up arrows) in the cells of interest might be detected at low abundance overall due to low expression (down arrows) in background cells. Cells of interest must be physically isolated or tagged to measure the cell-specific proteome. Physical isolation measures steady-state levels of intracellular proteins, whereas labeling methods can be time-resolved and used to identify secreted proteins.

Cell-type-specific transcriptomics experiments have revealed mRNA expression patterns in a wide array of biological systems, but mRNA and protein levels are often dissonant (1). Moreover, some important elements of proteome dynamics, including post-

translational modification, degradation, and localization, cannot be addressed by mRNA measurements alone (2, 3). Until recently, changes in protein abundance in specific cells could be measured only in targeted, low-throughput experiments, but innovations in mass spectrometry and computational algorithms have facilitated the identification and quantification of thousands of proteins simultaneously from complex biological samples (4-6).

In this chapter, we highlight recent developments in determining cell-type-specific proteomes and recommend experimental design strategies that are guided by the question at hand.

1.3 Cell-selective translatomics and ribosome profiling

Translatomic studies, which select for ribosome-associated transcripts, have yielded stronger correlations between transcript and protein abundances than experiments that measure steady-state mRNA levels (7). Cell-type-specific studies have been enabled by translating ribosome affinity purification (TRAP), a method in which epitope-tagged ribosomes and their associated transcripts are captured, enriched and subjected to amplification and deep sequencing (8). TRAP can be rendered cell-specific by placing expression of the tagged ribosome under control of a selective promoter.

More recently, Ingolia and Weissman have developed ribosome profiling, which identifies ribosome-protected mRNA footprints and allows investigators to determine ribosome occupancy with positional specificity. This information can be used to measure translation

levels and locate non-canonical start sites (7). Gonzalez *et al*. used TRAP to cell-selectively purify ribosome-bound transcripts, and employed ribosome profiling to identify the translatome of gliomas and to reveal decreased translation in glial progenitors compared to the tumor microenvironment (9). Ribosome profiling is a powerful technique that we expect to find increasing use upon further development of cell-specific methods.

While translatomic studies provide greater depth of coverage than current proteomic measurements, ribosome binding does not ensure that a transcript is undergoing active translation (10).

1.4 Separating cells for steady-state proteomic analysis

The earliest strategies to determine cell-specific proteomes relied on separating and purifying the cells of interest prior to analysis. Cells can be sorted on the basis of expression of a transgene under control of a cell-specific promoter or by antibody staining of marker epitopes. These tools are well established and have been thoughtfully reviewed (10, 11). Physical methods have been used for years to isolate cell types from mammalian tissues for subsequent downstream analyses (12, 13). More recently these methods have been used to measure growth rates and elucidate proteomic signatures of *Salmonella* during murine infection (14).

Physical separations remain the best method for analyzing clinical specimens and genetically intractable organisms. However, imperfect separations and long sample processing times can diminish selectivity and increase the likelihood of artifacts.

Furthermore, such methods intrinsically yield steady-state proteomic information. In contrast, metabolic labeling strategies enable cell-specific proteomic analysis to be accomplished in time-resolved fashion.

1.5 Metabolic labeling: trade-offs between sensitivity and perturbation

Metabolic labeling methods are temporally resolved and use an arsenal of amino acid isotopologs, non-canonical amino acids, and analogs of protein synthesis inhibitors (Figure 2). Each of these strategies can be placed under control of cell-specific genetic elements to afford cellular resolution. The choice of promoter(s) is key for these systems, and the degree of protein labeling needs to be weighed against the possibility of perturbing the system. Results should be validated via independent assays because labels may affect protein expression, stability, and/or function.

Cell-type-specific labeling using amino acid precursors (CTAP)

Stable isotope labeling by amino acids in cell culture (SILAC) relies on the incorporation of isotopically labeled amino acids into proteins. To make SILAC cell-selective, Gauthier *et al*. introduced cell-type-specific labeling using amino acid precursors (CTAP), a method that exploits the fact that lysine is an essential amino acid in mammalian cells (15). Cell-selective expression of biosynthetic enzymes allows L-lysine isotopologs to be synthesized *in situ* starting from isotope-labeled precursors. Only minor differences in gene expression resulted from feeding the heavy precursor to cells expressing the biosynthetic machinery versus supplementing cells directly with L-lysine.

In principle, both exchange of L-lysine between cells and extracellular processing of the precursor can compromise the cell-specificity of the CTAP method. When Lavis and coworkers employed an analogous strategy to unmask fluorophores in targeted cells, they noted that the unmasked small molecule diffused through gap junctions. This effect can be exploited to study cell-cell connectivity, but would confound cell-specific protein labeling if the small molecule were to diffuse to cells lacking the decaging enzyme (16). To address these concerns, Tape *et al*. optimized CTAP for eukaryotic cell types and achieved ~90% cell-specific labeling in ten-day co-cultures (17). Using their optimized method, Tape *et al*. combined CTAP with phosphoproteomics to study heterocellular $KRAS^{G12D}$ signaling in pancreatic ductal adenocarcinoma cells (18). By restricting their proteomic analysis to cells that expressed $KRAS^{G12D}$, the authors showed that the oncogene regulates AKT through reciprocal signaling – not through the accepted cell-autonomous pathway.

Bioorthogonal Noncanonical Amino acid Tagging (BONCAT)

CTAP is most suitable for cell-specific experiments conducted in culture on timescales of 3-7 days (19). For studies that require better time resolution, the bioorthogonal non-canonical amino acid tagging (BONCAT) method, introduced by Dieterich and coworkers, offers a good alternative (20, 21). In its original form, BONCAT exploits the capacity of the endogenous aminoacyl-tRNA synthetases to charge non-canonical amino acids (ncAAs) to their cognate tRNAs for incorporation into proteins. ncAAs bearing bioorthogonal chemical handles, often azides or alkynes, enable conjugation to affinity tags and separation of tagged proteins from the rest of the protein pool. The methionine surrogates azidohomoalanine (Aha) and homopropargylglycine (Hpg) have been used to

probe proteome dynamics in bacterial (22-26) and mammalian (27) systems, and notably, to enrich and quantify secreted proteins (28). Depletion of cellular methionine is not necessary for Aha labeling; Bagert *et al*. showed that a 30:1 ratio of Aha to Met yielded excellent protein labeling while minimizing perturbations that might be expected to arise from methionine starvation (29). Other studies have shown that ncAA labeling for periods of up to two days do not perturb embryonic growth in live mice (30). In designing a BONCAT experiment, the investigator should choose concentrations of the ncAA label and its natural counterpart that reflect the relative rates of activation of the amino acids by the cognate synthetase.

In 2009, Ngo and coworkers developed a cell-selective version of BONCAT by engineering an *E. coli* methionyl-tRNA synthetase (*Ec*MetRS) variant that activates azidonorleucine (Anl). Because Anl is a poor substrate for wild-type *Ec*MetRS, labeling is essentially restricted to cells that express the mutant synthetase. In the first example of the cell-specific BONCAT method, Ngo *et al*. reported specific labeling of *E. coli* cells co-cultured with murine alveolar macrophages (31). Grammel *et al*. expanded on this method by enriching for proteins synthesized during *Salmonella typhimurium* infection (32), and Mahdavi and coworkers used BONCAT to determine the order in which *Yersinia enterocolitica* effector proteins are injected into HeLa cells in the course of infection (33).

Cell-selective BONCAT has now been extended to proteomic analysis in live animals, highlighting its potential utility in creating cell-specific proteomic “atlases”. In 2015 we reported a mutant phenylalanyl-tRNA synthetase (PheRS) that enables the use of *p*-

azidophenylalanine (Azf) as a BONCAT probe in *Caenorhabditis elegans* (34). Combining cell-selective BONCAT with stable isotope labeling, we used the *myo-2* promoter to direct expression of the mutant synthetase to the 20 pharyngeal muscle cells of the worm. We were able to quantify 2270 proteins by this method, and to verify the pharyngeal expression patterns of several previously uncharacterized proteins.

Dieterich and coworkers have adapted cell-selective BONCAT labeling to *Drosophila melanogaster* through controlled expression of the *Dm*MetRS L262G mutant (35). Chronic administration of Anl in developing flies expressing the mutant synthetase caused slight impairments in larval growth and behavior, but shorter (48 h) labeling times led to no noticeable defects. Importantly, administration of the amino acid in flies that did not express the mutant MetRS caused no discernible effect. Using this strategy, Niehues *et al*. measured reduced neuronal protein synthesis rates in a *Drosophila* model of Charcot-Marie-Tooth (CMT) neuropathy (36). Mahdavi *et al.* and Muller *et al.* have employed the analogous (L274G) mouse synthetase in mammalian cell culture and in a neuron-glia co-culture system, respectively (37, 38). The latter experiments enabled the investigators to monitor changes in the astrocytic proteome in response to treatment with brain-derived neurotrophic factor (BDNF).

Split synthetases have been developed to enable cell-selective analysis of systems in which no single promoter restricts expression of the mutant enzyme to the cells of interest (39). Notably, all amino acids and enrichment media needed for BONCAT experiments are commercially available.

Stochastic Orthogonal Recoding of Translation (SORT)

Chin and coworkers have developed a residue-specific ncAA-labeling technology termed stochastic orthogonal recoding of translation (SORT), which – like BONCAT – allows chemoselective modification and enrichment of newly synthesized cellular proteins. SORT relies on expression of a pyrrolysyl-tRNA synthetase and its cognate tRNA (40, 41). Using this method, Elliott *et al*. cell-selectively labeled and identified proteins made during different stages of larval growth in *Drosophila*. Importantly, SORT allows the anticodon of the cognate tRNA to be changed to direct the ncAA to different sets of codons in the labeled proteins. Elliott *et al*. have characterized the enrichment process and found that tagging at different codons leads to the enrichment of overlapping, but distinct sets of proteins (42). The authors noted that simultaneous expression of multiple tRNAs (i.e., tRNA-Ala, -Ser and -Met) increases labeling efficiency. Furthermore, Elliott *et al.* found that enrichment after tagging improves detection of low-abundance proteins.

Cell-selective O-propargyl-puromycin (OP-Puro) labeling

The *O*-propargyl-puromycin (OP-Puro) method also incorporates "clickable" handles into nascent proteins (43). Cohen and coworkers recently achieved cell-targeted OP-puromycin labeling by using a phenylacetyl-caged analog that is uncaged by cell-selective expression of penicillin G acylase (PGA) (44). The OP-puro method is the fastest of the metabolic labeling methods and the best suited for studies requiring ultra-short labeling times (45). Prolonged labeling with OP-puro would be expected to perturb cellular behavior through

inhibition of global translation. Furthermore, premature truncation renders this method ineffective for the identification of secreted proteins.

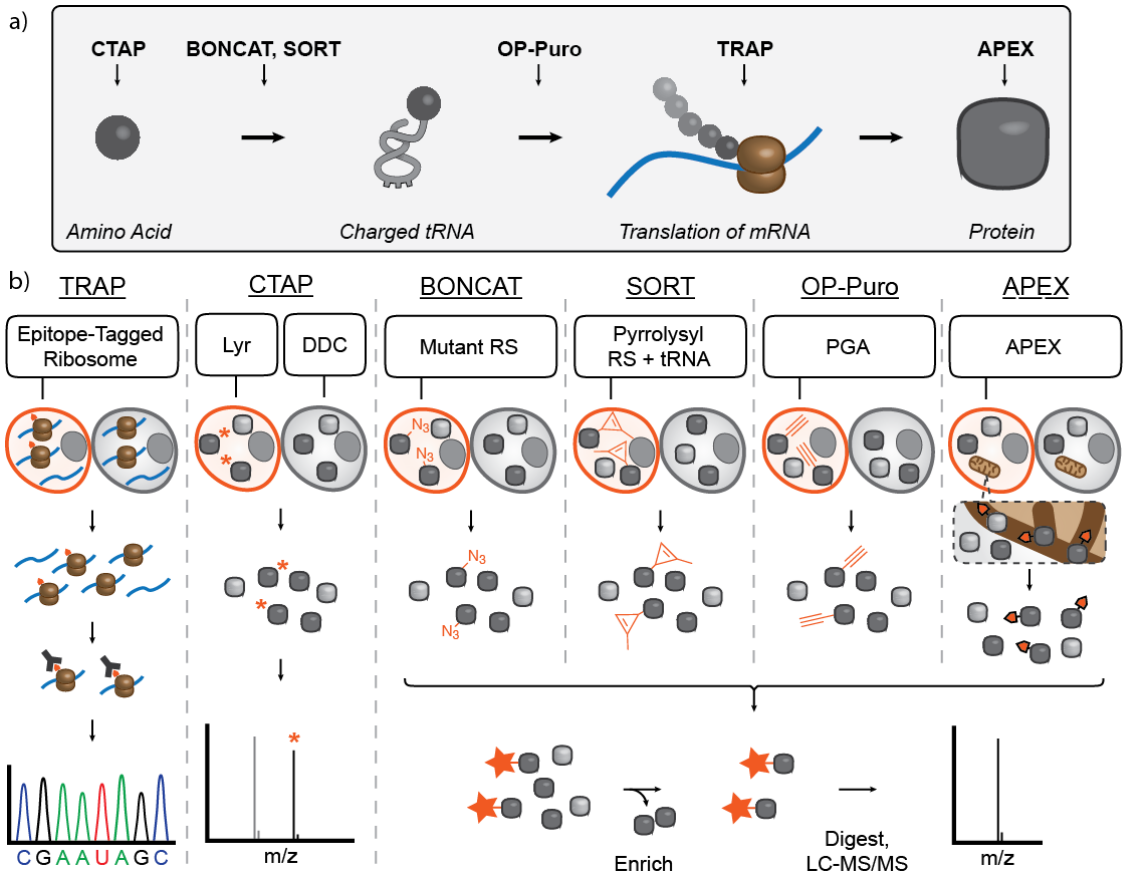


Figure 1.2: Labeling strategies for cell-selective proteomics. a) The process by which amino acids are incorporated into proteins, and the step exploited by each of the labeling methods discussed in this chapter. b) Schematic of each technique. Translating ribosome affinity purification: TRAP; Cell type-specific labeling using amino acid precursors: CTAP; Bioorthogonal noncanonical amino acid tagging: BONCAT; Stochastic orthogonal recoding of translation: SORT; *O*-propargyl puromycin: OP-Puro; ascorbate peroxidase: APEX; Lysine racemase: Lyr; diaminopimelate decarboxylase: DDC; aminoacyl-tRNA synthetase: RS; penicillin-G-acylase: PGA.

1.6 Spatially restricted & subcellular proteomics

Ting and coworkers first used a mutant ascorbate peroxidase (APEX) to selectively tag proteins localized to the mitochondrial matrix (46, 47). Unlike the cell-selective metabolic labeling methods just described, this method labels all proteins, including pre-existing proteins, within a subcellular volume. Chen *et al*. used this elegant strategy to characterize multiple cell types in *Drosophila*, including the mitochondrial matrix of muscle tissue (48). The Weissman laboratory has combined the APEX labeling method with ribosome profiling to characterize localized protein synthesis in yeast (49, 50); extension of their method to cell-selective analysis is readily imagined.

1.7 Choosing a cell-selective proteomic method

The choice of a cell-selective method of proteomic analysis should reflect careful consideration of the advantages and disadvantages of each of the available approaches (Table 1).

Physical sorting methods allow straightforward characterization of the steady-state proteome of the cell type of interest. However, removing cells from their natural environments prior to analysis raises concerns about artifacts, leads to limited temporal information, and sacrifices information about secreted proteins.

Ribosome profiling, when combined with cell-selective TRAP, provides significantly higher coverage of the gene expression profile than any direct proteomic measurement. But ribosome profiling is not a perfect proxy for protein synthesis and yields no

information regarding protein secretion (51). Moreover, only direct proteomic methods allow detection of post-translational modifications.

CTAP simplifies quantitative proteomic measurements for samples of relatively low complexity, but enrichment-based strategies (i.e., BONCAT, SORT or OP-Puro) are likely to be superior for short labeling times or for analysis of rare cells in complex tissues. Only APEX yields snapshots of the steady-state proteome with sub-cellular resolution. All cell-selective, enrichment-based experiments require the use of genetically tractable organisms.

Optimization of enrichment-based strategies requires careful consideration of alternative purification chemistries. Attachment to the resin used for purification can be accomplished either by direct covalent ligation or by a two-step process of affinity-tagging (e.g., with biotin reagents) and non-covalent binding (e.g., to streptavidin resins). Following appropriate washing steps, samples can be released from the resin by competitive binding, by proteolysis, or by selective cleavage of the affinity reagent. APEX appends biotin to surrounding molecules, so streptavidin-based resins are used to enrich for labeled proteins (47). OP-Puro requires an azide-based affinity handle or resin for enrichment (44). SORT uses cyclopropene labels and tetrazine linkers in a ligation reaction reported to be 100 to 1000 times faster than the strain-promoted azide-alkyne cycloaddition (42). BONCAT labels with either alkynes or azides, and enriches with complementary azide or alkyne reagents. A special consideration arises in the analysis of lysates labeled with azides: Free thiols, which are known to react with cyclooctynes, must be blocked with capping

reagents such as iodoacetamide or N-ethylmaleimide to avoid high background (34). Many azide and alkyne resins and linkers are commercially available, and tetrazine-based reagents are beginning to appear on the market.

If the investigator wishes to identify the sites at which protein labeling has occurred, linkers with cleavable moieties can be used (52). For many experiments, though, identification of labeling sites is not necessary, and on-bead digestion of enriched proteins is often simpler and more straightforward. In our hands, directly conjugating azide-labeled lysates to cyclooctyne resins has allowed us to identify larger numbers of relevant proteins (34). Because enrichments are never perfect, running mock enrichments of unlabeled sample along with labeled samples provides a useful indication of background reactivity and non-specific protein contamination. Samples with abundant contaminating biopolymers such as pectin, serum proteins, or mucin may need an additional step to remove or degrade these contaminants and facilitate successful enrichment.

1.8 Conclusions & future outlook

Recent years have witnessed the introduction of powerful techniques that allow investigators to monitor protein synthesis with unprecedented resolution in space and time. Cell-specific proteomic analyses will play a key role in the identification of the mechanisms that govern cell specialization and that allow complex organisms to respond to changing environments.

	Cell-Specific Method	Biomolecule identified	Organisms demonstrated in?	Temporal Resolution	Secreted?	PTM?	Advantages	Disadvantages	References
Translatomics	TRAP	mRNA	Prokaryotes, eukaryotes	Snapshot of translation	No	No	High sequence coverage, able to combine with ribosome profiling	Requires expression of tagged ribosome, miss translational control	*[7-11]*
Cell Separation	Manual	mRNA, Protein	Prokaryotes, eukaryotes, clinical samples	Steady-state proteome	No	Yes	Straightforward, inexpensive	Possible artifacts from sample preparation, time and labor intensive	*[3], [12-13]*
	FACS	mRNA, Protein	Prokaryotes, eukaryotes, clinical samples	Steady-state proteome	No	Yes	High-throughput	Requires dissociation of cells, need expression of transgene or recognizable epitope, need specialized equipment	*[14]*
Metabolic	CTAP	Protein	Cell culture	Up to 10 days continuous cell culture	Yes	Yes	Quantitative, compatible with long-term cell culture	Requires expression of Lyr/DDC, cells must be auxotrophic for lysine, restricted to cell culture	*[15], [17-19]*
	BONCAT	Protein	Prokaryotes, eukaryotes, cell culture	Newly synthesized proteins in minutes (prokaryotic, cell culture) to days (whole animal)	Yes	Yes	Commercially available reagents, high degree of temporal resolution	Requires expression of synthetase; only Met/Phe residue replacement available currently, requires delivery of the ncAA	*[20-38], [44]*
	SORT	Protein	Eukaryotes, cell culture	Newly synthesized proteins in minutes (cell culture) to days (whole animal)	Yes	Yes	Easy to change the residue of non-canonical amino acid incorporation, high degree of temporal resolution	Requires expression of synthetase/tRNA pair, reagents not currently commercially available, requires delivery of the ncAA	*[39-41]*
	OP-Puro	Protein	Cell culture	Newly synthesized proteins in minutes (cell culture)	No	No	Not residue-dependent, highest degree of temporal resolution	Requires expression of PGA, reagents not currently commercially available	*[42-44]*
Spatial	APEX	mRNA, Protein	Eukaryotes, cell culture	Subcellular, steady-state proteome	No	Yes	High degree of spatial resolution (subcellular)	Requires expression of APEX/HRP, no temporal resolution, requires delivery of peroxide + biotin-phenol reagent, possible toxicity of peroxide over longer timescales, needs oxidative environment	*[45-49]*

Figure 1.3: Advantages and disadvantages of cell-specific proteomic methods.

References

1. Vogel C, Marcotte EM: Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Gen* 2012, 13:227-232.
2. Maier T, Guell M, Serrano L: Correlation of mRNA and protein in complex biological samples. *FEBS Lett* 2009, 583:3966-3973.
3. Kislinger T, Cox B, Kannan A, Chung C, Hu PZ, Ignatchenko A, Scott MS, Gramolini AO, Morris Q, Hallett MT, et al.: Global survey of organ and organelle protein expression in mouse: Combined proteomic and transcriptomic profiling. *Cell* 2006, 125:173-186.
4. Zhang YY, Fonslow BR, Shan B, Baek MC, Yates JR: Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem Rev* 2013, 113:2343-2394.
5. Aebersold R, Mann M: Mass-spectrometric exploration of proteome structure and function. *Nature* 2016, 537:347-355.
6. Yuet KP, Tirrell DA: Chemical Tools for Temporally and Spatially Resolved Mass Spectrometry-Based Proteomics. *Ann Biomed Eng* 2014, 42:299-311.
7. Ingolia NT: Ribosome Footprint Profiling of Translation throughout the Genome. *Cell* 2016, 165:22-33.
8. Sanz E, Yang L, Su T, Morris DR, McKnight GS, Amieux PS: Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc Natl Acad Sci* 2009, 106:13939-13944.
9. Gonzalez C, Sims JS, Hornstein N, Mela A, Garcia F, Lei L, Gass DA, Amendolara B, Bruce JN, Canoll P, et al.: Ribosome Profiling Reveals a Cell-Type-Specific Translational Landscape in Brain Tumors. *J Neurosci* 2014, 34:10924-10936.

10. Handley A, Schauer T, Ladurner AG, Margulies CE: Designing Cell-Type-Specific Genome-wide Experiments. *Mol Cell* 2015, 58:621-631.
    - This review delineates recent advances in cell-type-specific genomic methods, and thoughtfully recommends experimental approaches and controls.
11. Okaty BW, Sugino K, Nelson SB: Cell Type-Specific Transcriptomics in the Brain. *J Neurosci* 2011, 31:6939-6943.
12. Sharma K, Schmitt S, Bergner CG, Tyanova S, Kannaiyan N, Manrique-Hoyos N, Kongi K, Cantuti L, Hanisch UK, Philips MA, et al.: Cell type- and brain region-resolved mouse brain proteome. *Nat Neuro* 2015, 18:1819-1831.
13. Azimifar SB, Nagaraj N, Cox J, Mann M: Cell-Type-Resolved Quantitative Proteomics of Murine Liver. *Cell Metab* 2014, 20:1076-1087.
14. Claudi B, Sprote P, Chirkova A, Personnic N, Zankl J, Schurmann N, Schmidt A, Bumann D: Phenotypic Variation of Salmonella in Host Tissues Delays Eradication by Antimicrobial Chemotherapy. *Cell* 2014, 158:722-733.
15. Gauthier NP, Soufi B, Walkowicz WE, Pedicord VA, Mavrakis KJ, Macek B, Gin DY, Sander C, Millerephrin ML: Cell-selective labeling using amino acid precursors for proteomic studies of multicellular environments. *Nat Meth* 2013, 10:768-+.
16. Tian L, Yang YL, Wysocki LM, Arnold AC, Hu A, Ravichandran B, Sternson SM, Looger LL, Lavis LD: Selective esterase-ester pair for targeting small molecules with cellular specificity. *Proc Natl Acad Sci* 2012, 109:4756-4761.
17. Tape CJ, Norrie IC, Worboys JD, Lim L, Lauffenburger DA, Jorgensen C: Cell-specific Labeling Enzymes for Analysis of Cell-Cell Communication in Continuous Co-culture. *Mol Cell Prot* 2014, 13:1866-1876.
18. Tape CJ, Ling S, Dimitriadi M, McMahon KM, Worboys JD, Leong HS, Norrie IC, Miller CJ, Poulogiannis G, Lauffenburger DA, et al.: Oncogenic KRAS Regulates Tumor Cell Signaling via Stromal Reciprocation. *Cell* 2016, 165:910-920.
    - The authors combine cell-specific proteomic labeling (CTAP) and multivariate phosphoproteomics to study pancreatic ductal adenocarcinomas in heterocellular contexts. This study reveals reciprocal $KRAS^{G12D}$ signaling, which would have gone undetected if the carcinomas were studied in isolation.
19. Tape CJ: Systems Biology Analysis of Heterocellular Signaling. *Trends in Biotech* 2016, 34:627-637.
20. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM: Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci* 2006, 103:9482-9487.
21. Dieterich DC, Lee JJ, Link AJ, Graumann J, Tirrell DA, Schuman EM: Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nat Prot* 2007, 2:532-540.
22. Bagert JD, van Kessel JC, Sweredoski MJ, Feng L, Hess S, Bassler BL, Tirrell DA: Time-resolved proteomic analysis of quorum sensing in *Vibrio harveyi*. *Chem Sci* 2015, 7:1797-1806.

23. Feng L, Rutherford ST, Papenfort K, Bagert JD, van Kessel JC, Tirrell DA, Wingreen NS, Bassler BL: A qrr noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. *Cell* 2015, 160:228-240.
24. Kramer G, Sprenger RR, Back J, Dekker HL, Nessen MA, van Maarseveen JH, de Koning LJ, Hellingwerf KJ, de Jong L, de Koster CG: Identification and quantitation of newly synthesized proteins in *Escherichia coli* by enrichment of azidohomoalanine-labeled peptides with diagonal chromatography. *Mol Cell Prot* 2009, 8:1599-1611.
25. Sinai L, Rosenberg A, Smith Y, Segev E, Ben-Yehuda S: The molecular timeline of a reviving bacterial spore. *Mol Cell* 2015, 57:695-707.
26. Hatzenpichler R, Connon SA, Goudeau D, Malmstrom RR, Woyke T, Orphan VJ: Visualizing in situ translational activity for identifying and sorting slow-growing archaeal-bacterial consortia. *Proc Natl Acad Sci* 2016, 113:E4069-E4078.
27. Howden AJ, Geoghegan V, Katsch K, Efstathiou G, Bhushan B, Boutureira O, Thomas B, Trudgian DC, Kessler BM, Dieterich DC, et al.: QuaNCAT: quantitating proteome dynamics in primary cells. *Nat Meth* 2013, 10:343-346.
28. Eichelbaum K, Winter M, Diaz MB, Herzig S, Krijgsveld J: Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nat Biotech* 2012, 30:984-+.
29. Bagert JD, Xie YSJ, Sweredoski MJ, Qi YT, Hess S, Schuman EM, Tirrell DA: Quantitative, Time-Resolved Proteomic Analysis by Combining Bioorthogonal Noncanonical Amino Acid Tagging and Pulsed Stable Isotope Labeling by Amino Acids in Cell Culture. *Mol Cell Prot* 2014, 13:1352-1358.
30. Calve S: Incorporation of non-canonical amino acids into the developing murine proteome. *Sci Rep* 2016.
31. Ngo JT, Champion JA, Mahdavi A, Tanrikulu IC, Beatty KE, Connor RE, Yoo TH, Dieterich DC, Schuman EM, Tirrell DA: Cell-selective metabolic labeling of proteins. *Nat Chem Biol* 2009, 5:715-717.
32. Grammel M, Zhang MZM, Hang HC: Orthogonal Alkynyl Amino Acid Reporter for Selective Labeling of Bacterial Proteomes during Infection. *Angew Chem* 2010, 49:5970-5974.
33. Mahdavi A, Szychowski J, Ngo JT, Sweredoski MJ, Graham RLJ, Hess S, Schneewind O, Mazmanian SK, Tirrell DA: Identification of secreted bacterial proteins by noncanonical amino acid tagging. *Proc Natl Acad Sci* 2014, 111:433-438.
34. Yuet KP, Doma MK, Ngo JT, Sweredoski MJ, Graham RLJ, Moradian A, Hess S, Schuman EM, Sternberg PW, Tirrell DA: Cell-specific proteomic analysis in Caenorhabditis elegans. *Proc Natl Acad Sci* 2015, 112:2705-2710.
    - The authors present a mutant PheRS to label and visualize proteins expressed in various cell types in the nematode *Caenorhabitis elegans*. The authors combine BONCAT and SILAC to enrich, identify and quantify proteins produced in the 20 pharyngeal muscle cells of *C. elegans*.

35. Erdmann I, Marter K, Kobler O, Niehues S, Abele J, Muller A, Bussmann J, Storkebaum E, Ziv T, Thomas U, et al.: Cell-selective labelling of proteomes in Drosophila melanogaster. *Nat Comm* 2015, 6.
    - This study utilizes a mutant MetRS to perform cell-type-specific BONCAT analysis in *Drosophila*.
36. Niehues S, Bussmann J, Steffes G, Erdmann I, Kohrer C, Sun LT, Wagner M, Schafer K, Wang GX, Koerdt SN, et al.: Impaired protein translation in Drosophila models for Charcot-Marie-Tooth neuropathy caused by mutant tRNA synthetases. *Nat Comm* 2015, 6.
37. Mahdavi A, Hamblin GD, Jindal GA, Bagert JD, Dong C, Sweredoski MJ, Hess S, Schuman EM, Tirrell DA: Engineered Aminoacyl-tRNA Synthetase for Cell-Selective Analysis of Mammalian Protein Synthesis. *J Am Chem Soc* 2016, 138:4278-4281.
38. Muller A, Stellmacher A, Freitag CE, Landgraf P, Dieterich DC: Monitoring Astrocytic Proteome Dynamics by Cell Type-Specific Protein Labeling. *PLoS One* 2015, 10.
39. Mahdavi A, Segall-Shapiro TH, Kou SZ, Jindal GA, Hoff KG, Liu S, Chitsaz M, Ismagilov RF, Silberg JJ, Tirrell DA: A Genetically Encoded AND Gate for Cell-Targeted Metabolic Labeling of Proteins. *J Am Chem Soc* 2013, 135:2979-2982.
40. Elliott TS, Townsley FM, Bianco A, Ernst RJ, Sachdeva A, Elsasser SJ, Davis L, Lang K, Pisa R, Greiss S, et al.: Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat Biotech* 2014, 32:465-U186.
    - This study introduces the use of SORT to label, image and identify proteins from germ cells in *Drosophila melanogaster* ovaries.
41. Elliott TS, Bianco A, Chin JW: Genetic code expansion and bioorthogonal labelling enables cell specific proteomics in an animal. *Curr Opin Chem Biol* 2014, 21:154-160.
42. Elliott TS, Bianco A, Townsley FM, Fried SD, Chin JW: Tagging and Enriching Proteins Enables Cell-Specific Proteomics. *Cell Chem Biol* 2016, 23:805-815.
43. Liu J, Xu YQ, Stoleru D, Salic A: Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proc Natl Acad Sci* 2012, 109:413-418.
44. Barrett RM, Liu HW, Jin HH, Goodman RH, Cohen MS: Cell-specific Profiling of Nascent Proteomes Using Orthogonal Enzyme-mediated Puromycin Incorporation. *ACS Chem Biol* 2016, 11:1532-1536.
    - This study describes cell-type-specific metabolic labeling by *O*-propargyl-puromycin.
45. Dieck ST, Kochen L, Hanus C, Heumueller M, Bartnik I, Nassim-Assir B, Merk K, Mosler T, Garg S, Bunse S, et al.: Direct visualization of newly synthesized target proteins in situ. *Nat Meth* 2015, 12:411-+.
46. Rhee HW, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA, Ting AY: Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* 2013, 339:1328-1331.

47. Hung V, Udeshi ND, Lam SS, Loh KH, Cox KJ, Pedram K, Carr SA, Ting AY: Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat Prot* 2016, 11:456-475.
48. Chen CL, Hu YH, Udeshi ND, Lau TY, Wirtz-Peitz F, He L, Ting AY, Carr SA, Perrimon N: Proteomic mapping in live *Drosophila* tissues using an engineered ascorbate peroxidase. *Proc Natl Acad Sci* 2015, 112:12093-12098.
    - This study reports the development of APEX, an engineered ascorbate peroxidase that biotinylates proximal proteins. The authors further report a database that inventories mitochondrial proteins annotated at the sub-compartment level.
49. Jan CH, Williams CC, Weissman JS: Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science* 2014, 346:716-+.
50. Williams CC, Jan CH, Weissman JS: Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling. *Science* 2014, 346:748-751.
51. Tian RJ: Exploring intercellular signaling by proteomic approaches. *Proteomics* 2014, 14:498-512.
52. Szychowski J, Mahdavi A, Hodas JJL, Bagert JD, Ngo JT, Landgraf P, Dieterich DC, Schuman EM, Tirrell DA: Cleavable Biotin Probes for Labeling of Biomolecules via Azide-Alkyne Cycloaddition. *J Am Chem Soc* 2010, 132:18351-18360.

Chapter 2

HOW TO DO A CELL-SELECTIVE PROTEOMICS EXPERIMENT USING BIOORTHOGONAL NONCANONICAL AMINO ACID TAGGING (BONCAT)

2.1 Abstract

This protocol describes how to cell-selectively tag and identify newly-synthesized proteins using azide-containing noncanonical amino acids. Tagged proteins can be analyzed with conventional biochemical methods or liquid chromatography-tandem mass spectrometry (LC-MS/MS). The protocol involves an initial cloning step, but the tagging, detection, and enrichment steps can proceed over 3-5 days. Notably, this protocol does not require depletion of any amino acids in the media and can even be used in animals.

This work was a collaboration with Judy Shon, Graham D. Hamblin, Weslee S. Glenn, Brett M. Babin, Kai P. Yuet, Alborz Mahdavi, Annie Moradian, Michael J. Sweredoski, and Sonja Hess.

2.2 Introduction

Experiments that require proteomic analysis of subpopulations of cells within a complex multicellular environment often can be like looking for needles in a haystack. Whether an investigator is looking at proteins secreted by a pathogen into host cells (1), or defining the proteome of particular cell types such as neurons or glia within an entire organism (2, 3), cell-selective metabolic labeling methods are useful tools for distinguishing the needles from the hay.

Our lab has developed cell-selective bioorthogonal noncanonical amino acid tagging (BONCAT) as a metabolic labeling method that can target specific subpopulations of cells. Achieving this specificity requires the expression of a mutant aminoacyl-tRNA synthetase (aaRS) under a promoter that is only active in the cell type of interest (Fig 2.1). The mutant aaRS has been engineered to incorporate a noncanonical amino acid (ncAA) in a residue-selective manner. For example, we have created both a mutant methionyl-tRNA synthetase (MetRS) (1, 4, 5) and phenylalanyl-tRNA synthetase (PheRS) (3), which can incorporate azido analogs of methionine and phenylalanine residues, respectively (Fig 2.2). The azide moiety allows the newly synthesized proteins of only distinct cell types to be identified and enriched from the rest of the proteins using strain-promoted azide-alkyne cycloaddition (SPAAC) (Fig 2.1B) (4, 6). Cells within the population that do not express the mutant aaRS cannot incorporate the ncAA into their proteins. Because the mutant aaRSs charge tRNAs with the ncAA at high enough rates to compete with charging of canonical substrates, this method does not require depletion

of amino acids or use of minimal medium, and can be accomplished in whole organisms. Our goal is to make this method straightforward so that even non-specialized laboratories can perform a cell-selective proteomic experiment. Mutant aaRSs are available through Addgene, and all reagents are commercially available. Analyzing subpopulations of cells does not have to be nearly as daunting as our original needles and haystack analogy suggests.

Applications of the method

While complete depletion of a canonical amino acid may perturb biological systems, BONCAT is non-toxic and has no obvious effect on protein synthesis or degradation in media that has canonical amino acids present (7, 8). Previous BONCAT experiments that were not cell-selective relied on the cell's endogenous protein synthesis machinery to incorporate ncAAs such as azidohomoalanine (Aha) and homopropargylglycine (Hpg) (Fig 2.2). To achieve high incorporation with these ncAAs, cells were often first depleted of their natural methionine (Met) reserves. Cell-selective BONCAT does not require such depletion, as the mutant aaRSs generally prefer the ncAAs over the canonical residue. Depending on the expression level and activity of the endogenous aaRS in the organism of interest, optimal labeling is usually achieved at either equimolar concentrations of free natural and noncanonical residues or reasonable 10- to 50-fold excesses of the noncanonical residue. For example, the mutant *E. coli* NLL-MetRS prefers Anl 1.2 times more than Met (4), and the mutant PheRS prefers 4-azidophenylalanine (Azf) 42.7 times more than phenylalanine (Phe) (9). In comparison,

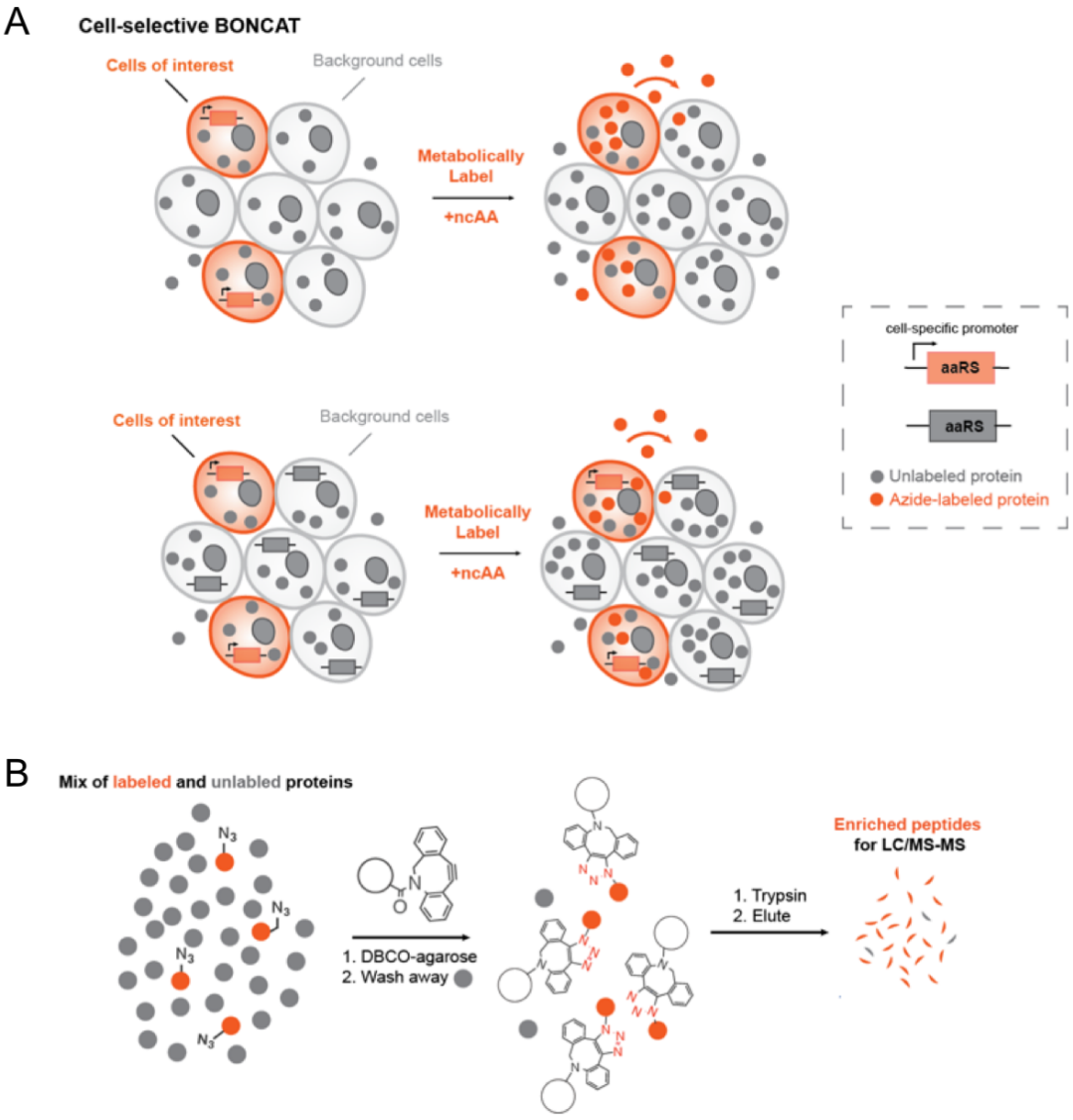


Figure 2.1: A) Scheme depicting two different versions of cell-specific BONCAT. In the first scheme, cells have the mutant aaRS in the presence of cells that do not. In the second scheme, all cells have the mutant aaRS but a cell-specific promoter drives expression only in the cells of interest. B) Scheme depicting strain-promoted azide-alkyne cycloaddition (SPAAC) based enrichment strategy for detection of tagged proteins. Azide-tagged proteins are conjugated to DBCO-agarose beads and untagged proteins can be washed away. On-bead digestion of enriched proteins reveals many peptides that can be analyzed using liquid chromatography tandem mass spectrometry (LC-MS/MS).

the k_{cat}/ k_M (a measure of substrate preference) is 400-fold higher for the natural substrate of wild-type *E. coli* MetRS than for the ncAA analog incorporates Aha (6), which explains why previous BONCAT experiments needed to deplete cells of Met and add prohibitively high Aha concentrations in order to achieve a high degree of labeling. The mouse L274GMetRS was engineered to incorporate Anl about 4 times less than its natural substrate Met when the two compounds are in equal concentrations (2, 10). Cell-selective BONCAT can work under normal cell growth conditions and has even been used to label proteins inside living animals (2).

While we use strain-promoted azide-alkyne cycloaddition (SPAAC) to enrich azide-labeled proteins in this protocol, we envision similar design strategies may also be useful for other click chemistries, such as tetrazine ligation or copper-catalyzed alkyne-azide cycloadditions (CuAAC). Furthermore, this method is compatible with other metabolic labeling methods, including stable isotope labeling with amino acids in culture (SILAC) for quantitative proteomics (3).

Comparisons with other methods

The design strategies outlined here apply to other useful metabolic labeling methods in the literature, such as stochastic orthogonal recoding of translation (SORT) (11, 12). Physical sorting methods like fluorescence activated cell sorting (FACS) are required when cells come from clinical samples or are genetically intractable; however, these methods can produce artifacts during the sample preparation process.

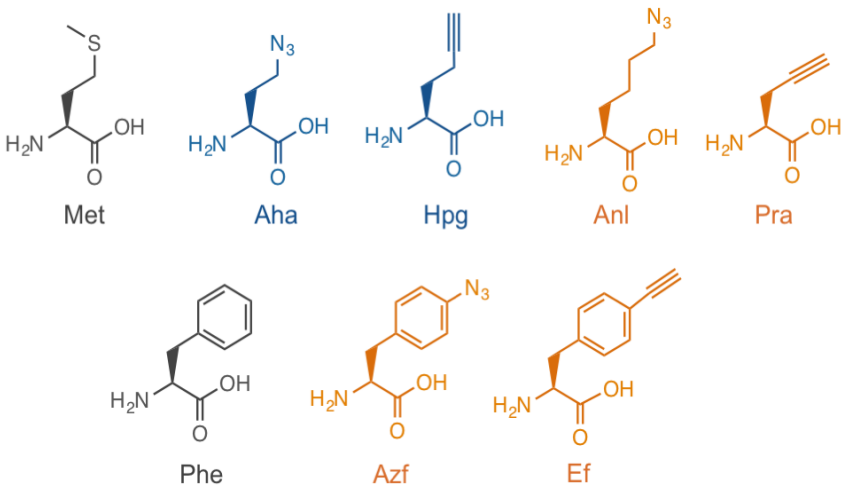


Figure 2.2: Noncanonical amino acids (ncAAs) and the corresponding residue they replace. Orange residues require the expression of a mutant synthetase in order to be incorporated into proteins, and thus, can be used for cell-selective BONCAT. Blue residues can be incorporated by cell's endogenous protein synthesis machinery and cannot be used for cell-selective BONCAT. **Met**: methionine. **Aha**: azidohomoalanine. **Hpg**: homopropargylglycine. **Anl**: azidonorleucine. **Pra**: propargylglycine. **Phe**: phenylalanine. **Azf**: 4-azido-phenylalanine. **Ef**: 4-ethynyl-phenylalanine

Limitations of the method

Cell-selective BONCAT can only be used in genetically-tractable systems due to its reliance on the expression of a mutant aaRS. Promoters that are only active in the cellular subset of interest must also be established or characterized prior to use. In order to initiate labeling, one needs a way to deliver the ncAA to the cells of interest in sufficient quantities. Finally, the investigator must have the organism's genome available to theoretically digest the proteome and compare it to identified fragments.

Design of a cell-specific BONCAT experiment

Figure 2.3 illustrates the workflow of the entire procedure, and the details are described below. The investigator must first choose a unique promoter from the cell type of interest and clone the mutant synthetase to be under the control of that promoter. Next, expression of the synthetase and labeling efficiency should be tested *in vitro*; an appropriate negative control is required. Labeling should be compared with cells lacking the mutant synthetase by using click chemistry in combination with either in-gel fluorescence or fluorescence microscopy. After confirmation of cell-type specific expression, the experiment should be run in (at least) biological triplicate with enough material for the proteomics enrichment.

The choice of synthetase may be limited by the cell system of interest. We recommend testing both the MetRS and PheRS to see which system exhibits strong labeling without compromising cellular growth. Elliot *et al.* suggest that some bias in the set of proteins identified can depend on the residue chosen for noncanonical replacement and bioorthogonal reaction (11); labeling with the mutant PheRS or mutant MetRS may also exhibit a similar bias.

It is important to note that total replacement of the chosen residue is not necessary or probably even desirable for detection and enrichment of proteins from the rest of the cellular milieu. In fact, one ncAA per protein is ideal as it is enough to attach the protein to the enrichment resin, and would minimize perturbation of function and dynamics.

This approach also leaves the rest of the protein to be digested by trypsin for detection and identification in the mass spectrometer. In our experience, searching for the site of ncAA incorporation is difficult and unnecessary for identification and quantification of cell-specific proteins.

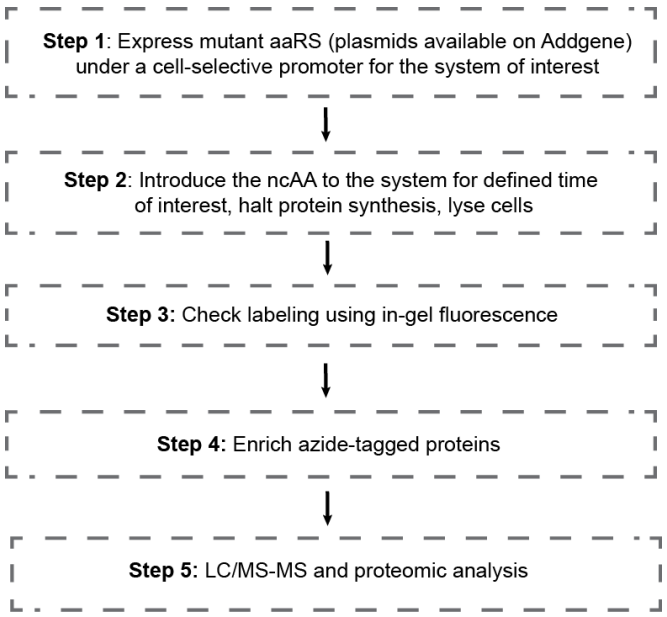


Figure 2.3: Workflow of cell-selective BONCAT protocol outlined in the text.

Materials

Reagents:

- DNA for cloning the mutant aaRS into cells of interest (MetRS and PheRS variants for prokaryotic and eukaryotic expression available at https://www.addgene.org/David_Tirrell/)
- Corresponding azido-noncanonical
    - L-azidolysine hydrochloride (Iris Biotech cat. no. HAA1625)
    - 4-azido-L-phenylalanine hydrochloride (Iris Biotech cat. no. HAA1850)
- 4-12% NuPage Bis-Tris polyacrylamide gels (Thermo Fisher cat. no NP0322BOX) or similar
- Complete EDTA-free Protease Inhibitor Tablets (Roche, cat. no. 1873580)
- Dithiothreitol (DTT)
- Benzonase (Sigma, cat. no. E1014)
- Iodoacetamide (Sigma, cat. no. I1149): 0.5 M in molecular biology grade water, prepare fresh before use
- DBCO-TAMRA (Click Chemistry Tools, cat. no. A131)
- 4x Laemmli Sample Buffer (Bio Rad cat. no. 161-0747)
- InstantBlue (Expedeon cat. no. ISB1L)
- Protein ladder (such as SeeBlue)
- Destain solution (10% acetic acid, 40% methanol, 50% ddH_2O)
- SDS wash buffer: 0.8% SDS, 0.15 M NaCl, in 100 mM Tris pH 8.0
- Urea wash buffer: 8 M urea, 0.15 M NaCl, in 100 mM Tris pH 8.0
- DBCO-Agarose (Click Chemistry Tools, 1034-2)
- Poly-Prep Chromatography Columns (Bio-Rad #7311550)
- Sequencing-grade Trypsin (Promega, cat. no. V5111)
- Ammonium bicarbonate
- Formic Acid
- DMSO
- Acetonitrile
- HiPPR Detergent Removal Spin Column Kit (Thermo Fisher cat. no. 88305)
- Bicinchoninic acid (BCA) protein quantification (Pierce, cat. no 23227)

Equipment:

- Protein gel electrophoresis system
- Gel imaging system (such as GE Healthcare's Typhoon gel imager)

- Rotator
- Rocker
- Temperature-controlled shaker for microcentrifuge tubes (such as Thermomixer from Eppendorf)
- Speedvac
- Orbitrap MS

Software tools:

- MaxQuant (freely available (13))

Procedure

Preparation of solutions

- DBCO-TAMRA stock solution (5 mM): Dissolve DBCO-TAMRA in DMSO to obtain a final concentration of 5 mM (1000x stock). Prepare 10 µL aliquots of the stock solution in individual tubes and store them at −20 °C for up to 2 years. DBCO-TAMRA is light sensitive and should be kept in the dark.
- Trypsin: Resuspend 500 µg in 50 µL 1 N HCl in HPLC grade water. Prepare 5 µL aliquots of this 10 µg/µL solution, which can be stored at -20 °C for up to 1 year.

Step 1: Plasmid construction and expression of the mutant aaRS

1. Obtain the desired cell line or organism required for the experiment.
2. Insert the aaRS gene in a vector of your choice under a cell-specific promoter. As a negative control, do not insert the aaRS.
   Depending on the species of interest and the species of the aaRS, codon-optimization may be necessary. If the aaRS does not express well in your species of interest, check the codon usage of each.
   In this protocol, we have cloned a codon-optimized version of the *E. coli* NLL-MetRS (Addgene #51401) into *Staphylococcus aureus* under a tetracycline-

inducible promoter (Addgene #26252). Our control is this vector without the inserted MetRS.

Critical Step: The choice of promoter is very important, as it will determine the level of specificity of the experiment. It is recommended to check expression and cell specificity of the chosen promoter throughout the labeling period using a fluorescent protein or Western blot. Even a small degree of non-specific expression of the aaRS will result in off-target synthesis of tagged proteins and hinder results.

Step 2: Cell-specific labeling

5. Culture the cells as appropriate. There is no need to deplete the cells of amino acids.
6. Add the ncAA for the time period of interest.

   The amount of ncAA to add depends on the labeling time, the concentration of free amino acids in the growth medium, the rate of protein synthesis, and the natural expression level of endogenous MetRS. For most systems, we generally add 30x the amount of free Met when using Anl and equimolar amounts of Azf and Phe in solution. Using higher concentrations will increase labeling but also increase the probability of deleterious effects such as inhibition of growth. Investigators may wish to determine the optimal ncAA concentration and labeling time for each cell type on a small scale first.

   For example, the free methionine concentration in serum is ~30 μM; therefore, concentrations of 1 mM Anl procure sufficient cell-specific labeling for downstream analysis (13).
7. To stop labeling, add both a protein synthesis inhibitor, such as chloramphenicol (10 μg/mL) for bacterial cells or cycloheximide for mammalian cells (100 μg/mL), and protease inhibitors (1x) for 5 min.

   Without this step, protein labeling will continue during lysis and separation, and stress-response proteins will be found highly expressed.

8. Lyse cells in the presence of an alkylation agent.

   For *E. coli*, we often spin down the cells at 3500 rcf (relative centrifugal force, g) for 5 min, then lyse in 10% of the original volume in 2% SDS in 100 mM Tris (pH 8.0) with 100 mM chloroacetamide. For mammalian cells, the addition of 2% SDS in 100 mM Tris (pH 8.0) and 100 mM chloroacetamide lyses the cells. Heat the lysates at 65 °C to ensure alkylation of thiols.

   Critical: The addition of chloroacetamide will alkylate free thiols, which would nonspecifically react with cyclooctyne reagents (14). We have also had success with iodoacetamide and use the two interchangeably. Both of these solutions need to be made fresh on the day of the experiment and kept from light.

   If the lysates are highly viscous due to the presence of DNA, sonication or benzonase treatment should be performed.

   We recommend saving 10% of the volume of these labeled lysates in a separate small aliquot to test labeling efficiency using DBCO-TAMRA prior to enrichment. Additionally, they can be used as an "unenriched" control to approximate the degree of enrichment by using filter-aided sample preparation (FASP) (15) prior to LC-MS/MS analysis.

**Pause point: Labeled cell lysates can be stored at -80 °C for several months without any harmful effect on click chemistry enrichments.

Step 3: Check protein concentration and degree of labeling of small aliquot

9. Thaw lysates (if frozen). Spin 5 min at 12-14k rcf to clarify and keep the supernatant.
10. Perform BCA assay or other method of determining protein concentration as per manufacturer's instructions.
11. Take 20 µg of protein from each sample and bring to 5 µM DBCO-TAMRA for 15 min.
12. Add 4x Laemmli sample buffer with 50 mM DTT and boil for 2 min.

13. Allow samples to cool to <50 °C, then load ~10 μg into protein gel wells, along with a protein ladder.
14. Run protein gel at constant voltage (170 V) for 1 hour.
15. Carefully remove protein gel from cast and submerge in Destain solution (enough to cover the gel). Leave the gel in this solution on a rocker, covered from light, for at least 4 hours at room temperature, to both remove leftover unbound dye and fix the proteins within the gel.

**Pause point: The gel can be left for up to several days in this Destain solution as long as it is kept from light prior to imaging.

16. After disposing of the Destain solution, allow the gel to rehydrate in deionized water (dH_2O) for 15 min, then visualize the gel using a gel imaging system with appropriate laser and bandpass filter settings. For TAMRA (λ_{ex} = 555 nm and λ_{em} = 580 nm), we excite with a green laser at 532 nm and detect signal with a 580 band-pass 30 nm filter. Only the cells that expressed the mutant aaRS should show labeling in the protein lanes, exemplified in Fig 2.4.
17. Stain for total protein by adding enough InstantBlue to cover the gel for 15 min. Protein bands should become visible by eye, and can be imaged on the gel imaging system. Protein loading should be about equal between lanes.

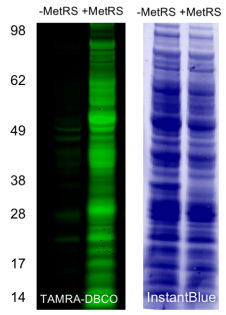


Figure 2.4: TAMRA gel of a cell-selective BONCAT experiment. Cells expressing the NLL-MetRS incorporate azides into their proteins, which can be conjugated to DBCO-TAMRA

Step 4: Enrichment

18. Thaw lysates (if frozen). Spin 5 min at 12-14k rcf to clarify and keep the supernatant.
19. Add equal volume 8 M urea/0.15 M NaCl/protease inhibitor in PBS, made fresh so the protein is resuspended in buffer with 0.5-1% SDS, and 4 M urea.
20. Wash ~25 µL per sample DBCO-agarose in 1 mL of 1% SDS three times. Resuspend in original volume of 1% SDS.
    Critical: Always centrifuge the resin at 1500 rcf or less. Spinning at faster speeds can result in destruction of the resin. We generally use ~25 µL of resin / 5mg enrichment. Using high amounts of resin results in much higher proteomic background.
21. Add washed resin to samples and rotate end over end for 16-24 hours.
22. Wash resin, now with covalently bound proteins, with 1 mL wash buffer to remove unbound proteins
    We often keep the supernatants at this step, in case the proteins did not bind.
23. Reduce bound proteins with 0.5 mL 5 mM DTT in SDS wash buffer for 30 min. Remove supernatant.
24. Alkylate bound proteins with 100 mM chloroacetamide or iodoacetamide in the dark for 45 min at 50 °C.
25. Transfer resin to Poly-prep chromatography column.

26. Wash resin with the following solutions:
    a. 8 x 5 mL SDS wash buffer
    b. 8 x 5 mL 8 M urea in 100 mM Tris pH 8.0
    c. 8 x 5 mL 20% acetonitrile
27. Transfer beads to Eppendorf tubes using 10% ACN in 50 mM Ammonium Bicarbonate (AmBi).
28. Spin 5 min at 1500 rcf, remove supernatant down to 100 µL.
29. Add 100 ng trypsin to each sample and incubate overnight at 37 °C.
30. Collect supernatant, then wash resin with 150 µL 20% acetonitrile twice and combine washes with supernatant. Be careful to avoid carrying over resin during transfer steps.
31. Speedvac to dryness.

**Pause point: Peptides can be stored at -20 °C for several months without any harmful effects.

32. Follow StageTip protocol to desalt peptides (17).
33. Resuspend in 8 µL 0.2% formic acid for injection onto the LC-MS/MS

Step 5: LC-MS/MS and proteomic analysis (~8-48 hours, depending on the number of samples and replicates)

34. Inject ~100 ng of the enriched lysate onto a liquid chromatography system coupled to an Orbitrap mass spectrometer, equipped with a nano-electrospray ion source.
35. Separate the peptide using a chromatographic separation for 1-3 hours using an elution gradient from 2 to 30% acetonitrile at a flow rate of 220 nL/min, and operate the mass spectrometry in data-dependent mode (18).
    We typically collect full scan mass spectra with 400-1600 m/z, and collect the top 20 most intense ions from the survey scan for fragmentation in the linear ion

trap by collision-induced dissociation (CID). We use precursor ion charge state screening to reject singly charged and unassigned charge states.

36. Take the raw files from the LC-MS/MS run and process using MaxQuant as previously described (13, 19).

    Set carbamidomethylation of cysteine as a fixed modification, and protein N-terminal acetylation, N-terminal formylation, and methionine oxidation as variable modifications. We also include variable modifications of methionine corresponding to Anl (+23.0450) and reduced Anl (-2.9455), but often do not find many sites of Anl labeling because they are left on the resin during the digestion step.

Table 2.1: Troubleshooting Table

Step	Problem	Possible Reason	Solution
8	Viscous samples after lysis	Insufficient genomic DNA lysis	Use more Benzonase to facilitate lysis of genomic DNA, shear DNA with a syringe and a needle, or sonicate samples
17	Low levels of tagged proteins	Labeling time too short or not enough ncAA added	Add the ncAA for a longer time or increase concentration
35	Few proteins found, LC trace has most signal towards end of run	Leftover detergent in samples	Use HiPPR detergent removal columns or increase the number of washes of the resin
35	PEG present in LC-MS	PEG contamination	Use as few transfers of the material as possible into new Eppendorf tubes, as the sample will pick up trace plasticizers over time, do not use autoclaved plastics as this increases the amount of plastics

Timing

Step 1: Cloning of aaRS into cells of interest (variable)

Varies depending on system chosen

Step 2: Cell labeling and lysis (1-24 hours)

Varies depending on time window of interest to label newly synthesized proteins

Step 3: Testing labeling using in-gel fluorescence (6 hours)

Step 4: Enrichment and preparation of proteins for MS (48 hours)

Step 5: LC-MS/MS and data analysis (2 days to 2 weeks**)**

Anticipated Results

Using this method, we often find 1000-5000 proteins from the cell-type of interest.

References

1. Mahdavi A, *et al*. (2014) Identification of secreted bacterial proteins by noncanonical amino acid tagging. *Proc Natl Acad Sci* 111(1):433-438.
2. Erdmann I, *et al*. (2015) Cell-selective labelling of proteomes in *Drosophila melanogaster*. *Nat Comm* 6.
3. Yuet KP, *et al*. (2015) Cell-specific proteomic analysis in *Caenorhabditis elegans*. *Proc Natl Acad Sci* 112(9):2705-2710.
4. Tanrikulu IC, Schmitt E, Mechulam Y, Goddard WA, & Tirrell DA (2009) Discovery of *Escherichia coli* methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine *in vivo*. *Proc Natl Acad Sci* 106(36):15285-15290.
5. Ngo JT, *et al*. (2009) Cell-selective metabolic labeling of proteins. *Nat Chem Biol* 5(10):715-717.
6. Kiick KL, Saxon E, Tirrell DA, & Bertozzi CR (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc Natl Acad Sci* 99(1):19-24.
7. Bagert JD, *et al*. (2014) Quantitative, Time-Resolved Proteomic Analysis by Combining Bioorthogonal Noncanonical Amino Acid Tagging and Pulsed Stable Isotope Labeling by Amino Acids in Cell Culture. *Mol Cell Prot* 13(5):1352-1358.
8. Wang JG, *et al*. (2017) Nonradioactive quantification of autophagic protein degradation with (L)-azidohomoalanine labeling. *Nat Protoc* 12(2):279-288.
9. Carrico IS (2004) Protein engineering through *in vivo* incorporation of phenylalanine analogs. Thesis Dissertation. Chapter 4.

10. Mahdavi A, *et al*. (2016) Engineered Aminoacyl-tRNA Synthetase for Cell-Selective Analysis of Mammalian Protein Synthesis. *J Am Chem Soc* 138(13):4278-4281.
11. Elliott TS, Bianco A, Townsley FM, Fried SD, & Chin JW (2016) Tagging and Enriching Proteins Enables Cell-Specific Proteomics. *Cell Chem Biol* 23(7):805-815.
12. Elliott TS, *et al*. (2014) Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat Biotech* 32(5):465-U186.
13. Stein WH & Moore S (1954) The Free Amino Acids of Human Blood Plasma. *J Biol Chem* 211(2):915-926.
14. van Geel R, Pruijn GJM, van Delft FL, & Boelens WC (2012) Preventing Thiol-Yne Addition Improves the Specificity of Strain-Promoted Azide-Alkyne Cycloaddition. *Bioconj Chem* 23(3):392-398.
15. Wisniewski JR, Zougman A, Nagaraj N, & Mann M (2009) Universal sample preparation method for proteome analysis. *Nat Meth* 6(5):359-U360.

Chapter 3

CELL-SELECTIVE PROTEOMICS OF METHICILLIN-RESISTANT *STAPHYLOCOCCUS AUREUS* (MRSA) INFECTION IN MICE IDENTIFIES A NOVEL ANTI-VIRULENCE TARGET

3.1 Abstract

Methicillin-resistant *Staphylococcus aureus* (MRSA) poses a threat to human health and is becoming increasingly resistant to current antibiotics. Characterizing the MRSA proteome during an infection *in vivo* would reveal important information about the process of infection and potentially reveal new strategies to fight the disease. Currently available proteomic techniques, however, are not capable of efficiently measuring pathogen proteins among the highly abundant host protein background that is inherent to infection models. In this work we use bioorthogonal noncanonical amino acid tagging (BONCAT) to perform cell-selective proteomic analysis of MRSA in a mouse skin infection model, identifying 766 MRSA proteins synthesized during infection. Quantitative analysis of our dataset identifies novel virulence factors that are upregulated during the course of infection. Deletion of AdhE, one of the top BONCAT hit proteins, led to a significant decrease in MRSA virulence phenotype during infection. Furthermore, a point mutation within AdhE reduced virulence to levels similar to that of the full deletion, indicating a novel site for targeted anti-virulence therapies. Overall, this work demonstrates the importance of cell-selective chemoproteomic labeling *in vivo* and provides insight into the pathogenesis of MRSA infections, revealing several promising targets for anti-virulence therapy development.

This work was a collaborative effort with Arya Khosravi, Brett M. Babin, Michael J. Sweredoski, John D. Bagert, Will H. DePas, Annie Moradian, Alborz Mahdavi, Hwan Kim, Fabiana Falugi, Sonja Hess, Dianne K. Newman, Olaf Schneewind, Dominique Missiakas, and Sarkis K. Mazmanian.

Significance Statement

This work uses a novel chemoproteomic method to tag and detect newly synthesized proteins by a pathogen in a mouse model of skin infection. By specifically incorporating a label into proteins synthesized by methicillin-resistant *Staphylococcus aureus* (MRSA) as it infects its mouse host, proteins important for nutrient acquisition, metabolism, and pathogenesis were identified. We tested whether these proteins contributed to infection by assaying strains of MRSA with single deletions in them. One deletion demonstrated a significant decrease in infection, indicating that this protein contributes to MRSA pathogenesis and could be a novel target for antibiotics. Overall, our findings underscore the ability of bioorthogonal noncanonical amino acid tagging (BONCAT) to "fish out" pathogenic proteins from more abundant host ones, suggesting that this strategy could be used for other infectious pathogens.

3.2 Introduction

Methicillin-resistant *Staphylococcus aureus* (MRSA), a pathogen that causes life-threatening infections, poses a serious human health threat due to the increasing numbers of multidrug resistant strains and a lack of new antibiotics (1, 2). Anti-infective drugs, which target bacteria during invasion, have been suggested as alternative or adjunct therapies to antibiotics, as they increase bacterial susceptibility (3). By targeting factors uniquely expressed *in vivo*, such antibiotics would not affect bacteria growing outside the host, thus decreasing the possibility of evolved resistance (4). A catalog of virulence factors uniquely expressed during infection is required for continual efforts to

develop anti-infective agents. Elucidating the *in vivo* staphylococcal proteome during infection, however, is challenging due to the technical limitations of shotgun proteomics: the relative overabundance of host proteins occupy most of a mass spectrometer's bandwidth, which strongly represses the detection of proteins from the pathogen. Previous studies have attempted to overcome this limitation by enriching for bacteria using cell sorting. These approaches, however, are prone to artifacts during sample preparation, are blind to secreted proteins, and are not amenable to certain biological systems (5, 6). Other studies attempt to mimic the nutrient status of the host in test tubes, but do not fully recapitulate the complex environment of the host (7, 8).

Bioorthogonal noncanonical amino acid tagging (BONCAT) is a chemoproteomic technique that enables the cell-selective and temporal labeling of the cellular proteome (9). With this technique, a mutant aminoacyl-tRNA synthetase expressed solely by the cells of interest allows for selective incorporation of a noncanonical amino acid (ncAA) with a bioorthogonal handle into proteins for subsequent detection and identification. Because cellular proteins are only labeled during the ncAA pulse, BONCAT provides a temporally precise overview of the cell's proteome. The chemical handles on the ncAAs, often azides or alkynes, can then be conjugated to fluorophores or enrichment tags using copper-catalyzed alkyne-azide cycloaddition (CuAAC) or strain-promoted alkyne-azide cycloaddition (SPAAC) (10, 11). Notably, all of these reagents including the ncAAs are commercially available. BONCAT has previously been applied to studies of microbial pathogenesis in *in vitro* models of infection (11-15). Herein, we adapted BONCAT to a

mouse model of skin infection and identified the proteome of MRSA *in vivo* (Fig 3.1).

Our findings highlight the ability of BONCAT to selectively enrich for pathogenic proteins from the highly abundant host protein background in the complex milieu of an active skin infection. Labeling is pathogen-specific and can be achieved within 4 hours and subsequently used to enrich for staphylococcal proteins made within a host. Testing potential candidates from this list for virulence defects *in vivo* led to the discovery of a novel protein important for MRSA infection. We expect that this unbiased approach to label pathogenic proteins could be used as a global discovery tool for novel anti-infective strategies since it is compatible with other pathogens and modes of infection, and uses commercially available reagents.

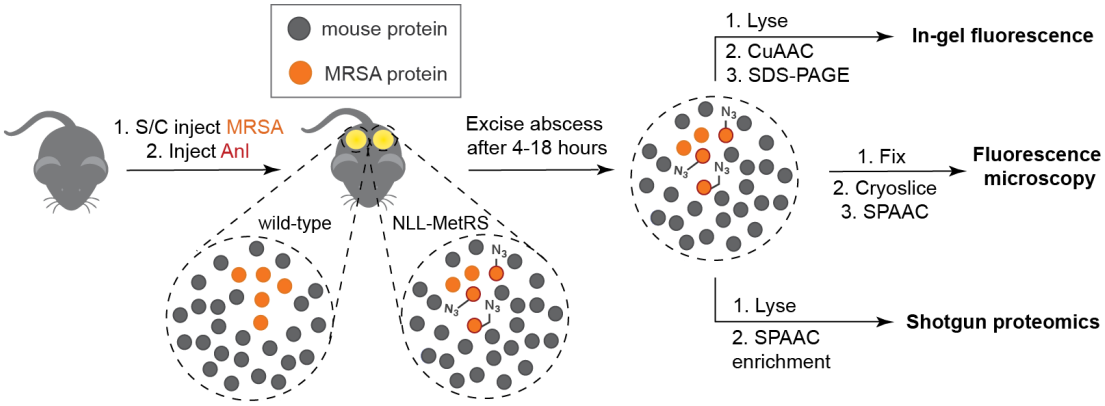


Figure 3.1: Schematic depicting BONCAT to label MRSA in a mouse model of skin infection. Strains of MRSA that do not express NLL-MetRS are unable to incorporate azidonorleucine (Anl) into proteins. Azide-labeled proteins are subsequently visualized using fluorescence microscopy or enriched for proteomic detection.

3.3 Results

Adapting cell-selective BONCAT for MRSA. We first needed to adapt cell-selective BONCAT for use in MRSA. We opted to use a mutant methionyl-tRNA synthetase (NLL-MetRS) previously employed in studies of host-pathogen interactions that allows for the incorporation of azidonorleucine (Anl) in place of methionine residues (16). The *Escherichia coli (Ec)* NLL-MetRS codon-optimized for *Staphylococcus aureus* was inserted into a staphylococcal shuttle plasmid under control of the hprK/lgt promoter, which has been shown to be constitutively active, including during infections *in vivo* (17-19). This plasmid (+NLL) was transformed into the virulent MRSA strain USA300, which accounts for up to 98% of reported skin and soft tissue infections (SSTIs) in hospitals (1). A plasmid without the inserted NLL-MetRS (–NLL) was used as an empty-vector control for all characterizations.

To assess the ability of the *Ec* NLL-MetRS to charge MRSA tRNAs with azidonorleucine (Anl), +NLL and –NLL cultures were grown in tryptic soy broth (TSB) and 2 mM of Anl was added for 1 hour during exponential growth. Protein lysate from treated cells was first conjugated to alkyne-functionalized tetramethylrhodamine (alkyne-TAMRA, Fig S3.1) using CuAAC and then separated by SDS-PAGE for subsequent in-gel fluorescence scanning (Fig S3.2A). Gels revealed selective labeling in the +NLL samples only, signifying that the *Ec* NLL-MetRS had successfully charged Anl onto tRNAs and labeled newly-synthesized proteins within MRSA. Nonspecific incorporation of Anl into the proteome of –NLL strains was not detected. Importantly,

growth did not appear to be affected, even with long (12 hour) pulses of Anl (Fig S3.2B).

Characterizing cell-selective BONCAT labeling during skin infection in a mouse. The +NLL/-NLL strains of MRSA were then used in a mouse model of early skin infection (4). Following subcutaneous (S/C) injection of the +NLL or –NLL strains along with 25 μL of 25 mM Anl, we excised and homogenized entire abscesses after 4, 12, and 16 hours of labeling. Conjugation to alkyne-TAMRA followed by in-gel fluorescence revealed Anl labeling at every time point for +NLL MRSA with the strongest labeling occurring 12 to 16 hours post infection (Fig 3.2B, Fig S3.3). As the mouse and the –NLL strain do not express the mutant aminoacyl-tRNA synthetase, Anl is unable to be incorporated into their proteins. Importantly, the size of the lesion formed and the number of colony-forming units (CFUs) obtained from each lesion did not depend on the strain used or on the presence of Anl, suggesting that BONCAT does not perturb the progress of infection (Fig 3.2B).

Visualization of newly synthesized MRSA proteins within skin abscesses. To ensure that Anl labeling was confined to MRSA, we infected with +NLL or –NLL strains of MRSA and subcutaneously injected Anl at the time of infection. After waiting 16 hours to allow for incorporation of Anl into newly synthesized proteins, we excised and cryosectioned whole skin abscesses. We detected Anl labeling by conjugation via SPAAC to aza-dibenzocyclooctyne functionalized with AlexaFluor488 (DBCO-AF488,

Fig S3.1) and immunostained for MRSA proteins using a polyclonal anti-*S. aureus* antibody. We observed DBCO-AF488 staining of proteins in the abscess only when the NLL-MetRS was expressed within MRSA and that this staining co-localized with the anti-*S. aureus* counterstain, suggesting that MRSA protein synthesis could be cell-selectively visualized within the abscesses (Fig 3.2D, Fig S3.4).

Enrichment of Anl-labeled proteins results in greater proteomic coverage of MRSA proteins synthesized *in vivo*. To identify MRSA proteins within the skin abscesses, we labeled proteins in triplicate during the first 16 hours of subcutaneous infection, excised the abscesses, and homogenized the entire abscess. After cell lysis, we enriched for azide-labeled proteins using aza-dibenzocyclooctyne-conjugated agarose beads (DBCO-agarose). A portion of the sample was reserved to serve as an "unenriched" control for comparison. Enriched and unenriched samples were digested with trypsin and subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

Prior to enrichment, the LC-MS/MS runs were overwhelmed with highly abundant mouse proteins, masking the signal from nearly all MRSA proteins. For example, of all the spectra detected in the unenriched samples, only 3% were assigned to MRSA proteins, and the remaining were attributed to mouse proteins and common contaminants. This finding led to the identification of only 178 MRSA proteins in the unenriched samples.

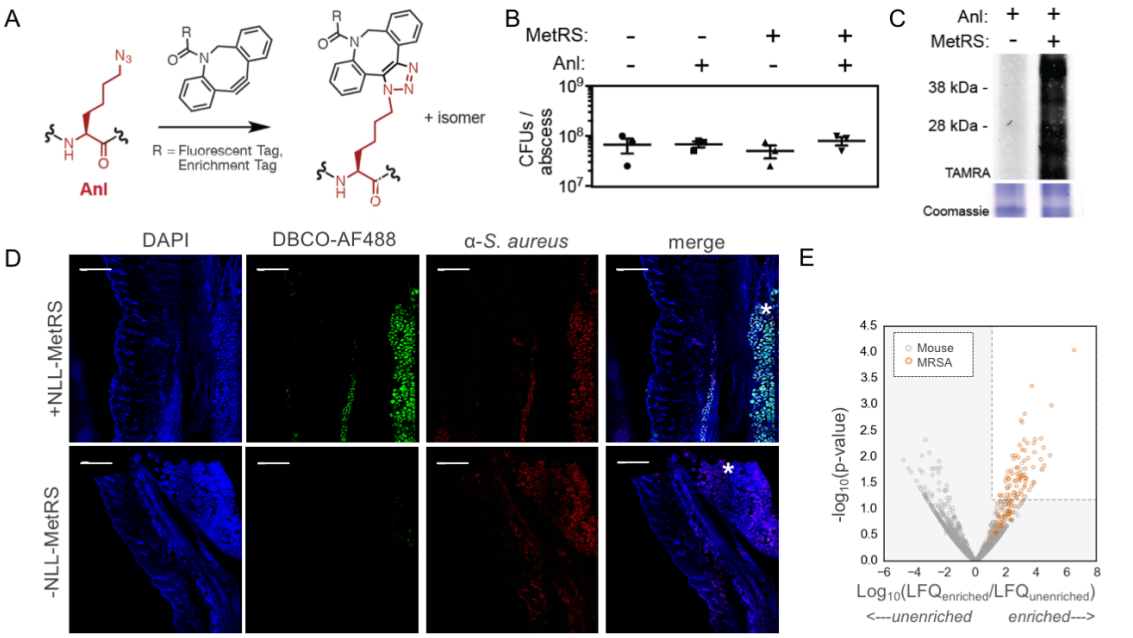


Figure 3.2: BONCAT specifically labels newly synthesized MRSA proteins within skin abscesses. A) Strain-promoted alkyne-azide cycloaddition (SPAAC) for visualization or enrichment of *in vivo* synthesized proteins. B) BONCAT does not affect growth of MRSA *in vivo*. C) Copper-catalyzed alkyne-azide cycloaddition (CuAAC) with TAMRA-alkyne followed by SDS-PAGE enables in-gel fluorescence of labeled proteins. D) Skin infected with +NLL or –NLL MRSA and labeled with Anl for 18 hours. After the abscessed was excised and cryogenically sliced, the samples were blocked, and then reacted to DBCO-AF488, followed by an anti-*S. aureus* antibody. *Indicate abscess. Scale bar is 100 μm. E) Volcano plot of proteins found in both unenriched samples and enriched samples shows that MRSA proteins (orange) are enriched compared to mouse proteins (blue). Most MRSA proteins were uniquely found in enriched samples and do not have fold-change values to exhibit.

With enrichment, 23% of all spectra could be assigned to MRSA proteins, leading to the identification of 766 proteins with at least 1 peptide, which is the largest number of identified MRSA proteins in an *in vivo* infection model to date (Fig S3.5). Additionally,

the identified proteins had greater average intensities, leading to more confident identifications (IDs) and quantifications (Fig S3.5C, S3.5D). Linear Models for Microarray Data (LIMMA), an empirical Bayes (eBayes) method, was then employed to detect significant changes in protein abundances for proteins found before and after enrichment (20, 21). MRSA proteins were significantly more abundant in the enriched samples (Fig 3.2E). Thus, labeling with Anl followed by enrichment increases the number of identified MRSA proteins.

Identifying MRSA proteins important for infection using comparative proteomics. We sought to identify proteins that are expressed exclusively or highly upregulated within the skin and may serve as promising anti-infective targets. We therefore compared the detected MRSA proteins during skin infection to those made during aerobic and anaerobic growth. The +NLL MRSA strain was grown aerobically in triplicate and labeled with Anl for 1 hour during mid-log growth, then lysed and enriched for proteins synthesized during this time. The same experiment was performed using anaerobic cultures, except for being labeled for 2 hours during mid-log growth. Using this strategy, we quantified 655 proteins expressed during skin infection (at least 2 peptides/protein), 925 proteins expressed during aerobic growth, and 1,055 proteins expressed during anaerobic growth (Fig 3.3A, Table S3.4). The 530 proteins detected in all three environments were then compared using LIMMA and confirmed that NLL-MetRS was expressed at similar levels between the replicates (Fig S3.6A), highlighting the constitutive nature of the chosen promoter.

To determine if the proteins found upregulated in the skin were associated with virulence, we used Gene Ontology (GO) analysis. The log-fold changes of the proteins identified during the skin infection were compared to either aerobically or anaerobically grown cells, and a Mann-Whitney U test was performed to determine if the median values for proteins with the given annotation tended to be higher or lower than those without the annotation. Compared to aerobically and anaerobically grown samples, three GO terms were significantly enriched in both: "cellular catabolic process", "pathogenesis", and "small molecule catabolic process" (Fig 3.3B). After plotting the log-fold changes in the skin compared to anaerobic vs. aerobic, proteins annotated with the GO term "pathogenesis" were generally located in the upper right quadrant (Fig 3.3C). The only protein not found in this quadrant, yet annotated as important to pathogenesis, was enolase (Eno), a glycolytic pathway enzyme reported to bind laminin during infection (22). Catabolic processes have been established as important in skin and soft tissue infections due to their low abundance of glucose (23).

Compared to *in vitro* growth, we found many known virulence factors and contributors to infection to be upregulated in our *in vivo* skin infection model (Fig 3B). For example, virulence factors gamma-hemolysin (Hlg), staphylococcal superantigen-like protein 7 (Ssl7), and immunoglobulin-binding protein (Sbi) were uniquely detected during infection (18). Immunoglobulin-binding protein A (Spa) was significantly upregulated *in vivo* compared to aerobic and anaerobic growth conditions. Additionally, staphylococcal accessory regulator SarA and catalase (KatA) were upregulated *in vivo* compared to

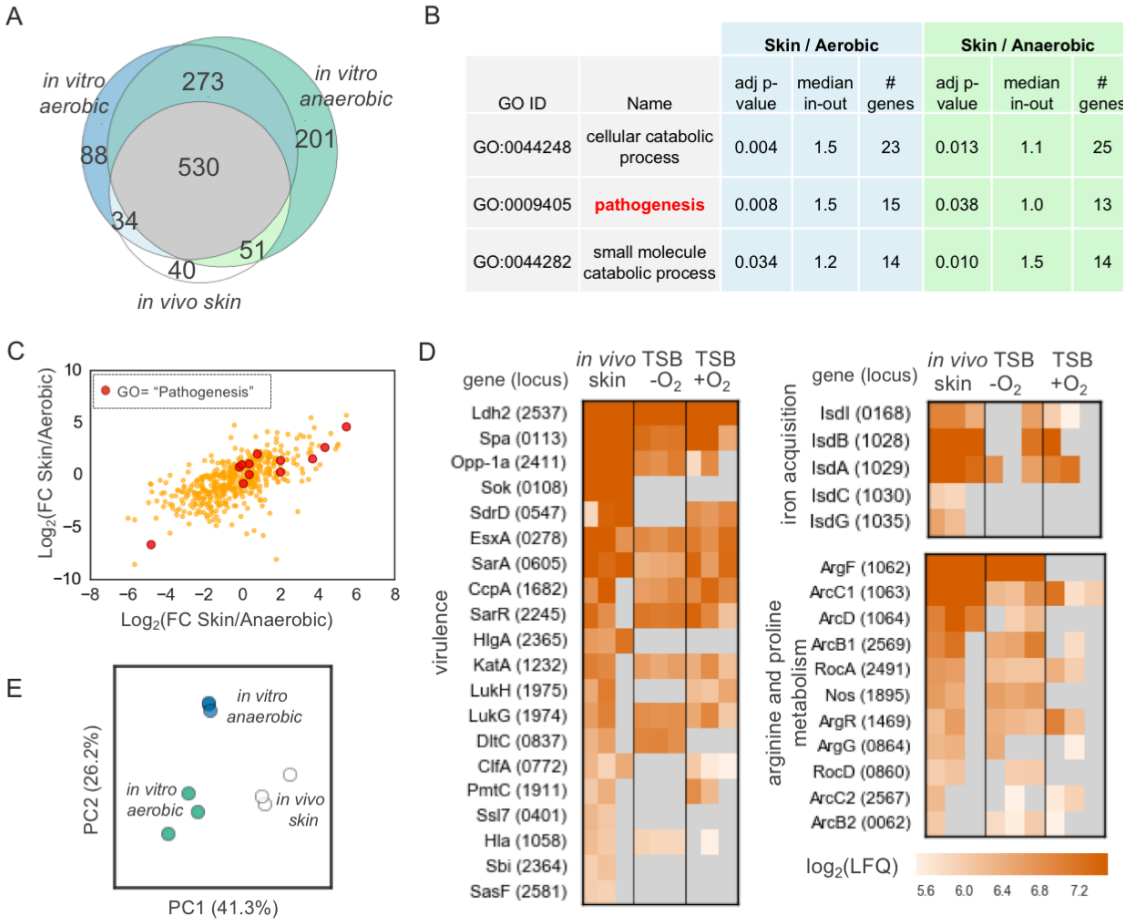
anaerobic growth (Fig 3.3C, S3.7-8) (24).

In addition to these known virulence factors, many genes involved in nutrient scavenging were identified using BONCAT. Five of the genes involved in iron acquisition from heme (IsdA, IsdB, IsdC, IsdG, and IsdI) were upregulated during infection (Fig 3.3D, Fig S3.9), which are important due to low iron availability within the mammalian host (25). Proteins upregulated in the arginine catabolic mobile element (ACME: ArcC2 and ArcB) that are important to withstand the acidic environment of the skin (26), and in the endogenous arginine catabolic pathway (ArgF and ArcC1) that degrade arginine under anaerobic conditions (Fig 3.3D, Fig S3.10), were also detected. Additionally, two proteins involved in catabolizing N-acetylneuraminate, the predominant sialic acid in humans (sialic acid lyase (NanA) and NanR), were found only in the skin (Fig 3.4A, Table S3.4).

S. aureus requires distinct sets of proteins for successful infection of different tissues, whereas few genes are required for the pathogenesis of all types of *S. aureus* infections, suggesting a limitation of why several reported virulence genes were not detected using BONCAT (27). Additionally, certain proteins may be too low abundance, not contain any methionine residues besides the initiator, or not ionize well on the mass spectrometer. For example, accessory gene regulator A (AgrA), fibronectin binding protein A (FnbA), nuclease (Nuc), phenol-soluble modulins (PsmA1, PsmA4: which do not contain any methionine residues besides the initiator Met) (28), and the *S. aureus*

exoprotein expression two-component system SaeRS were not detected in the *in vivo* samples.

Because levels of oxygen vary widely within different tissues in the host (17), we wondered whether our infection model more closely resembled aerobically or anaerobically grown samples. We therefore performed a principal component analysis (PCA) on the proteins found in all samples; the replicates clustered in three separate subspaces (Fig 3.3E, Fig S3.16). Based on the proteomic profile, the *in vivo* environment is a distinct growth setting compared to *in vitro* conditions. Performing PCA on the proteins found in aerobically and anaerobically grown samples revealed a single component that accounted for 71% of the variance between the data sets. A projection of the data from the skin infection replicates into this dimension formed a distinct group (Fig S3.5D).



		Skin / Aerobic			Skin / Anaerobic		
GO ID	Name	adj p-value	median in-out	# genes	adj p-value	median in-out	# genes
GO:0044248	cellular catabolic process	0.004	1.5	23	0.013	1.1	25
GO:0009405	pathogenesis	0.008	1.5	15	0.038	1.0	13
GO:0044282	small molecule catabolic process	0.034	1.2	14	0.010	1.5	14

Figure 3.3: Comparative proteomics reveals MRSA proteins expressed during infection. A) Venn diagram of proteins quantified during aerobic or anaerobic growth *in vitro* or during skin infection *in vivo*. B) Gene Ontology (GO) analysis of upregulated pathways during skin infection revealed the terms "catabolic process" and "pathogenesis". C) Proteins annotated with the GO term "pathogenesis" (red circles) are upregulated compared to aerobic or anaerobic growth. D) Heat maps of proteins annotated to be important to virulence, iron acquisition, or arginine catabolism during all three conditions. E) Principal component analysis (PCA) of proteins found in all samples.

Screening hits from BONCAT analysis reveals AdhE as a novel factor that contributes to virulence. From the list of proteins found upregulated or unique to infection (Table S3.4), 10 proteins that had not been previously associated with infection (Fig 3.4A) (6, 29) were selected. Isogenic knockout mutants of each of these strains, along with known virulence factor sortase (*ΔsrtA*) (2, 22, 25-27), were then used for subsequent studies. SrtA is a target for novel anti-infective therapies due to its ability to display 20 different adhesins on the cell-surface (25, 30). To ascertain if each protein singularly contributed to skin infection, we infected mice using the same model of skin infection that we used to obtain our proteomic results and monitored the size of the lesion formed. On day five of the infection, we excised and homogenized each abscess and performed serial dilutions to count the number of CFUs present.

Of the mutants tested, only one demonstrated a decrease in the severity of infection compared to wild-type MRSA in both the size of the lesion formed (Fig 3.4B, Fig S3.6) and the number of CFUs within the abscess (Fig 3.4C, S3.6). This mutant with a hindered ability to infect has an insertion within the bifunctional aldehyde-alcohol dehydrogenase (AdhE) gene (SAUSA300_0151) (*ΔadhE*). Mutants that exhibited no decrease in infection included an iron-containing alcohol dehydrogenase (*Δadh*), threonine dehydratase (*ΔtdcB*), alanine dehydrogenase (*Δald1*), SAUSA300_2375, SAUSA300_2132, sialic acid lyase (*ΔnanA*), pseudouridine synthase (*ΔpsuG*), and SAUSA300_1393. Notably, deletion of the AdhE protein decreases the pathogen's virulence to a more significant degree than the *ΔsrtA* knockout in this infection model.

To control for possible pleiotropic effects of the insertion, we re-inserted *adhE* into the chromosome under its endogenous promoter using an engineered *attB2* integrase site that does not disrupt any known genes to generate *ΔadhE/adhE*$^{+}$ (31). The virulence of the *ΔadhE* mutant was restored in both lesion size (Fig 3.4D), and number of CFUs recovered (Fig 3.4E) by this complementation strain, suggesting that indeed, this protein was contributing to virulence.

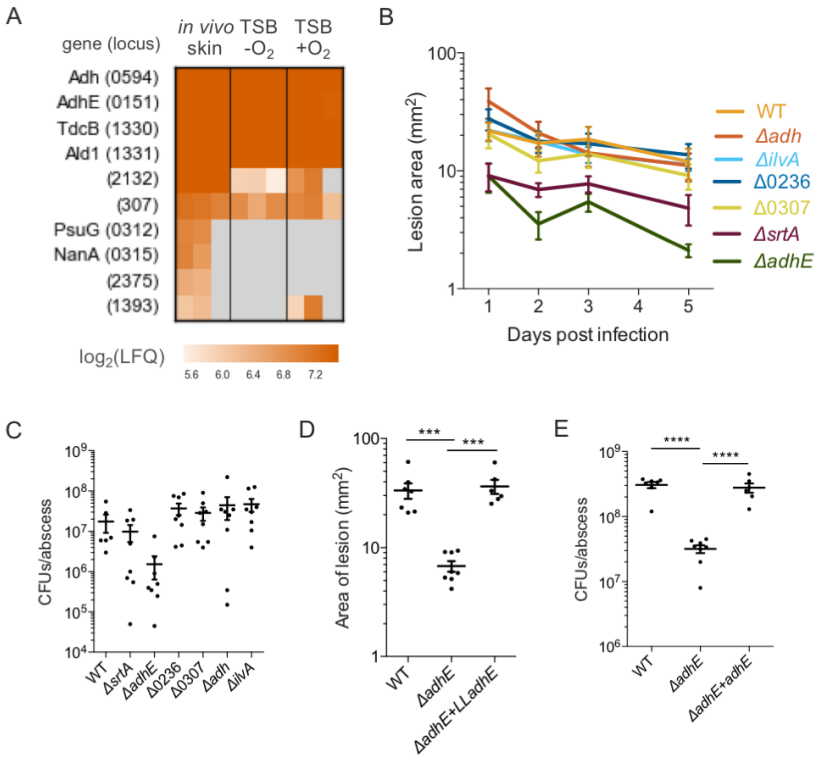


Figure 3.4: Screening of mutant hits from BONCAT analysis reveals AdhE as a novel proteomic factor important for MRSA infection. A) Proteins found during skin infection chosen to test for contributions to virulence. B) Quantification of lesion areas formed over five days after subcutaneous (S/C) injection with 2e6 CFUs of wild-type (WT) USA300 strain JE2 or selected transposon mutants found in BONCAT screen. C) CFUs recovered from abscess in 3.5B. D) Quantification of lesion areas formed with complemented strain of AdhE on day two of skin infection. E) CFUs recovered from abscess in 3.5D.

The *ΔadhE* mutant shows no phenotype *in vitro*. AdhE is an 869-amino acid-long protein important for fermentative growth under redox stress (32) (Fig 3.5A). We found AdhE to be significantly upregulated during anaerobic growth and *in vivo* compared to aerobic conditions. Additionally, it was one of the top proteins that PCA revealed contributed to differences in expression profiles between *in vitro* and *in vivo* samples (Fig S3.16). Although AdhE has not been studied in great detail with respect to *S. aureus*, studies of its homologs in other bacteria such as *E. coli* may afford clues towards its function *in vivo*. AdhE is important for adhesion and peroxide resistance in *E. coli* (33, 34), adhesion in *Listeria monocytogenes (35)*, and virulence in *Streptococcus pneumoniae* (36).

However, *S. aureus ΔadhE* strains showed no defect in hemolysis, pigment formation, or mannitol fermentation (37). Notably, the *ΔadhE* knockout grew the same as wild-type *S. aureus* on all carbon sources tested, including during aerobic and anaerobic growth (Fig S7A), did not change its growth rate in response to peroxide (Fig S3.7B), and did not appear to affect its ability to adhere to cells in culture (Fig S3.7C). This lack of phenotype highlights the inability of *in vitro* models of infection to accurately reflect the phenotype *in vivo* and validates the method of proteomic analysis during infection *in vivo* rather than during infection *in vitro*.

A single point mutation in AdhE decreases *S. aureus* infection. We sought to determine whether the catalytic activity of the AdhE protein contributed to infection. In

Entamoeba histolytica, the alcohol dehydrogenase activity of the C-terminal domain was dependent on the presence of an active N-terminal aldehyde dehydrogenase domain (38). We mutated the cysteine homologous to the reported catalytic cysteine within the aldehyde dehydrogenase domain to an alanine residue (C258A, Fig 3.5B) and inserted this mutant into *ΔadhE* to see if it could recapitulate infection as well as wild-type or the AdhE complemented strain (39). The C258A mutation reduced the levels of infection to the levels of the *ΔadhE*, indicating that a mutation within a single residue of AdhE is enough to decrease virulence (Fig 3.5D, 3.5E, S3.8)

Disulfiram decreases the severity of infection. Knowing that a single mutation drastically decreased the severity of MRSA skin infection, we postulated that targeting this residue could serve as a novel anti-infective strategy. The anti-alcoholism drug disulfiram (Antabuse) is reported to irreversibly carbamylate the conserved catalytic cysteine in the human homolog of the aldehyde dehydrogenase domain of AdhE, ALDH1A1 (40). Disulfiram was also shown to decrease the ability of *E. histolytica* to ferment through inhibition of its AdhE activity (38). The *E. histolytica* AdhE shares 71% positive homology with *S. aureus* AdhE so we tested if it would also decrease MRSA virulence within the skin, even though it has reported weak antibiotic activity *in vitro* (41). We also did not note differences in growth rate in the presence of disulfiram (Fig S3.13). Injection of disulfiram into mice as they were infected with MRSA led to a significant reduction in both the size of the lesion and number of CFUs recovered compared to vehicle (Fig 3.5E, 3.5F).

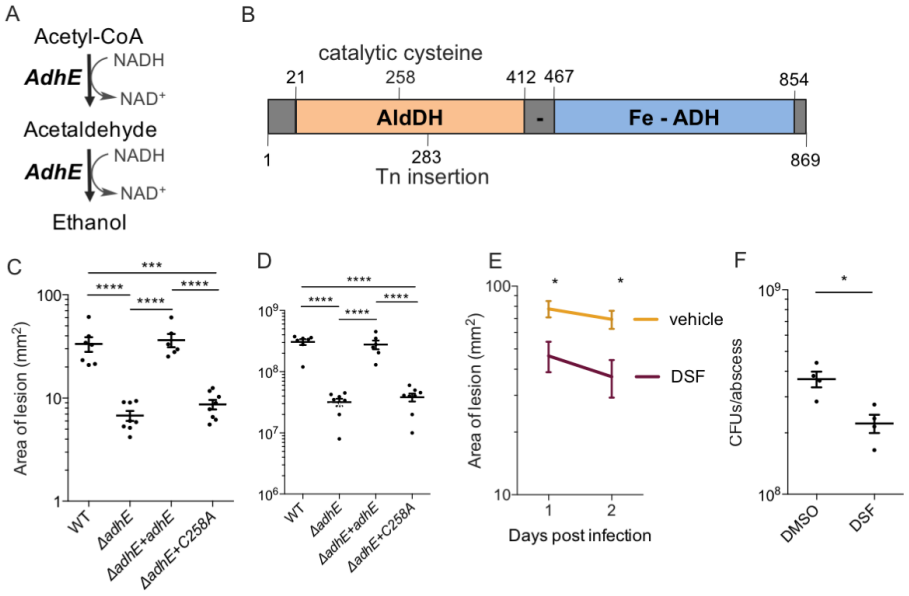


Figure 3.5: The metabolic role of AdhE during infection. A) Fermentative pathway AdhE is annotated as important during low oxygen environments. B) Cartoon representation of the bifunctional enzyme. C) Area of lesion formed during complementation with the C258A mutant compared to wild-type (WT) after two days of infection. D) CFUs / abscess obtained in experiment 3.5D. E) Area of lesion formed after DSF treatment. F) CFUs / abscess obtained in experiment 3.5E.

3.4 Discussion

Elucidation of the proteome of a pathogen during an *in vivo* infection is difficult; *in vitro* cultures that mimic the host environment are often used instead. However, these *in vitro* cultures do not sufficiently reflect the conditions within hosts and important factors may be missed. By using the cell-specific BONCAT method, we were able to label, visualize, and enrich proteins involved in pathogenesis within a mouse model of MRSA skin infection, overcoming some of the prior challenges associated with discovery-based proteomics of pathogens in a mammalian host. In addition to validating many known virulence factors, the proteomics results provide an unbiased snapshot of MRSA as it

infects.

This portrait of infecting MRSA reveals a pathogen hiding from the innate immune system while scavenging for nutrients by expressing proteins important for host cell lysis, antibody evasion, iron acquisition, and amino acid catabolism (Fig 3.3). While many of these proteins have previously been identified to be important for infection, we also identified some that had not been previously associated with contributing to infection. We postulate that many of these could be candidate anti-infective targets, but did not have the capabilities to screen every protein found.

By screening ten proteins significantly upregulated or uniquely found within the mouse compared to *in vitro* growth, we hoped to find proteins that are important for infecting MRSA, but are not indispensable for normal growth and viability. Targeted inhibitors of these proteins may impose less selection pressure for the microbe and suppress the development of antibiotic resistance. We were fortunate to find a protein, AdhE, which fit these characteristics. Disrupting AdhE does not cause defects to growth *in vitro*, which may have led to it not previously being identified as important for MRSA infection (Fig S3.11). A single point mutation to the catalytic cysteine within AdhE significantly decreased the levels of infection, which is surprising as the MRSA genome contains many redundancies (5). For example, we also found a related alcohol dehydrogenase (Adh) highly upregulated in the skin infection proteome, but when we tested an *Δadh* mutant *in vivo*, we saw no decrease in its ability to infect (Fig 3.4A-C).

Understanding the molecular mechanism of AdhE's contributions to infection and fermentation the abscess will require further study. The redox sensor Rex regulates expression of AdhE, suggesting that during adaptation to early infection, *S. aureus* expresses proteins important in maintaining the redox balance of $NADH/NAD^+$ (42). Previous studies have indicated that the pathogenic nature of MRSA depends on its ability to ferment within the low oxygen environment of the abscess (23, 43, 44). While L-lactate fermentation is critical for renal infection, especially during nitrosative stress (45), fermenting MRSA have also been shown to produce ethanol and formate (41). Supporting this theory, a study of a *S. aureus* infection in a human prosthetic joint detected the presence of both lactate and ethanol in the joint fluid (46). While disulfiram significantly minimized the infection, the drug also inhibits the mammalian ALDH1A1. We postulate that more specific therapies related to disulfiram's structure could be created to target the C258 residue in *S. aureus* in particular with higher specificity and affinity. Overall, targeting AdhE and pathogenic fermentation in general has therapeutic potential as it augments the innate immune system and decreases infection severity.

Finally, we hypothesize that application of this approach to other pathogens and animal models of infection would be straightforward and could yield important insights into both the dynamic environment that pathogens see during infection and their subsequent response to these fluctuations. We envision that by placing the NLL-MetRS under different promoters cell- and state-selective BONCAT may be achieved to label

additional subpopulations within infections.

3.5 Materials & Methods

Ethics statement. Animal experiments were performed in accordance with the regulations for the Institutional Animal Care and Use Committee at Caltech. The investigators were not blinded to the experimental conditions.

Culture conditions. Bacteria were grown at 37 °C in LB (*E. coli*), or tryptic soy (*S. aureus*) medium or on agar plates supplemented with 100 μg/mL ampicillin, 10 μg/mL erythromycin, 50 μg/mL spectinomycin, 10 μg/mL chloramphenicol (RN4220), 20 μg/mL chloramphenicol (USA300), or 5 μg/mL tetracycline as needed. Azidonorleucine hydrochloride (Iris Biotech) was dissolved in PBS and brought to a pH of 7.5 using 1 M NaOH for a final concentration of 100 mM. This stock solution was diluted in PBS for all subsequent labeling studies.

Bacterial strains and plasmids. A complete list of the strains and primers used can be found in Supplementary Tables 1 and 2. *S. aureus* mutant strains were generated from the *S. aureus* clinical isolate JE2 (USA300). Wild-type strains and all isogenic mutants were provided by the Network on Antimicrobial Resistance in *Staphylococcus aureus* (NARSA) for distribution by BEI Resources, NIAID, NIH (47). Transposon insertions were verified through sequencing prior to use (47).

The *E. coli* mutant NLL methionyl-tRNA synthetase was codon-optimized for *S. aureus* (Genscript) and inserted into the backbone of pmRFPmars (a gift from Martin Fraunholz (18), Addgene plasmid # 26252) using ligation-independent cloning to remove the mRFP and insert an inducible NLL-MetRS (pSS20_tet). The tetR promoter was removed and the hprK promoter amplified from the USA300 genome was inserted in front of the NLL-MetRS using the SalI and KpnI sites to make the plasmids pSS20_hprK. These plasmids were confirmed through sequencing (Table S3.3). This plasmid was electroporated into the cloning strain of *S. aureus* RN4220 and the isolated plasmid from this strain was electroporated into strain JE2 (USA300).

Chromosomal integration of AdhE was achieved using the system created by Lei et al (31). We amplified AdhE and its endogenous promoter from the USA300 chromosome and inserted this fragment into the multiple cloning site of plasmid pLL102 to create pLL102_adhE. We inserted this plasmid by electroporation into a strain of RN4220 that has the engineered integration site and expresses the corresponding integrase (strain CYL12349). This chromosomal insertion was then transduced using Φ85 phage into the JE2 ΔAdhE mutant strain. Using the same method, the empty pLL102 plasmid was also incorporated into wild-type JE2 and the ΔAdhE strain for controls.

BONCAT labeling in TSB. Overnight cultures of *S. aureus* (+NLL or –NLL) were diluted 100-fold into 2 mL of fresh TSB with chloramphenicol. When the cultures reached an OD of ~0.3, Anl was added to a final concentration of 2 mM for 1 hour. The samples

were spun down, resuspended in 200 µl TSM (50 mM Tris, 500 mM sucrose, 10 mM $MgCl_2$), and lysed with 1 µg of lysostaphin at 37 °C for 1 hour. The samples were brought to 1% SDS and the manufacturer's instructions were followed for the Click-It Kit with alkyne-TAMRA (Thermo Scientific).

Mouse infection and *in vivo* labeling studies. Skin infections were performed according to previously described methods (4). Briefly, overnight cultures of *S. aureus* were diluted 100-fold into 5 mL of fresh TSB. The culture was grown at 37 °C for 2-2.5 hours with shaking until the OD_{600} = 0.48. Cultures were washed with PBS and resuspended to achieve a concentration of ~$2x10^7$ cells/ 50 µL. For each experiment, a dilution series of every strain to be injected was plated in duplicate for an estimate of cell count injected into each animal.

8-week-old female C57BL/6J mice (Jackson Labs, #000664) housed under specific-pathogen free (SPF) conditions were shaved and treated with Nair on their flanks at least a day prior to infection. On the day of infection, mice were anesthetized with isoflurane and inoculated by subcutaneous (S/C) injection in each flank with ~$2x10^7$ CFU in 50 µL D-PBS. For Anl labeling experiments, 25 µL of 50 mM Anl was injected along with *S. aureus*. Each day the lesions on the mice were measured by calipers and photographed for image processing using ImageJ. At the indicated timepoint, mice were killed by CO_2 asphyxiation, and abscesses were excised and homogenized in 1 mL of 1% Triton X-100 in PBS. Dilution series of the homogenates were plated on agar for enumeration of viable

MRSA. Statistical analysis was performed with ANOVA corrected for multiple hypothesis testing with Dunnett's method using Prism (GraphPad Software, La Jolla, CA).

Cryosectioning, DBCO-staining, immunostaining, and imaging. Abscesses were excised, placed in optimum cutting temperature (O.C.T.) compound, and frozen at -80 °C. They were then cut into ~15 μm sections and deposited onto glass slides. The slides were fixed with 2% paraformaldehyde (PFA), permeabilized with 0.1% Triton-X, and treated with lysostaphin at 37 °C. The samples were next blocked in 100 mM iodoacetamide for 30 min in the dark, then reacted with 5 μM DBCO-AlexaFluor488 (Click Chemistry Tools) for 15 min. After washing, the samples were blocked with 1% mouse serum, then treated with a 1:3000 dilution of anti-*Staphylococcus aureus* antibody (polyclonal, rabbit) (Ab37644, Abcam). A goat anti-rabbit secondary antibody conjugated to AlexaFluor555 (Thermo Scientific) was then added at 1:10,000 dilution. Samples were washed with PBS, and Vectashield with DAPI was added prior to coverslips. Imaging was performed using a Zeiss LSM 880 confocal microscope with a Plan-Apochromat 10/0.45-numerical aperture M27 objective (working distance [wd], 2.0 mm). Image reconstruction and analysis was performed in the FIJI distribution of ImageJ (48).

Enrichment of azide-labeled proteins.

Proteins were enriched in parallel from +NLL or –NLL infected samples to account for nonspecific binding to beads. Homogenized abscesses in PBS were brought to 1.5% SDS and boiled. Insoluble proteins were removed and the clarified lysate was reduced for 10 min in 10 mM dithiothreitol (DTT), followed by incubation with 100 mM iodoacetamide

for 45 min in the dark to block any reactive thiols. One volume equivalent of 8 M urea was added, followed by 30 µL of washed DBCO-agarose, and the samples were rotated end over end in the dark for 16-24 hours. Unbound proteins were washed from the agarose resin in gravity columns using 1% SDS (10 x 5 mL), 8 M urea (10 x 5 mL), and 20% acetonitrile (10 x 5 mL). An on-bead digest was performed by adding 100 ng trypsin in 10% acetonitrile (ACN) in 100 mM ammonium bicarbonate buffer. Released peptides were dried and desalted using C18 StageTips as previously described (49).

LC-MS/MS analysis. All LC-MS/MS experiments were performed on an EASY-nLC1000 coupled to a hybrid LTQ-Orbitrap Elite mass spectrometer with a nano-electrospray ion source (Thermo Fisher Scientific). Peptides were re-suspended in 0.1% (v/v) formic acid and loaded on a 15-cm reversed phase analytical column (75 µm internal performance diameter, ID) packed in-house with 3 µm C18AQ beads (ReproSil-Pur C18AQ, Dr. Maisch-GmbH). For LC, solvent A consisted of 97.8% H_2O, 2% ACN, and 0.2% formic acid and solvent B consisted of 19.8% H_2O, 80% ACN, and 0.2% formic acid. The samples were run on a 120 minute elution gradient from 0% (100% solvent A) to 30% Solvent B (70% solvent A) at a flow rate of 350 nL/minute. The mass spectrometer was operated in data-dependent mode to switch automatically between MS and MS/MS scans (50). Survey full scan mass spectra were acquired in the Orbitrap (400–1600 m/z) with a resolution of 60,000 at 400 m/z. The top twenty most intense ions from the survey scan were isolated and fragmented in the linear ion trap by collisionally induced dissociation (CID). Precursor ion charge state screening was enabled, and all

singly charged and unassigned charge states were rejected. The dynamic exclusion list was set with a maximum retention time of 90 seconds and a relative mass window of 10 ppm.

MS data analysis. Raw files were searched using MaxQuant(51) against the *S. aureus* strain USA300 UniProt entry (2,607 sequences), *Mus musculus* C57BL/6J UniProt entry (51,544 sequences), and an in-house contaminant database (259 sequences). The digestion enzyme was specified as trypsin with up to three missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification, and protein N-terminal acetylation, N-terminal formylation, and methionine oxidation were variable modifications. We also included variable modifications of methionine corresponding to Anl or reduced Anl.

Statistical data analysis. Skin infection data was analyzed and plotted using Prism (GraphPad). All data are expressed as mean ± SEM. Differential protein expression analysis was performed with the Bioconductor package LIMMA 3.14.1. Proteins with a corresponding fold-change P value (adjusted for multiple hypothesis testing with the Benjamini-Hochberg method) lower than 0.05 were accepted as differentially expressed.

Disulfiram testing. Mice were gavaged with 160 mg/kg disulfiram (Pharmaceutical grade, Sigma) or vehicle the day before and the day of infection.

Supporting Information. For details regarding experimental methods and supplementary figures, see Supporting Materials and Methods (Appendix A).

References

1. Thurlow LR, Joshi GS, & Richardson AR (2012) Virulence strategies of the dominant USA300 lineage of community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA). *Fems Immunol Med Mic* 65(1):5-22.
2. Thomer L, Schneewind O, & Missiakas D (2016) Pathogenesis of *Staphylococcus aureus* Bloodstream Infections. *Annu Rev Pathol-Mech* 11:343-364.
3. Wang Y, *et al*. (2017) Mouse model of hematogenous implant-related *Staphylococcus aureus* biofilm infection reveals therapeutic targets. *Proc Natl Acad Sci* 114(26):E5094-E5102.
4. Kennedy AD, *et al*. (2010) Targeting of alpha-hemolysin by active or passive immunization decreases severity of USA300 skin infection in a mouse model. *J Infect Dis* 202(7):1050-1058.
5. Diep BA, *et al*. (2014) Identifying Potential Therapeutic Targets of Methicillin-resistant *Staphylococcus aureus* Through *in Vivo* Proteomic Analysis. *J Infect Dis* 209(10):1533-1541.
6. Becker D, *et al*. (2006) Robust *Salmonella* metabolism limits possibilities for new antimicrobials. *Nature* 440(7082):303-307.
7. Otto A, van Dijl JM, Hecker M, & Becher D (2014) The *Staphylococcus aureus* proteome. *Int J Med Microbiol* 304(2):110-120.
8. Olive AJ & Sassetti CM (2016) Metabolic crosstalk between host and pathogen: sensing, adapting and competing. *Nat Rev Microbiol* 14(4):221-234.
9. Stone SE, Glenn WS, Hamblin GD, & Tirrell DA (2017) Cell-selective proteomics for biological discovery. *Curr Opin Chem Biol* 36:50-57.
10. Yuet KP, *et al*. (2015) Cell-specific proteomic analysis in *Caenorhabditis elegans*. *Proc Natl Acad Sci* 112(9):2705-2710.
11. Mahdavi A, *et al*. (2014) Identification of secreted bacterial proteins by noncanonical amino acid tagging. *Proc Natl Acad Sci* 111(1):433-438.
12. Ngo JT, Schuman EM, & Tirrell DA (2013) Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells. *Proc Natl Acad Sci* 110(13):4992-4997.

13. Grammel M, Dossa PD, Taylor-Salmon E, & Hang HC (2012) Cell-selective labeling of bacterial proteomes with an orthogonal phenylalanine amino acid reporter. *Chem Commun* 48(10):1473-1474.
14. Wier GM, McGreevy EM, Brown MJ, & Boyle JP (2015) New Method for the Orthogonal Labeling and Purification of *Toxoplasma gondii* Proteins While Inside the Host Cell. *Mbio* 6(2):e01628-01614.
15. Chande AG, *et al*. (2015) Selective enrichment of mycobacterial proteins from infected host macrophages. *Sci Rep-Uk* 5:13430.
16. Ngo JT, *et al*. (2009) Cell-selective metabolic labeling of proteins. *Nat Chem Biol* 5(10):715-717.
17. Bateman BT, Donegan NP, Jarry TM, Palma M, & Cheung AL (2001) Evaluation of a tetracycline-inducible promoter in *Staphylococcus aureus in vitro* and *in vivo* and its application in demonstrating the role of sigB in microcolony formation. *Infect Immun* 69(12):7851-7857.
18. Paprotka K, Giese B, & Fraunholz MJ (2010) Codon-improved fluorescent proteins in investigation of *Staphylococcus aureus* host pathogen interactions. *J Microbiol Methods* 83(1):82-86.
19. Bubeck Wardenburg J, Williams WA, & Missiakas D (2006) Host defenses against *Staphylococcus aureus* infection require recognition of bacterial lipoproteins. *Proc Natl Acad Sci* 103(37):13831-13836.
20. Kammers K, Cole RN, Tiengwe C, & Ruczinski I (2015) Detecting Significant Changes in Protein Abundance. *EuPA Open Proteom* 7:11-19.
21. Ritchie ME, *et al*. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7): e47.
22. Carneiro CRW, Postol E, Nomizo R, Reis LFL, & Brentani RR (2004) Identification of enolase as a laminin-binding protein on the surface of *Staphylococcus aureus*. *Microbes Infect* 6(6):604-608.
23. Halsey CR, *et al*. (2017) Amino Acid Catabolism in *Staphylococcus aureus* and the Function of Carbon Catabolite Repression. *Mbio* 8(1): e01434-16.
24. Yajjala VK, *et al*. (2016) Resistance to Acute Macrophage Killing Promotes Airway Fitness of Prevalent Community-Acquired *Staphylococcus aureus* Strains. *J Immunol* 196(10):4196-4203.
25. Mazmanian SK, *et al*. (2003) Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* 299(5608):906-909.
26. Thurlow LR, *et al*. (2013) Functional Modularity of the Arginine Catabolic Mobile Element Contributes to the Success of USA300 Methicillin-Resistant *Staphylococcus aureus*. *Cell Host Microbe* 13(1):100-107.
27. Bae T, *et al*. (2004) *Staphylocloccus aureus* virulence genes identified by bursa aurealis mutagenesis and nematode killing. *Proc Natl Acad Sci* 101(33):12312-12317.
28. Chatterjee SS, *et al*. (2013) Essential *Staphylococcus aureus* toxin export system. *Nat Med* 19(3):364-367.
29. Cheng AG, *et al*. (2009) Genetic requirements for *Staphylococcus aureus* abscess formation and persistence in host tissues. *Faseb J* 23(10):3393-3404.

30. Cascioferro S, Totsika M, & Schillaci D (2014) Sortase A: An ideal target for anti-virulence drug development. *Microb Pathogenesis* 77:105-112.
31. Lei MG, *et al*. (2012) A single copy integration vector that integrates at an engineered site on the *Staphylococcus aureus* chromosome. *BMC Res Notes* 5:5.
32. Fuchs S, Pane-Farre J, Kohler C, Hecker M, & Engelmann S (2007) Anaerobic gene expression in *Staphylococcus aureus*. *J Bacteriol* 189(11):4275-4289.
33. Beckham KSH, *et al*. (2014) The metabolic enzyme AdhE controls the virulence of *Escherichia coli* O157:H7. *Mol Microbiol* 93(1):199-211.
34. Echave P, Tamarit J, Cabiscol E, & Ros J (2003) Novel antioxidant role of alcohol dehydrogenase E from *Escherichia coli*. *J Biol Chem* 278(32):30193-30198.
35. Burkholder KM & Bhunia AK (2010) *Listeria monocytogenes* Uses Listeria Adhesion Protein (LAP) To Promote Bacterial Transepithelial Translocation and Induces Expression of LAP Receptor Hsp60. *Infect Immun* 78(12):5062-5073.
36. Luong TT, *et al*. (2015) Ethanol-Induced Alcohol Dehydrogenase E (AdhE) Potentiates Pneumolysin in *Streptococcus pneumoniae*. *Infect Immun* 83(1):108-119.
37. Lorenz N, Clow F, Radcliff FJ, & Fraser JD (2013) Full functional activity of SSL7 requires binding of both complement C5 and IgA. *Immunol Cell Biol* 91(7):469-476.
38. Espinosa A, *et al*. (2001) The bifunctional *Entamoeba histolytica* alcohol dehydrogenase 2 (EhADH2) protein is necessary for amebic growth and survival and requires an intact C-terminal domain for both alcohol dehydrogenase and acetaldehyde dehydrogenase activity. *J Biol Chem* 276(23):20136-20143.
39. Zheng TY, *et al*. (2015) Cofactor Specificity of the Bifunctional Alcohol and Aldehyde Dehydrogenase (AdhE) in Wild-Type and Mutant *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*. *J Bacteriol* 197(15):2610-2619.
40. Koppaka V, *et al*. (2012) Aldehyde Dehydrogenase Inhibitors: a Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application. *Pharmacol Rev* 64(3):520-539.
41. Ejim L, *et al*. (2011) Combinations of antibiotics and nonantibiotic drugs enhance antimicrobial efficacy. *Nat Chem Biol* 7(6):348-350.
42. Pagels M, *et al*. (2010) Redox sensing by a Rex-family repressor is involved in the regulation of anaerobic gene expression in *Staphylococcus aureus*. *Mol Microbiol* 76(5):1142-1161.
43. Park MK, Myers RAM, & Marzella L (1992) Oxygen-Tensions and Infections - Modulation of Microbial-Growth, Activity of Antimicrobial Agents, and Immunological Responses. *Clin Infect Dis* 14(3):720-740.
44. Vitko NP, Spahich NA, & Richardson AR (2015) Glycolytic Dependency of High-Level Nitric Oxide Resistance and Virulence in *Staphylococcus aureus*. *Mbio* 6(2): e00045-15.

45. Spahich NA, Vitko NP, Thurlow LR, Temple B, & Richardson AR (2016) *Staphylococcus aureus* lactate- and malate-quinone oxidoreductases contribute to nitric oxide resistance and virulence. *Mol Microbiol* 100(5):759-773.
46. Xu YJ, *et al*. (2016) *In vivo* gene expression in a *Staphylococcus aureus* prosthetic joint infection characterized by RNA sequencing and metabolomics: a pilot study. *BMC Microbiol* 16:80.
47. Fey PD, *et al*. (2013) A Genetic Resource for Rapid and Comprehensive Phenotype Screening of Nonessential *Staphylococcus aureus* Genes. *Mbio* 4(1).
48. Schindelin J, *et al*. (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9(7):676-682.
49. Rappsilber J, Mann M, & Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2(8):1896-1906.
50. Kalli A & Hess S (2012) Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer. *Proteomics* 12(1):21-31.
51. Cox J & Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26(12):1367-1372.

Chapter 4

VISUALIZING PATHOGENIC PROTEIN SYNTHESIS DURING INFECTION

4.1 Abstract

We extend cell-selective bioorthogonal noncanonical amino acid tagging (BONCAT) to visualize staphylococcal protein synthesis in three dimensions within skin abscesses. We use BONCAT methodology combined with the hybridization chain reaction (HCR) to visualize both proteins and rRNA within cleared abscesses. We hypothesize that this methodology can be readily applied to diverse microbial systems to study the biogeography of host-microbe interactions.

This work was a collaboration with Will H. DePas, Bryan Yoo, Sarkis K. Mazmanian, and Dianne K. Newman.

4.2 Introduction

The biogeography, or spatial arrangement of microbes within a host, depends on the features of their microenvironment such as nutrient availability, oxygen levels, and host-cell interactions. Characterizing these spatial arrangements is important to understanding and controlling microbes as they interact with the host (1, 2). While thin sectioning of abscesses has led to remarkable insights into host-microbe interactions, tools to study the three-dimensional biogeography of infection are currently limited. Using tissue-clearing techniques to visualize infection may produce insights previously unobtainable by thin slicing alone.

Recent advances in tissue clearing techniques such as the passive CLARITY technique (PACT) have been used to render tissues transparent and allow for visualization of fine structure in three-dimensions (3). Importantly, PACT preserves the spatial relationships of molecules within the cell by encasing them within an acrylamide-based hydrogel, and is compatible with most immunohistochemistry or *in situ* hybridization techniques (4-6). While most researchers using PACT have focused on neuroscience applications due to the method's ability to preserve neural structure within the brain, the Tobin lab has used PACT to visualize granulomas of *Mycobacterium tuberculosis* expressing a fluorescent protein in clarified murine lung tissue and whole zebrafish (7, 8).

Recently, the Newman lab developed the microbial identification after passive CLARITY technique (MiPACT) to study polymicrobial infections within sputum samples from

patients with cystic fibrosis (CF) (9). This method uses hybridization chain reaction (HCR) to identify and visualize microbial species by labeling their ribosomal RNA (rRNA). HCR uses short DNA probes complementary to a target RNA to trigger the polymerization of fluorophore-labeled DNA hairpins in a hybridization chain reaction, effectively amplifying signal from *in situ* hybridization (10, 11). By using HCR and PACT, they were able to both quantify growth rates of CF pathogens and visualize substructures of multi-community biofilms in sputum.

We hypothesized that cell-selective bioorthogonal noncanonical amino acid tagging (BONCAT) could be used with tissue-clearing to explore both the protein localization and substructures formed by *Staphylococcus aureus* during skin infection. In this chapter, we apply MiPACT to staphylococcal abscesses that have been cell-selectively labeled using BONCAT to visualize both pathogenic proteins and rRNA during infection.

4.3 Results

Staphylococcal abscesses can be clarified using PACT. As detailed in Chapter 3, we first labeled proteins synthesized by methicillin-resistant *Staphylococcus aureus* (MRSA) within murine skin abscesses using azidonorleucine (Anl) (Fig 4.1A). After excision and fixation of the abscess, we tested various ratios of bisacrylamide, acrylamide, and paraformaldehyde (PFA) for embedding conditions to encase the skin abscess within a stabilizing hydrogel. The conditions we settled on were 4% of a 29:1 acrylamide:bis-acrylamide (v/v) mixture, 1% paraformaldehyde (PFA), and 0.25% of the thermal-initiator VA-044 in 1X PBS, which we called “B4P1”.

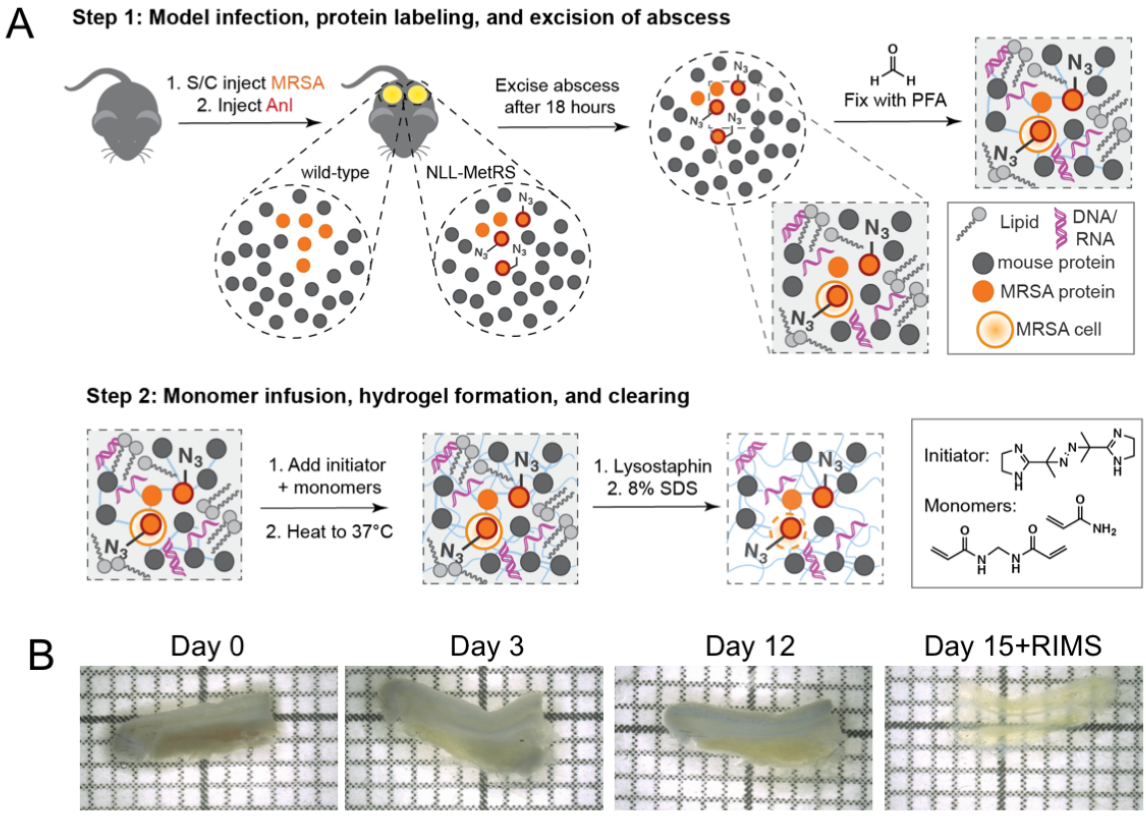


Figure 4.1: A) Scheme used to label skin abscess during infection, embed within hydrogel, and remove lipids. B) Pictures of the abscess during incubation in SDS and after submersion in refractive index matching solution (RIMS). Each square represents 1 mm^2.

Other conditions tested, with smaller quantities of bisacrylamide or PFA were less robust to the subsequent protocols: the inner staphylococcal abscesses would often separate from the epidermal layer. After polymerization, we cut the abscess into ~1 mm thick slices and removed the lipids from the tissue using the passive clarity technique (PACT). We took pictures of these skin abscess pieces over 15 days to monitor clearing (Fig 4.1B). We did not notice additional clearing after day 12, and on day 15 we resuspended the tissue in refractive index matching solution (RIMS) for 24 hours. The skin abscess appeared

significantly more transparent (Fig 4.1B). Collagenase treatment prior to clearing steps did not show enhanced transparency.

BONCAT/HCR labeling is compatible with PACT. To determine optimal click chemistry and HCR conditions to use on the skin abscesses, we first tested several conditions on B4P1 hydrogel blocks of *S. aureus* +MetRS or –MetRS strains grown in tryptic soy broth (TSB) and labeled with azidonorleucine (Anl) for 1 hour during mid-exponential growth. We tested the copper-catalyzed alkyne-azide cycloaddition (CuAAC) with a terminal alkyne-functionalized tetramethylrhodamine (TAMRA) or the strain-promoted alkyne-azide cycloaddition (SPAAC) using aza-dibenzocyclooctyne functionalized with TAMRA, before and after processing the samples with HCR probes. We found that the only conditions that showed signal from both the click reaction and HCR was when we performed the click reaction first using SPAAC, followed by a washing step, then incubation with HCR probes (Fig S4.1, Fig 4.2A).

We then took the azide-tagged and clarified skin abscess pieces, performed SPAAC with DBCO-TAMRA, then performed HCR with 30 nM staphylococcal-specific initiator probe (STA3) and corresponding amplification hairpins tagged with AlexaFluor488 (120 nM, B4-AF488). After washing unbound probes away followed by 24 hours of incubation in RIMS, we visualized abscesses and saw signal for both ribosomal RNA and newly-synthesized proteins within the abscess (Fig 4.2C). Probes with a "mismatch" sequence (no complementarity to target RNA) showed little background signal in the AF488

channel (Fig S4.2A) and samples infected with a –MetRS strain showed no signal in the TAMRA channel (Fig 4.2B), suggesting that the signal was specific. First attempts to stain skin abscesses showed a high degree of background labeling in the thick epidermal layer, but this signal was decreased by adding 1 µM of a random DNA primer to the hybridization step, effectively blocking charged regions that nonspecifically bind the DNA probes within the tissue.

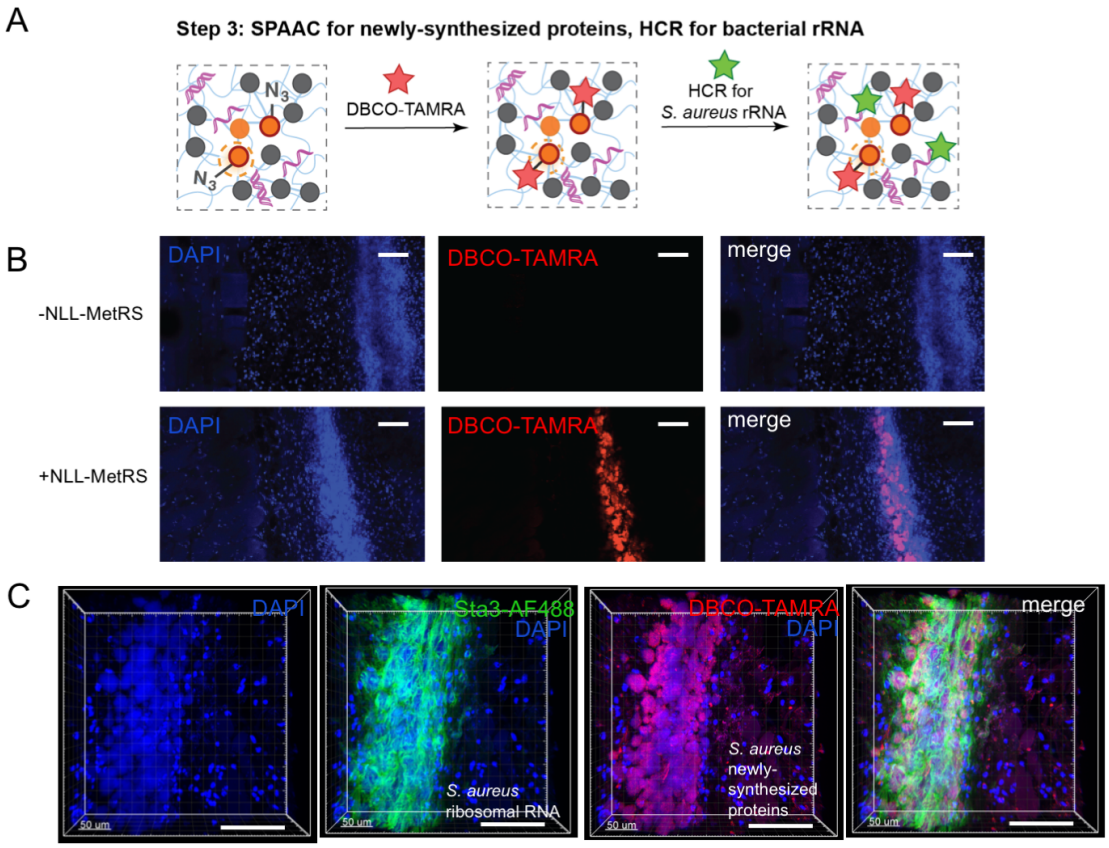


Figure 4.2: A) Scheme used to conjugate DBCO-fluorophore to newly synthesized proteins, followed by HCR for *S. aureus* ribosomal RNA (rRNA). B) Fluorescence microscopy of skin abscess pieces (volume ~30 mm^3) after following steps described in Fig 4.1A and 4.2A. Scale bars are 100 µm. C) Staining for azide-tagged proteins (red) overlaid in three-dimensional space with HCR for staphylococcal rRNA.

Notably, the colocalization of the HCR signal and the BONCAT signal was inconsistent: though the two signals did appear in the same general area of the abscess, their spatial arrangement and patterns were markedly different. This implied that the newly synthesized azide-tagged staphylococcal proteins were not in the same location as the rRNA used to synthesize them. In several experiments the rRNA probes appeared to surround the staphylococcal proteins in a fibrous “halo”, while in others the HCR signal appeared offset from the bulk of the MRSA abscess as visualized through BONCAT and DAPI. (Fig 4.2C). This result was puzzling, as we had expected the HCR signal to appear as puncta colocalized with protein visualization. We repeated the experiment but swapped the fluorophore used so that SPAAC was performed using DBCO-AF488 and the HCR amplification hairpins were conjugated to AlexaFluor647 (AF647). We again saw the HCR signal for staphylococcal rRNA slightly offset from the newly-synthesized staphylococcal proteins visualized with BONCAT (Fig 4.3).

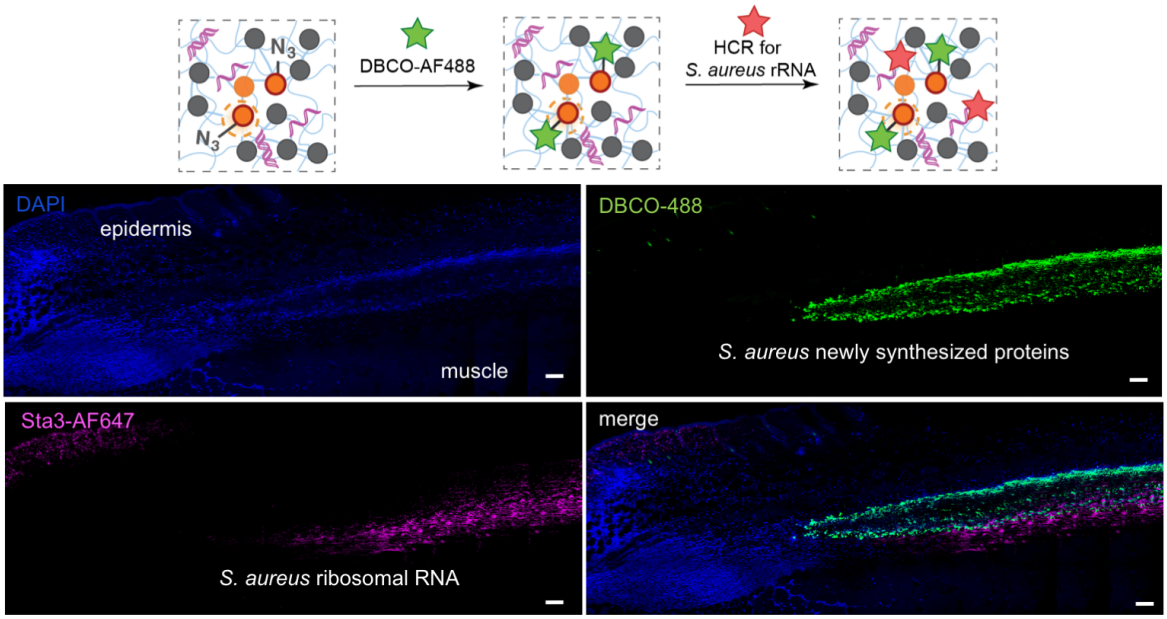


Figure 4.3: Wider field view of skin abscess following a similar protocol to Fig 4.2 except with different fluorophores. Scale bars are 100 µm.

Comparison of BONCAT/HCR labeling using PACT with thin sections. We aimed to compare the biogeography observed with PACT to that observed with traditional thin sectioning techniques. We cryosectioned fixed Anl-labeled staphylococcal abscesses to ~15 µm thick before applying BONCAT and HCR probes. Similar to the PACT results, we observed that the HCR signal colocalized only partially with the BONCAT signal. Furthermore, in the cryosectioned samples there was a strong BONCAT signal showing newly-synthesized staphylococcal proteins in a deeper layer of the skin abscess that we had previously not observed in the PACT samples (Fig 4.4). This BONCAT signal colocalized perfectly with polyclonal anti-*S. aureus* immunostaining, confirming the identity of staphylococcal proteins. We also compared these microscopy images to hematoxylin and eosin-stained (H&E) paraffin slices (~10 µm), in which the abscess

demonstrates a dark purple color due to the large amount of polymorphonuclear leukocytes (neutrophils, or PMNs) (Fig 4.4).

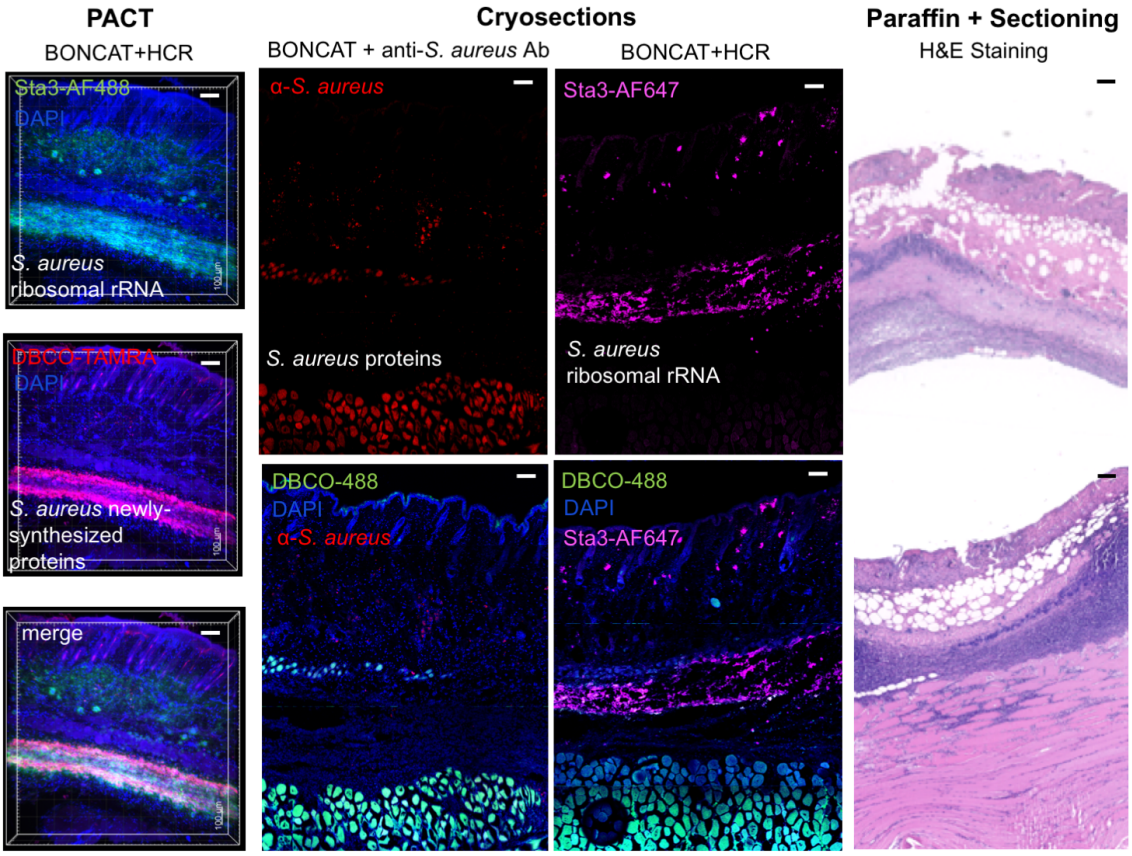


Figure 4.4: Comparisons of various techniques to visualize staphylococcal skin abscesses. Scale bars are 100 μm. In the samples cleared by PACT (~3x3x3 mm in volume), the HCR signal surrounded the BONCAT signal. In the cryosections, a similar pattern was seen to the PACT samples in the top portion of the skin (towards the epidermis), but an additional strong BONCAT signal (DBCO-488) was seen in the lower layers of the skin (towards the muscle). Additionally, the BONCAT signal for newly-synthesized proteins overlays with an anti-*S. aureus* polyclonal antibody. After paraffin embedding, thin-sectioning, and H & E staining, the neutrophils (dark purple) can be seen in several layers of the abscess within the skin.

4.4 Discussion and Future Directions

The colocalization of the anti- *S. aureus* immunostaining with signal from BONCAT suggests we have successfully tagged newly-synthesized staphylococcal proteins within the skin. HCR results are less clear as they do not exactly colocalize with signal from *S. aureus* proteins within this skin abscess infection model.

We have several testable hypotheses as to why part of the abscess lacked HCR signal: 1) the rRNA in this area is being degraded prior to detection, 2) the BONCAT and immunostaining is visualizing lysed *S. aureus* within host phagocytes, 3) the BONCAT and immunostaining is visualizing secreted proteins from *S. aureus*. During skin infection, *S. aureus* secretes numerous toxins to lyse host cells, sometimes while within phagocytes (12, 13), and BONCAT has previously been shown to tag secreted proteins (14). Chatterjee *et al*. showed that ~60% by mass of the proteins secreted by a S. aureus *in vitro* culture was a class of toxins known as phenol soluble modulins (PSMs) (15). Further immunohistochemistry experiments using antibodies against these toxins could confirm the identity of potential proteins secreted into the abscess.

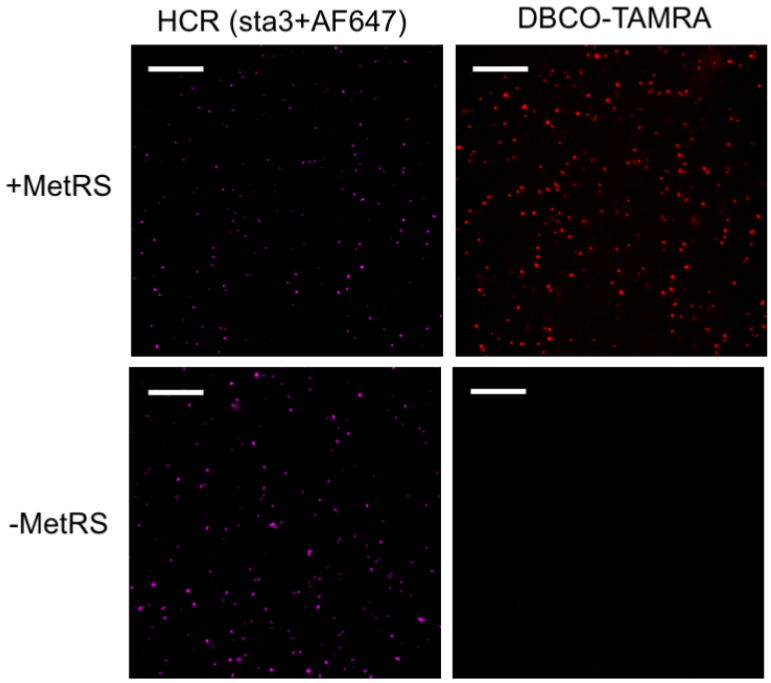
Additional RNA preservation steps (RNase inhibitors, for example) may be required for these hydrogel-embedded tissues. However, because the HCR signal was similar in both cleared tissues and cryoslices, we reasoned that penetration of the DNA probes in the abscess is not the issue, even within the dense hydrogel network. The strong BONCAT signal colocalized with anti-*S. aureus* immunostaining in the area not labeled by HCR was only seen in the cryoslices, and not in the clarified skin tissue. Future experiments could

elucidate whether this is because the hydrogel-embedded tissue is losing this part of the abscess, or if the clearing step is removing some newly-synthesized tagged proteins. Increased amounts of bisacrylamide and PFA may lend increased stiffness to the skin layers to preserve the intact abscess, but decreasing the hydrogel porosity may reduce probe penetration and increase artifacts such as swelling.

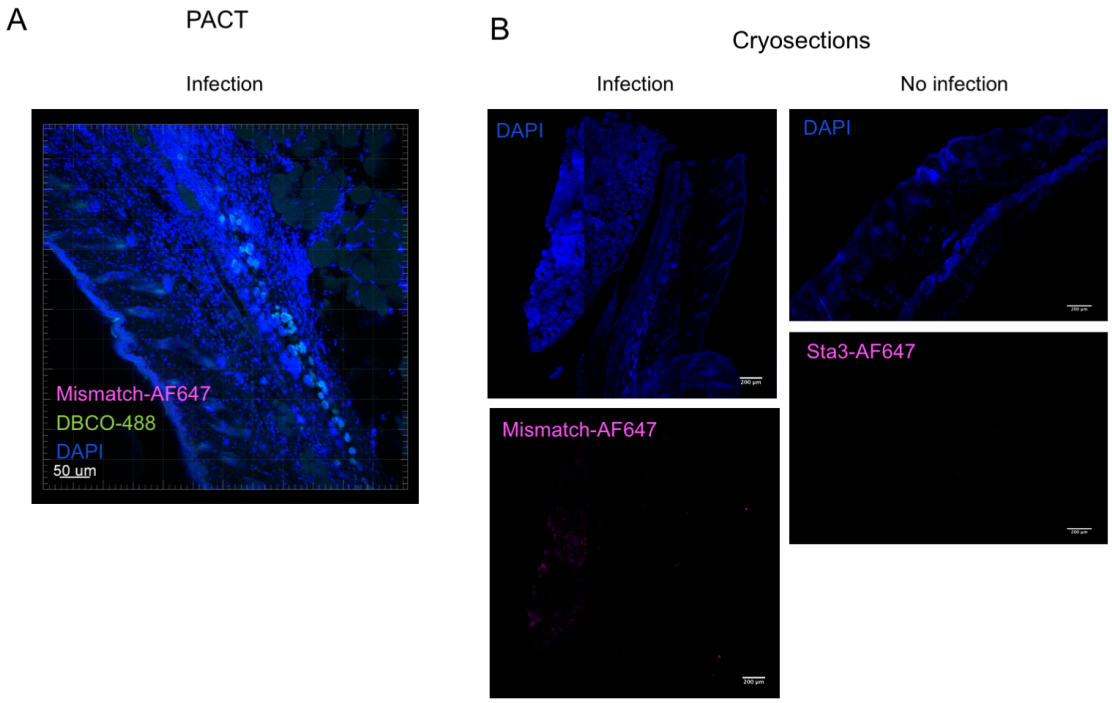
While tissue-clearing processes such as PACT have not reported on the loss of large amounts of proteins before, it is known that fluorescent lipophilic molecules like DyeI (MW: ~1 kDa) are incompatible with these techniques as they associate with the lipids removed from cell membranes. Several secreted staphylococcal toxins such as PSMs and the related delta-toxin are small (2-3 kDa), detergent-like cytolytic proteins that associate with host cell membranes and form transient pores within the membrane (16). It is possible that the tissue clearing detergents used in PACT remove this class of proteins, potentially revealing a limitation to this method in the visualization of pathogenic proteins in host tissue. We could test this hypothesis by assaying the PACT solution post clearing for proteins using Western blotting or LC-MS/MS (4).

We have shown the chemistries used for conjugation of azide-tagged molecules are compatible with the HCR and PACT techniques pioneered by MiPACT. We envision cell-selective labeling using BONCAT could be combined with tissue-clearing techniques to not only visualize host-pathogen interactions, but neural proteomic mapping as well.

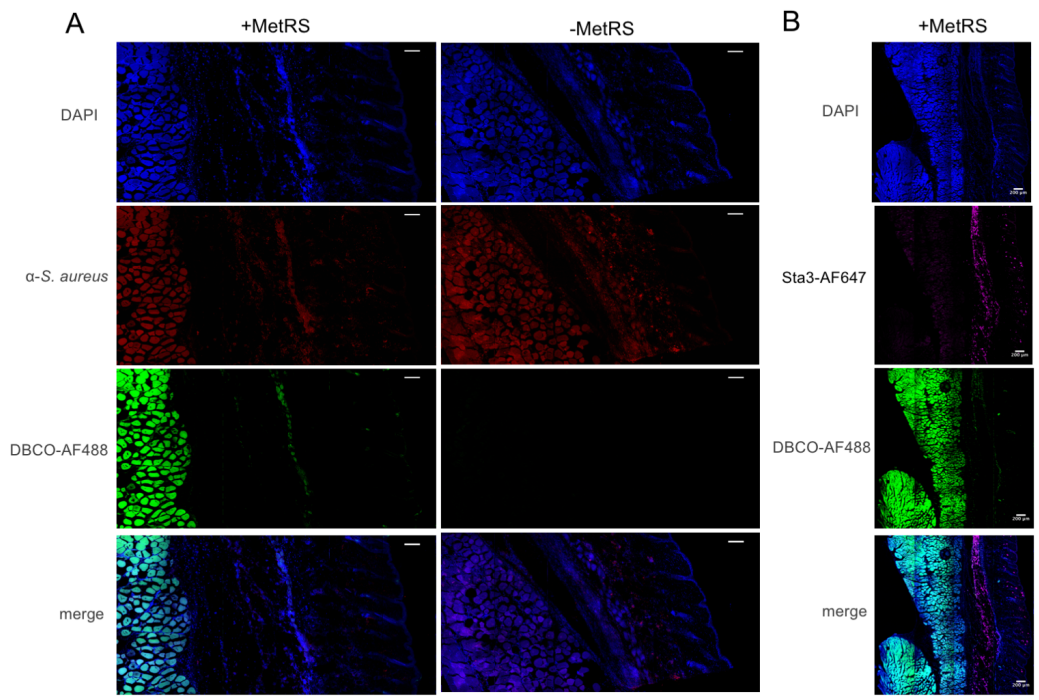
4.5 Supplementary Information



Supplemental Figure S4.1 DBCO-TAMRA labeling of incorporated azides and HCR staining for *S. aureus* in B4P1 hydrogels. Only the +MetRS samples show signal in the DBCO-TAMRA channel, while both the –MetRS and +MetRS samples show HCR signal for staphylococcal RNA. Scale bars are 50 µm.



Supplemental Figure S4.2: Controls for HCR in A) PACT abscesses and B) cryoslices. Mismatch-RNA probes show no binding to the abscess and staphylococcal probes do not bind to uninfected skin.



Supplemental Figure S4.3 A) BONCAT signal overlays with α-*S. aureus in* cryoslices. Only samples with +MetRS show signal in the DBCO-AF488 channel, while both +MetRS and MetRS samples stain positively for *S. aureus* proteins using an anti-*S. aureus* antibody. B) Staphylococcal rRNA signal does not entirely colocalize with BONCAT signal.

Experimental Procedures

Strains and growth conditions. The following strains were used in this study: *Staphylococcus aureus* USA300, JE2 (NARSA), grown in tryptic soy broth (TSB) aerobically with shaking at 250 rpm.

Embedding labeled bacteria grow*n in vitro* in hydrogel blocks. We used similar procedures as DePas *et al* with slight modifications. (9). *S. aureus* strains –MetRS (pWW412, see Appendix A) or +MetRS (pSS20_hprk, see Appendix A) were inoculated from single colonies from tryptic soy agar (TSA) plates in TSB with chloramphenicol (20

μg/mL). When the cultures reached mid-exponential phase, 2 mM of azidonorleucine (Anl) was added for 1 hour. The cultures were washed with PBS, and fixed overnight at 4 °C in 4% paraformaldehyde (PFA). Fixed samples were resuspended in 29:1 acrylamide:bis-acrylamide (v/v) (Bio-Rad 161-0146) and 0.25% VA-044 hardener (w/v) (Wako 27776-21-2) in 1X PBS for polymerization. After removing oxygen from the solution in an anaerobic chamber, blocks were polymerized at 37 ºC for three hours, without shaking, and then cut to ~3 mm^3. The staphylococcal cells were then digested with lysostaphin (50 μg/mL) in 50 mM Tris buffer for 2 hours at 37 °C. Samples were washed twice in PBS, then "cleared" for 5 days in 8% SDS in PBS at 37 °C.

Ethics statement. Animal experiments were performed in accordance with the regulations for the Institutional Animal Care and Use Committee (IACUC) at Caltech.

Embedding murine skin abscesses in hydrogel blocks. *S. aureus* strains -MetRS or +MetRS were prepared as described in Chapter 3 for skin infection. After labeling with Anl for 16 hours, the mice were culled (CO_2) and the abscesses excised. They were directly added to 4% paraformaldehyde (PFA) for fixation overnight at 4 °C then washed with PBS. Fixed samples were resuspended in 29:1 acrylamide:bis-acrylamide (v/v) (Bio-Rad 161-0146) 1% paraformaldehyde (PFA, EMS #15713), and 0.25% VA-044 hardener (w/v) (Wako 27776-21-2) in 1X PBS for polymerization. The staphylococcal cell wall was then digested with lysostaphin (50 μg/mL) in 50 mM Tris buffer for 6-12 hours at 37

°C. Samples were washed once in PBS, then cleared for 5-21 days in 8% SDS in PBS at 37 °C.

Click Reaction. After washing the SDS out of the cleared samples with three washes of PBS, free cysteines were blocked with 100 mM iodoacetamide in PBS at room temperature for 4-16 hours. Oftentimes, samples are reduced prior to blocking, but using 10 mM dithiothreitol (DTT) before iodoacetamide treatment showed higher background labeling. DBCO-TAMRA or DBCO-AF488 (Click Chemistry Tools) was added to a fresh solution of 100 mM iodoacetamide to a concentration of 5 μM and the reaction proceeded for 1 hour. The samples were washed twice in PBS (30 min each), rotated end over end in 50% DMSO in PBS overnight, then washed three more times in PBS.

HCR. 5'-ATTTCACATTTACAGACCTCAACCTACCTCCAACTCTCAC-3' was added to the 3' end of the DNA probe (termed "B4" (10)). DNA hairpins (Molecular Instruments) conjugated to either AlexaFluor-488 or AlexaFluor-647 as indicated were used with the appropriate initiator probe sets.

We initially used a previously-reported STA3 probe (17), but found high signal-to-noise with this sequence, so we increased the length of our probe from 16 nucleotides to 31 nucleotides. Microscopy images of the HCR in abscesses were all performed with this STA3_long probe. Sequence in Supplementary Table 1.

Hybridization: Samples were hybridized in 500 µL of HCR hybridization buffer (100 µL of 20X SSC, 100 mg dextran sulfate (Sigma D6001), 250 µL formamide, ddH_2O to 1 mL) with 30 nM initiator probe at 46 ºC, with shaking, for 24-48 hours. All solutions were filter sterilized. Excess probe was removed by washing each sample in 50 mL 84 mM FISH wash buffer (840 µL of 5 M NaCl, 1 mL of Tris-HCl [pH 7.6], 500 µL of 0.5 M EDTA [pH 7.2], 100 µL of 5% SDS, and Milli-Q H_2O to 50 mL) at 52 °C for 6 hours in a water bath.

Amplification: Hairpin pairs were first heated at 95 °C for 90 seconds in a thermocycler in separate PCR tubes, then cooled at room temperature for 30 minutes. Each hairpin in a pair was added to a final concentration of 115 nM. Amplification buffer with the appropriate hairpin mixture (120 µL) was then added to each sample in a 1.5 mL centrifuge tube. Samples were incubated at room temperature with gentle shaking for 24-48 hours. After amplification, samples were washed in 337.5 mM FISH wash buffer (3375 µL of 5 M NaCl, 1 mL of Tris-HCl [pH 7.6], 500 µL of 0.5 M EDTA [pH 7.2], 100 µL of 5% SDS, and Milli-Q H_2O to 50 mL) at 48 °C for 3 hours in a water bath. Samples were then incubated in 250 µL RIMS with 10 µg/mL DAPI (1:1000 from 10 mg/mL stocks in DMSO) at room temperature for at least 16 hours before imaging.

Microscopy. Prior to imaging, samples were incubated at RT, with shaking, overnight in RIMS with 1 µg/mL DAPI. Samples were then mounted on slides in 0.9 mm or 1.7 mm Coverwell perfusion chambers (Electron Microscopy Services) with a coverslip on the top. Imaging was performed using a Zeiss LSM 880 confocal microscope with a Plan-

Apochromat 10x/0.45 M27 objective (WD 2.0 mm). All images and Z-stacks were collected in 12-bit mode, with at least 1024x1024 scan format. Images were processed using Imaris imaging software (Bitplane) or the FIJI distribution of ImageJ (18).

Histological examination of mouse skin abscesses. Mouse skin was harvested 2 days after inoculation and fixed in 10% neutral-buffered formalin for 48 hours. Fixed tissues were embedded in paraffin, sectioned (5 μm), mounted on slides, and stained with hematoxylin and eosin (Pacific Pathology).

Cryosectioning, DBCO-staining, immunostaining, and imaging. Abscesses were excised, placed in optimum cutting temperature (O.C.T.) compound, and frozen at -80 °C. They were then cut into ~15 μm sections and deposited onto glass slides. The slides were fixed with 2% paraformaldehyde (PFA), permeabilized with 0.1% Triton-X, and treated with lysostaphin at 37 °C. The samples were next blocked in 100 mM iodoacetamide for 30 min in the dark, then reacted with 5 μM DBCO-AlexaFluor488 (Click Chemistry Tools) for 15 min. After washing, the samples were blocked with 1% mouse serum, then treated with a 1:3000 dilution of anti-*Staphylococcus aureus* antibody (polyclonal, rabbit) (Ab37644, Abcam). A goat anti-rabbit secondary antibody conjugated to AlexaFluor555 (Thermo Scientific) was then added at 1:10,000 dilution. Samples were washed with PBS, and Vectashield Antifade with DAPI was added prior to coverslips. Imaging was performed using a Zeiss LSM 880 confocal microscope with a Plan-Apochromat 10/0.45-

numerical aperture M27 objective (WD 2.0 mm) or a 25X objective. Image reconstruction and analysis was performed in the FIJI distribution of ImageJ (18).

Supplemental Table 1: Probes used in this study:

Species	Name	Sequence
Staphylococcus	STA3 (17)	GCACATCAGCGTCAGT
Staphylococcus	STA3_long	GATCCCCACGCTTTCGCACATCAGCGTCAGT
-	Mismatch	ACTCCTACGGGAGGCAGC
-	Random	AGCAGGTCGAACTCCTTGAG

References

1. Stacy A, McNally L, Darch SE, Brown SP, & Whiteley M (2016) The biogeography of polymicrobial infection. *Nat Rev Microbiol* 14(2):93-105.
2. Jorth P, *et al*. (2015) Regional Isolation Drives Bacterial Diversification within Cystic Fibrosis Lungs. *Cell Host Microbe* 18(3):307-319.
3. Lai HM, Ng WL, Gentleman SM, & Wu W (2017) Chemical Probes for Visualizing Intact Animal and Human Brain Tissue. *Cell Chem Biol* 24(6):659-672.
4. Chung K, *et al*. (2013) Structural and molecular interrogation of intact biological systems. *Nature* 497(7449):332-7.
5. Yang B, *et al*. (2014) Single-Cell Phenotyping within Transparent Intact Tissue through Whole-Body Clearing. *Cell* 158(4):945-958.
6. Shah S, *et al*. (2016) Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. *Development* 143(15):2862-2867.
7. Cronan MR, *et al*. (2016) Macrophage Epithelial Reprogramming Underlies Mycobacterial Granuloma Formation and Promotes Infection. *Immunity* 45(4):861-876.
8. Cronan MR, *et al*. (2015) CLARITY and PACT-based imaging of adult zebrafish and mouse for whole-animal analysis of infections. *Dis Model Mech* 8(12):1643-1650.
9. DePas WH, *et al*. (2016) Exposing the Three-Dimensional Biogeography and Metabolic States of Pathogens in Cystic Fibrosis Sputum via Hydrogel Embedding, Clearing, and rRNA Labeling. *Mbio* 7(5): e00796-16.
10. Choi HMT, Beck VA, & Pierce NA (2014) Next-Generation *in Situ* Hybridization Chain Reaction: Higher Gain, Lower Cost, Greater Durability. *ACS Nano* 8(5):4284-4294.
11. Dirks RM & Pierce NA (2004) Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci* 101(43):15275-15278.

12. Thurlow LR, *et al*. (2013) Functional Modularity of the Arginine Catabolic Mobile Element Contributes to the Success of USA300 Methicillin-Resistant *Staphylococcus aureus*. *Cell Host Microbe* 13(1):100-107.
13. Li M, *et al*. (2010) Comparative analysis of virulence and toxin expression of global community-associated methicillin-resistant *Staphylococcus aureus* strains. *J Infect Dis* 202(12):1866-1876.
14. Mahdavi A, *et al*. (2014) Identification of secreted bacterial proteins by noncanonical amino acid tagging. *Proc Natl Acad Sci* 111(1):433-438.
15. Chatterjee SS, *et al*. (2013) Essential *Staphylococcus aureus* toxin export system. *Nat Med* 19(3):364-367.
16. Otto M (2014) *Staphylococcus aureus* toxins. *Curr Opin Microbiol* 17:32-37.
17. Tavares A, Inacio J, Melo-Cristino J, & Couto I (2008) Use of fluorescence *in situ* hybridization for rapid identification of staphylococci in blood culture samples collected in a Portuguese hospital. *J Clin Microbiol* 46(9):3097-3100.
18. Schindelin J, *et al*. (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9(7):676-682.

Chapter 5

TOWARDS CELL-SPECIFIC PROTEOMICS OF THE GUT MICROBIOTA

5.1 Abstract

After developing cell-specific BONCAT in mice to study host-pathogen interactions (Chapter 3), we sought to extend the system to study a commensal microbe that lives within the gut. *Bacteroides fragilis* thrives within the crypts and mucus layers of the mammalian gut, promotes the development of a healthy immune system, and secretes numerous proteins involved in colonization. We utilized a mutant phenylalanyl-tRNA synthetase along with the noncanonical amino acids 4-azido-L-phenylalanine (Azf) and 4-ethynyl-L-phenylalanine (Ef) to tag proteins synthesized by the bacteria within the guts of living mice. This work demonstrates the first use of cell-selective BONCAT within an obligate anaerobe.

This work was a collaborative effort with Judy Shon, Gregory P. Donaldson, Kristie Yu, and Sarkis K. Mazmanian.

5.2 Introduction

Mammals and their resident microbes have formed strong alliances throughout their evolution together over time (1). The mammalian system supplies numerous molecules to its denizens, which produce signaling factors of their own to communicate with host organs (2, 3). Identifying these signaling molecules and interrogating their communication has become an important endeavor in the quest to improve human health: For example, a common gram-negative resident of the gut, *Bacteroides*, has been shown to regulate the host immune system and suppresses disease with its capsular polysaccharide (PSA) (4-6). A global proteomic study of these resident microbes in their host environment would be invaluable, but experimentally challenging due to the immense quantity of host and food proteins in the environment and the technical limitations of shotgun proteomics (7, 8).

To achieve targeted proteomic studies of *Bacteroides* within the gut, we developed cell-selective bioorthogonal noncanonical amino acid tagging (BONCAT) for use with this commensal microbe. The technique relies on the expression of a mutant tRNA-synthetase (aaRS) that can recognize a noncanonical amino acid (ncAA) and incorporate it in a residue-selective manner into newly-synthesized proteins (9). In this example, *Bacteroides* successfully expressed a PheRS mutant engineered to incorporate Azf or Ef ncAAs, effectively labeling its proteins with a bioorthogonal tag for detection and enrichment. Other proteins from the mouse or food will not have this tag and can be removed from the sample prior to analysis. This tool allows researchers to specifically label *Bacteroides* proteins synthesized *in vivo* in an effort to better understand the methods of

communication and survival this commensal microbe leverages in the harsh environment of the gut.

5.3 Results

Adaptation of cell-specific BONCAT to *Bacteroides fragilis*. We inserted the *E. coli* coding sequence of either a mutant MetRS (9, 10) or a mutant PheRS (11) into an *E. coli* – *Bacteroides* shuttle vector (pFD340) (12). Attempts to chromosomally incorporate these tRNA-synthetases using newer vectors were unsuccessful (13, 14). After conjugation of these constructed plasmids or an empty plasmid control into *B. fragilis*, we grew the strains anaerobically. After a 1:100 refresh into supplemented brain heart infusion (BHIS) media, we added PBS containing either 1 mM azidonorleucine (Anl), 200 μM 4-azido-phenylalanine (Azf), or 200 μM 4-ethynyl-phenylalanine (Ef) for 16 hours. The cultures were lysed and subjected to the copper-catalyzed azide-alkyne cycloaddition (CuAAC) with either alkyne-TAMRA (for the azide ncAAs Anl and Azf) or azide-TAMRA (for the alkyne ncAA Ef). While the strains expressing the mutant MetRS (+MetRS) did not show any incorporation of Anl (Fig S5.1), the strain expressing the mutant PheRS (+PheRS) was able to robustly incorporate both Azf and Ef (Fig 5.1C).

Notably, exposure to the ncAA azidohomoalanine (Aha), which does not require expression of a mutant aaRS as it is recognized by the endogenous methionyl-tRNA synthetases of most organisms, resulted in very strong labeling of *B. fragilis* in culture (Fig S5.1B). While Aha is consistently incorporated by endogenous MetRS, labeling

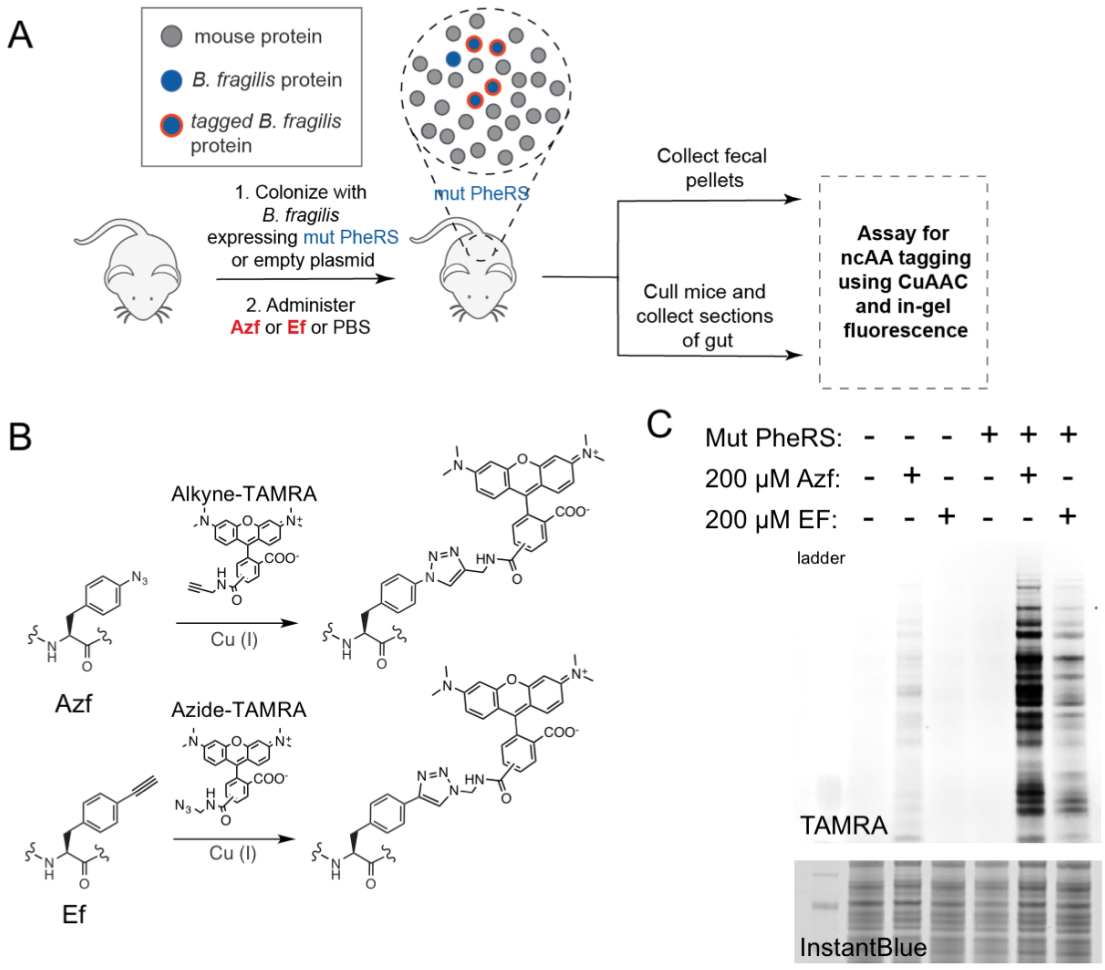


Figure 5.1: A) Scheme used to detect proteins synthesized by *Bacteroides fragilis* within the mouse gut. After colonization of mice with strains expressing a mutant phenylalanyl-tRNA synthetase (PheRS) or not, we administer azidophenylalanine (Azf) or ethynylphenylalanine (Ef). We then collect fecal pellets and homogenize prior to assessing for azide incorporation into proteins. B) Scheme of copper-catalyzed azide-alkyne cycloaddition (CuAAC) with Azf or Ef and the corresponding TAMRA. C) SDS-PAGE gel of *B. fragilis* after 16 hours of labeling during anaerobic growth in BHIS. Only cells that expressed the mutant PheRS and were labeled with Azf or EF showed fluorescence. Azf appears to have slightly higher background than Ef. InstantBlue shows equal protein loading across samples.

usually requires an environment with a depleted supply of methionine; *B. fragilis* showed very strong labeling with Aha even in the supplemented BHIS media without any

methionine depletion. Differences in the species' endogenous protein synthesis machinery, such as sequences of the MetRS or associated tRNAs may explain the difference in AHA incorporation observed between *B. fragilis* and *E. coli*, and the apparent incompatibility with the expressed *E. coli* mutant MetRS: in this case, the *B. fragilis* MetRS shares only 35% sequence homology with the *E. coli* version. Whenever adapting cell-selective BONCAT for a new species, we recommend characterizing labeling of both the PheRS and the MetRS mutant synthetases, and picking the one that shows strong labeling without notably affecting growth.

Colonization of germ-free (GF) mice with *B. fragilis* strains. Mice that had been maintained germ-free were colonized with the +PheRS strain of *B. fragilis* or the empty plasmid control. We allowed the colonization to expand within the gut over 2-4 weeks, then collected fecal pellets and verified by both serial dilution and colony PCR that each plasmid was maintained (Fig 5.2).

Labeling of *B. fragilis* within the gut of GF mice. An important consideration when performing *in vivo* cell-selective BONCAT is the method of delivery of the ncAA in proper concentrations to the cells of interest. Previously, we directly injected the site of infection with the ncAA to label bacteria (Chapter 3), but delivering the ncAA to the gut presented new challenges. Little is known about the pharmacokinetics of Azf and Ef ncAAs within animals and whether they would be stable in the harsh conditions of the stomach and gut. We assessed two different methods of delivering the ncAA to the gut: oral gavage (PO, or force feeding) and adding the ncAA to the mice's drinking water.

We gavaged the mice with 100 μL of 0, 0.5, 5.0, and 50 mM Azf dissolved in PBS (pH 7.5), then collected fecal pellets six hours and 24 hours after. These fecal pellets were resuspended in PBS, lysed, and subjected to CuAAC with an alkyne-TAMRA. Only the mice colonized with the +PheRS *B. fragilis* and gavaged with 50 mM Azf displayed signal six hours post-gavage via in-gel fluorescence (Fig 5.2C). After 24 hours, this signal disappeared, indicating that the Azf had been cleared from the system by this time. Importantly, without protease inhibitors in the lysis buffer, the proteins in solution were found to degrade very rapidly. Even with 10x protease inhibitors in the lysis buffer, we observed a very low signal associated with intact *B. fragilis* proteins, with the strongest TAMRA and InstantBlue protein labeling appearing in a broad band of low molecular weight fragments. The function of the gut is to break down protein: thus, high amounts of protease inhibitors must be used in these experiments in order to minimize protein degradation prior to analysis.

We also tested whether supplying the mice with Azf dissolved in drinking water would induce labeling. The mice were given opaque 50 mL conical tubes that had been converted to water bottles, with 25 mM Azf and sweetened with 2% sucrose to encourage consumption. After two days, we observed that they were only drinking ~1 mL/day, which is significantly less than mice's standard consumption of 3-5 mLs/day. We hypothesized that this could be due to the known bitter taste of phenylalanine and its derivatives (15), although we did not test this hypothesis with our own taste receptors. We collected fecal pellets from the mice that had been given this water and assessed labeling using the

CuAAC reaction. Two days after introduction of Azf to the water we observed strong labeling via in-gel fluorescence in the mice that were colonized with the +PheRS strain (Fig 5.2D). Although we successfully labeled *B. fragilis* proteins in 2 days, we reasoned that this method was not ideal as a constant supply of drinking water would 1) require a large quantity of Azf and 2) render us unable to label the proteins with sufficient time-resolution, as we did not monitor when each mouse drank from the water bottle. Oral gavage addresses both of these issues, but further experiments will have to be performed to assess the time resolution of this method and whether we can increase labeling intensity.

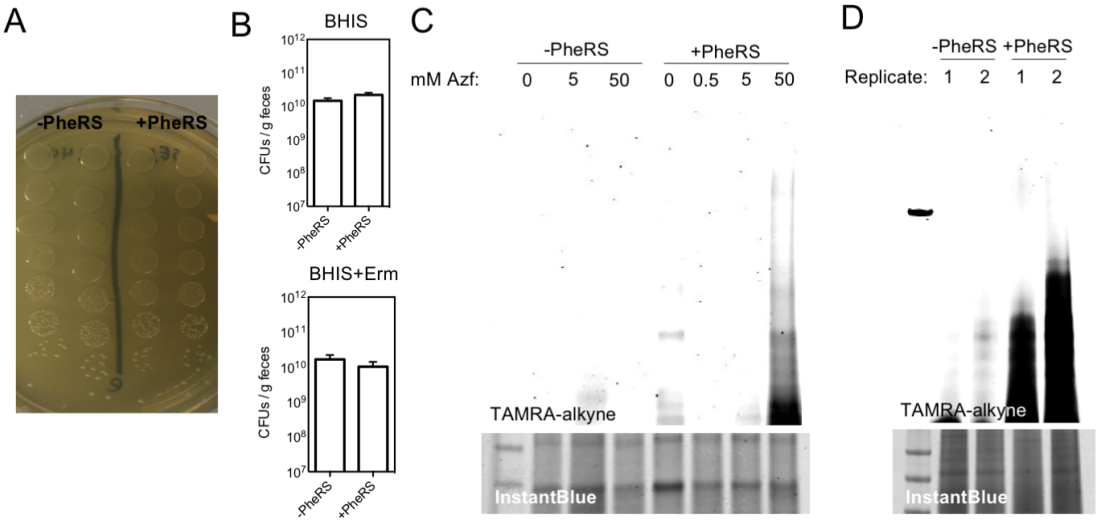


Figure 5.2: A) Dilution series of *B. fragilis* from fecal pellets 2 weeks post colonization of GF mice on erythromycin (Erm) plates. B) Bar graphs showing total CFU counts and Erm-resistant CFU counts. C) In-gel fluorescence gel of alkyne-TAMRA clicked fecal pellet obtained from GF mice 6 hours post 100 µL gavage of noted solutions D) Similar to C except 48 hours post introducing 25 mM Azf in water bottles.

5.4 Discussion and Future Directions

Surprisingly, the endogenous *B. fragilis* MetRS was able to incorporate Aha to a higher degree than other organisms even without depletion of available methionine. A comparison of charging rates of Aha onto tRNA for both mice and *Bacteroides* may be useful: If *Bacteroides* incorporate Aha at a significantly higher rate than the host, feeding Aha to mice may result in semi-specific labeling of the *Bacteroides*' proteins without the necessity to create a genetically-altered strain. In addition, the expression of a mutant *E. coli* MetRS did not show labeling with Anl, suggesting an incompatibility with either the Anl ncAA or between the exogenous synthetase and endogenous tRNA. Sequence analysis of the tRNA recognition sites may explain this incompatibility. Overall, these results highlight the value of having more than one choice of enzyme to perform cell-selective BONCAT. Testing different synthetases (aaRS) and ncAAs whenever it is adapted to a new species will be useful for future studies.

We developed cell-selective BONCAT for use in *Bacteroides*, an important constituent of the mammalian gut microbiota. Expression of a mutant PheRS enabled incorporation of Azf or Ef into microbial proteins while in the gut of mice, which could then be detected using click chemistry. This system can be used to label, visualize, and identify proteins made by the microbiota within a mammalian host, and complements other methods used to study host-microbe interactions *in vivo* (16, 17).

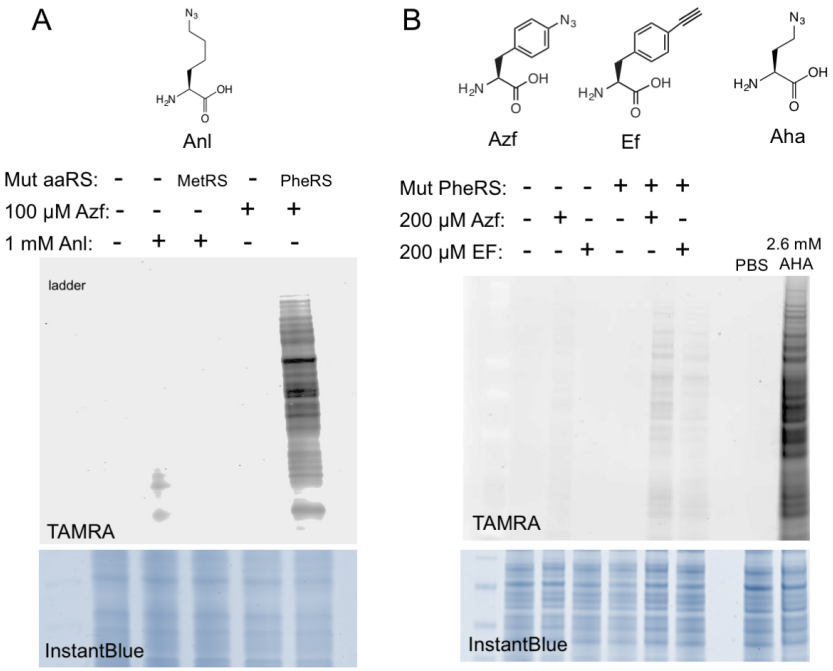
5.5 Experimental Procedures

Bacterial strains and media. *E. coli* were aerobically grown in Luria-Bertani broth (LB) at 37 °C and *B. fragilis* was grown anaerobically in brain heart infusion broth supplemented with vitamin K (0.5 μg/mL) and hemin (5 μg/mL) at 37 °C. Ligation-independent cloning was performed in *E. coli* DH10B cells and positive colonies were selected using ampicillin (100 μg/mL). Conjugations were into wild-type *B. fragilis* following protocols of Lee *et al*. (12) and selected for using erythromycin (10 μg/mL) and gentamicin (100 μg/mL).

Mice. 8-10 week-old germ-free (GF) Swiss Webster mice were purchased from Taconic Farms and bred in flexible film isolators. For colonization experiments, GF mice were transferred to freshly autoclaved microisolator cages, fed *ad libitum* with a standard autoclaved chow diet and given autoclaved water supplemented with 10 μg/mL of erythromycin and 100 μg/mL of gentamicin. All procedures were performed in accordance with the approved protocols using IACUC guidelines of the California Institute of Technology.

Colonization experiments. 8-12 week-old GF Swiss Webster mice were exposed to the *B. fragilis* strain of interest by adding 2 mL of an overnight culture to their bedding. After two weeks and four weeks, fresh fecal samples were collected, weighed, homogenized and serially diluted in PBS for plating on selective media to determine bacterial colony forming units (CFU) per gram of feces. Colonies from these plates were subsequently assayed by PCR to confirm the identity of the plasmid.

Labeling experiments *in vivo*. Colonized mice were gavaged with 100 µL of an Azf solution or PBS (as indicated), and fresh fecal samples were collected as in Colonization Experiments, except with the addition of 10X EDTA-free protease inhibitors (Roche). The resuspended bacterial solution in PBS was brought to 1% SDS and boiled for 10min to lyse the cells. The click reaction was performed on ~10 µg of protein from each sample using a Click-It Kit (Thermo Scientific) and the samples were run on an 8-12% Bis-Tris SDS-PAGE gel, followed by imaging on a Typhoon gel scanner.



Supplemental Figure S5.1: A) *B. fragilis* expressing the mutant MetRS shows no incorporation of azidonorleucine (Anl) after 18 hours of anaerobic growth in the presence of the ncAA. B) The addition of 2.6 mM azidohomoalanine (Aha) to BHIS over 18 hours of anaerobic growth leads to strong labeling. The entirety of the gel is included to compare the Aha incorporation to azidophenylalanine (Azf) or ethynylphenylalanine (Ef) incorporation.

References

1. McFall-Ngai M, *et al*. (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci* 110(9):3229-3236.
2. Yano JM, *et al*. (2015) Indigenous Bacteria from the Gut Microbiota Regulate Host Serotonin Biosynthesis (vol 161, pg 264, 2015). *Cell* 163(1):258-258.
3. Donia MS & Fischbach MA (2015) Small molecules from the human microbiota. *Science* 349(6246):1254766.
4. Shen Y, *et al*. (2012) Outer Membrane Vesicles of a Human Commensal Mediate Immune Regulation and Disease Protection. *Cell Host Microbe* 12(4):509-520.
5. Donaldson GP, Lee SM, & Mazmanian SK (2016) Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol* 14(1):20-32.
6. Wexler HM (2007) *Bacteroides*: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev* 20(4):593-621.
7. Mahowald MA, *et al*. (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc Natl Acad Sci* 106(14):5859-5864.
8. Wilson MM, Anderson DE, & Bernstein HD (2015) Analysis of the Outer Membrane Proteome and Secretome of *Bacteroides fragilis* Reveals a Multiplicity of Secretion Mechanisms. *PLoS One* 10(2): e0117732.
9. Ngo JT, Schuman EM, & Tirrell DA (2013) Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells. *Proc Natl Acad Sci* 110(13):4992-4997.
10. Mahdavi A, *et al*. (2014) Identification of secreted bacterial proteins by noncanonical amino acid tagging. *Proc Natl Acad Sci* 111(1):433-438.
11. Yuet KP, *et al*. (2015) Cell-specific proteomic analysis in *Caenorhabditis elegans*. *Proc Natl Acad Sci* 112(9):2705-2710.
12. Lee SM, *et al*. (2013) Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature* 501(7467):426-9.
13. Whitaker WR, Shepherd ES, & Sonnenburg JL (2017) Tunable Expression Tools Enable Single-Cell Strain Distinction in the Gut Microbiome. *Cell* 169(3):538-546.
14. Mimee M, Tucker AC, Voigt CA, & Lu TK (2015) Programming a Human Commensal Bacterium, *Bacteroides thetaiotaomicron*, to Sense and Respond to Stimuli in the Murine Gut Microbiota. *Cell Syst* 1(1):62-71.
15. Maehashi K & Huang L (2009) Bitter peptides and bitter taste receptors. *Cell Mol Life Sci* 66(10):1661-1671.
16. Geva-Zatorsky N, *et al*. (2015) *In vivo* imaging and tracking of host-microbiota interactions via metabolic labeling of gut anaerobic bacteria. *Nat Med* 21(9):1091-100.
17. Berry D, *et al*. (2013) Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing. *Proc Natl Acad Sci* 110(12):4720-4725.

Chapter 6

QUANTIFYING ENRICHMENT OF AZIDE-TAGGED PROTEINS

6.1 Abstract

We describe the development of SPIQE (Spike Peptide In to Quantify Enrichment), a novel method of quantifying the gain in enrichment of azide-tagged proteins from untagged proteins. The SPIQE peptide represents a small fraction of the proteomic samples but can be used to validate the method of enrichment of samples through various bio-conjugation strategies. We compare the enrichment values obtained by SPIQE to those obtained by stable isotope labeling of amino acids in culture (SILAC). Similar values indicated that SPIQE would be a useful addition to samples that do not have other characteristics to quantify enrichment.

This work was a collaborative effort with Judy Shon, Michael J. Sweredoski, Annie Moradian, Roxana Eggleston-Rangel, and Sonja Hess.

6.2 Introduction

Shotgun proteomics has revolutionized our ability to globally profile protein expression in complex cellular environments (1). This bottom-up proteomics strategy requires digestion of whole proteins within the cellular lysate, separation of the resulting peptides by liquid chromatography (LC), and subsequent detection by tandem mass spectrometry (MS/MS) (2). To broadly survey the entire complex lysate, the mass spectrometer selects the top precursor ions detected in the first MS scan (which are often the most abundant, since intensity and abundance are correlated) and subjects them to a second MS/MS fragmentation scan centered on that mass-to-charge ratio (m/z). The masses of these peptides are then compared to a theoretical digest of the organism's proteome using computational tools (3). Because of the nature of shotgun proteomics and its limited dynamic range, highly abundant proteins in the sample are often the most likely to be found (4). Sometimes investigators are interested in identifying only a small subset of these proteins, such as proteins that have been newly-synthesized, proteins with a particular post-translational modification (PTM), or proteins that are in an active enzymatic state. In these cases, if too many peptides from other proteins are in the sample, their signal will drown out the signal of the proteins of interest (Fig 6.1).

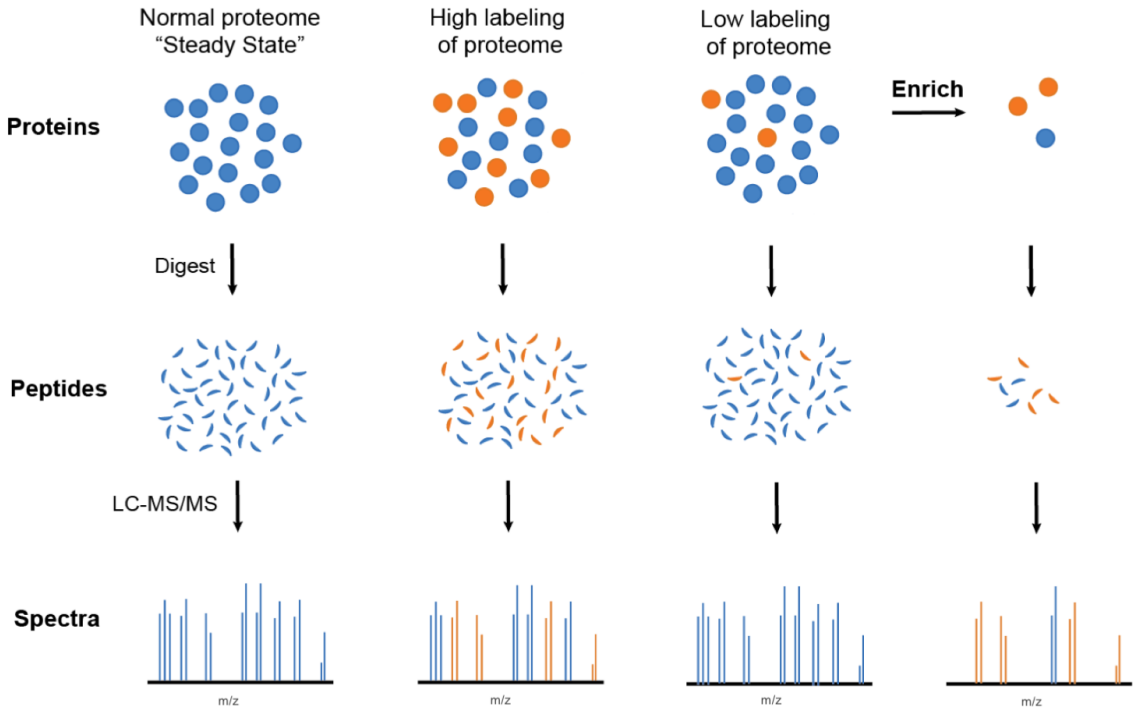


Figure 6.1: Scheme depicting shotgun proteomics and the difficulty in finding proteins of interest (orange) when they are not in high abundance compared to the background proteome (blue).

One way to enrich for these proteins of interest is to chemically append a tag that can subsequently be used to "fish out" the proteins of interest from the rest of the proteome (5). Many groups use the azide group as a tag because it is small, absent in biological systems, and able to specifically react with both terminal alkynes in the presence of Cu(I) (Copper-catalyzed alkyne-azide cycloaddition, or CuAAC) or strained cyclooctynes (Strain-promoted alkyne-azide cycloaddition, or SPAAC), which are also known as "click" reactions (6, 7). Once an azide has been introduced into a protein of interest, it can then be conjugated to a fluorophore or enrichment handle (such as biotin or directly to a resin) for detection and enrichment. In this manner, azides have been used to identify

newly-synthesized proteins (8-11), sites of post-translational modification (12-15), and active enzymes (16, 17).

After using CuAAC or SPAAC to append a linker attached to an affinity-handle (such as biotin) or directly to a resin, stringent washes remove the untagged proteins. The azide-tagged proteins can then be digested directly on the resin and eluted to collect their peptide fragments, or cleaved and eluted intact if the chosen linker is cleavable (18). While cleavable linkers are useful for finding the precise sites of azide incorporation into the peptide, they are sometimes unnecessary for the experiment. Identifying the site of azide-tagging on the protein is challenging, since any peptide with a tag adduct represents a very small fraction of the proteome – difficult to detect with a global search. Furthermore, azides are often heterogeneously incorporated into peptides; thus, if a cleavable resin is used the total ion current of each peptide may be distributed over multiple species and become more difficult to detect. If the goal of an experiment is to find newly-synthesized proteins in a sample, digesting the entire protein leads to more peptides that can be detected by the MS and higher confidence protein identifications. The drawback of this method is that it is more difficult to prove that the peptide came from an azide-tagged protein.

Enrichment of azide-tagged proteins from untagged ones is not a perfect process and depends on numerous factors such as the fraction of the proteins that are azide-tagged, and the relative rates of the click reaction to non-specific background reactions. Samples that

lack introduced azide tags when combined with resin will often elute nonspecifically bound proteins, which the mass spectrometer will detect. Running a "mock" enrichment alongside any experiment is often necessary to compare the amount of protein eluted from the resin in the absence or presence of azide-tagging. The ratio of these relative peptide concentrations can be used as a semi-quantitative method to determine the "strength" or level of enrichment gain. A quantitative method to determine the gain of an enrichment method on the mass spectrometer itself would be useful, which is why we developed SPIQE (Spike Peptide In to Quantify Enrichment) for reporting the gain of any azide-based proteomic enrichment. Herein we report the development of this peptide method and compare it to other methods often used to quantify enrichments.

Quantitative mass spectrometry

Mass spectrometry is not inherently quantitative; peptides do not ionize with identical efficiencies/rates in an ion trap and thus, the reported intensities of peptides with different sequences cannot be directly compared. Furthermore, the triggering of an MS/MS fragmentation is often stochastic for all but the highest abundance proteins. If a given peptide is not fragmented in one sample it does not indicate its absence from the sample, only that the peptide's rank order was not high enough to trigger fragmentation. Quantitative MS techniques seek to solve this problem by using isotope labels for valid comparisons of peptides between samples. For example, stable isotope labeling of amino acids in culture (SILAC) relies on the use of "light" (normal) or "heavy" amino acids within proteins (19). Heavy amino acids contain isotopes such as ^{15}N or ^{13}C, which are

indistinguishable from light amino acids within the cell, but easily identifiable using MS due to the change in their mass-to-charge (m/z) ratio. The light and heavy peptides elute at the same time during the LC separation, and the ratios can be directly compared as a means to quantify proteins between samples. Furthermore, if one of the peptides does not trigger an MS/MS fragmentation, the first MS scan can be used to estimate the abundance of the pair.

Both Eichelbaum *et al*. and Yuet *et al*. showed that combining SILAC with azide-based metabolic labeling methods allows for both quantification of enrichment and determination of whether a given protein came from an azide-tagged sample (10, 11). In this strategy, one sample contains only light amino acids and no azide tagging, and another sample contains the heavy amino acid(s) and azide-tagged proteins. Mixing these samples together 1:1 and subjecting them to a click-based enrichment step increases the fraction of peptides with heavy amino acids (20-22). The gain of the enrichment step can be sampled for each heavy:light peptide pair observed by the MS, resulting in a high confidence measure of enrichment (Fig 6.2). Another strategy called Heavy Isotope Labeled Azidohomoalanine Quantification (HILAQ) uses a heavy-isotope labeled azide-containing amino acid, which can both identify and quantify labeled sites (23). SILAC and HILAQ are valuable tools to quantify labeling when combined with azide-enrichment methods, but some samples are not amenable to this type of metabolic tagging or prohibitively expensive to label, such as in studies using animals.

6.3 General Approach

Our approach relies on synthetic peptides with two distinct features (Fig 6.2): One version of the peptide has no azide tag and a light arginine, while the other has an azide-containing amino acid in place of methionine and a heavy arginine. Both peptides share an internal lysine that a proteolytic enzyme (i.e. trypsin, Lys-C) can recognize and cleave. A sample spiked with these peptides models an experiment that combines an azide-tagging method with SILAC, but the sample does not need to be SILAC labeled. In a typical experiment, the light and heavy peptides are mixed 1:1 and spiked in low amounts into the sample to be enriched. During enrichment, the light peptide will not bind the resin since it has no azide, while the heavy peptide will conjugate to the resin via the click reaction. After proteolytic digestion, the released light or heavy peptides are then identified using MS and their ratios quantified to measure the gain of an enrichment method without having to label samples prior using SILAC. We call this peptide mixture SPIQE (Spike Peptide In to Quantify Enrichment).

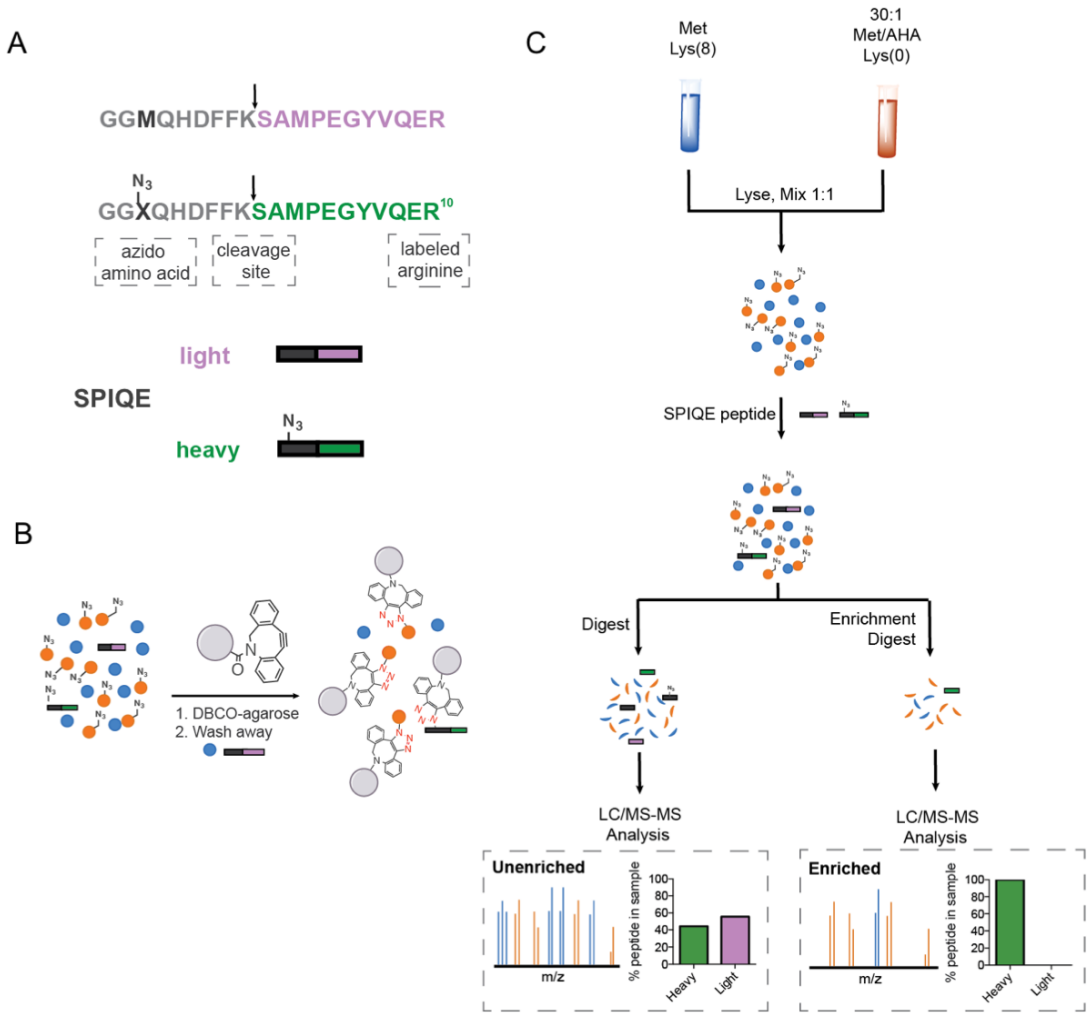


Figure 6.2: A) Sequence of SPIQE peptides and a cartoon depiction of each. B) Scheme of enrichment procedure using strain-promoted alkyne-azide cycloaddition (SPAAC). Overview of Spike-In for Quantitative Enrichment (SPIQE) method in shotgun proteomics

We aimed to compare the enrichment gains obtained from the SPIQE peptide to the SILAC strategy of quantifying enrichment, so we used SILAC-labeled lysates that had also been tagged with azides using bioorthogonal noncanonical amino acid tagging (BONCAT) (20, 21). BONCAT tags newly-synthesized proteins with azido-amino acids in a residue-specific manner and does not require the expression of any mutant aminoacyl-

tRNA synthetases or tRNAs. The method relies on the cell's endogenous machinery stochastically incorporating azidohomoalanine (Aha) instead of methionine during translation. Using a 30:1 mixture of Aha:Met removes any artifacts in quantitation associated with total methionine replacement (21). Samples tagged with azides were also fully labeled with light lysine, while untagged samples were labeled with heavy lysine. This experimental design allows us to measure the gain of our enrichment in two ways: 1) Calculating the light:heavy peptide ratios observed from the lysate using the traditional SILAC method 2) Calculating the heavy:light ratio of our SPIQE peptide within this complex mixture (Fig 6.2). We then compared the values obtained by each method to see if SPIQE could be used to measure enrichment gain when samples are not labeled by SILAC.

6.4 Results

Synthesis of SPIQE peptide. A model peptide from green fluorescent protein (GFP) (GGXQHDFFKSAMPEGYVQER, where X is methionine or azidohomoalanine, and R is light or heavy arginine (Fig 6.2), was synthesized with an N-terminal acetyl group, purified by high-performance liquid chromatography (HPLC), and characterized by MS (JPT Technologies; Berlin, Germany). We chose this peptide from GFP because it has been found often in standard MS experiments, indicating efficient ionization. The synthesized peptides were dissolved in 10% acetonitrile (ACN) and mixed together in a ~1:1 ratio.

Preparation of SILAC+BONCAT labeled lysates. *E. coli* DH10B cells were labeled in M9 minimal media with 18 amino acids (-Met, -Lys) and either 30:1 azidohomoalanine/methionine (Aha/Met) + Lys0 (light, tagged sample) or Met + Lys8 (heavy, untagged sample) for three days, with 3 separate 1:100 refreshes into new media to ensure quantitative replacement of lysine (21, 24). We verified the cells were tagged with azides using in-gel fluorescence (Fig 6.3), then digested the samples using trypsin, ran the resulting peptides on LC-MS/MS, and used MaxQuant quantitative proteomic software to search the *E. coli* proteome (UniProt) for peptides that include light or heavy lysine (3). The SILAC values were normalized and each evidence count (n = 1299) was plotted from lowest to highest SILAC ratios, which confirmed quantitative incorporation of light lysine in the Aha-tagged samples and heavy lysine in the methionine only samples. These light and heavy lysates were then mixed in a 1:1 "light + heavy" ratio and used for subsequent quantifications of enrichment gain to compare to those observed with the SPIQE peptide.

In addition to these samples, we also performed a "label-swap" experiment where cells were growth with 30:1 Aha/Met + Lys8 (heavy, tagged sample) or Met + Lys0 (light, untagged sample). We likewise confirmed quantitative incorporation of light or heavy lysine in each of the samples (Fig S6.1A). This label swap condition was used to determine the enrichment strategy that obtained the most peptides with low (or zero) SILAC ratios. By using the SILAC ratio and the number of light or heavy peptides, we found that DBCO-based resins and linker performed best for these azide-tagged samples

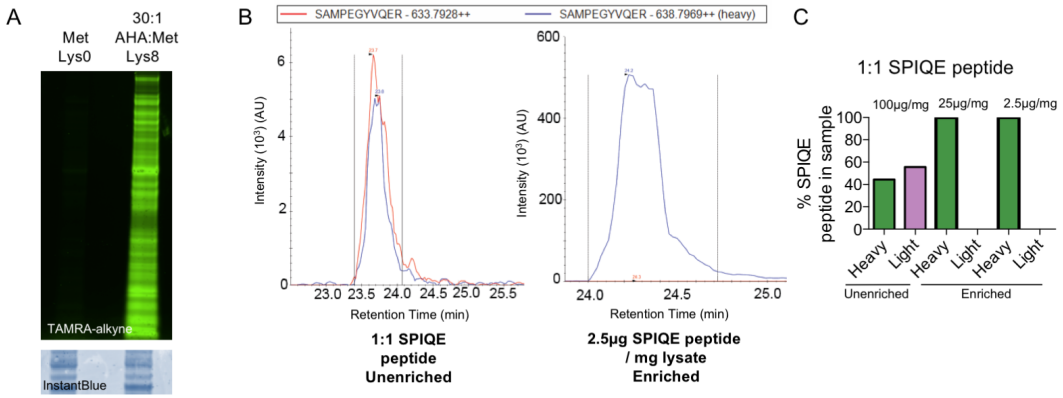
with enrichments of light peptides 100 to 1000-fold compared to pre-enriched samples (Fig S6.1B, S6.1C). We consequently used DBCO-agarose to test the SPIQE peptide enrichment.

MS analysis of unenriched SPIQE peptide mix in complex lysate. The 1:1 SPIQE peptide mix was added to 400 µL of the "light + heavy" *E. coli* lysate described above at either 25 µg peptide/mg of lysate or 2.5 µg peptide/mg of lysate. Prior to enrichment, we took 5% of this sample and reduced and alkylated the proteins prior to digestion and desalting. We then ran this sample using a directed search with a QTRAP mass spectrometer, acquiring MS/MS for the specific mass of the peptides (Fig 6.3B). The set-up and analysis was performed using Skyline, an application for both proteomic method creation and quantitative data analysis (25). By integrating the intensity of each peptide peak, we determined that this initial mixture without any enrichment was 44.4% "heavy" and 55.6% "light" (Fig 6.3C).

MS characterization of enrichment using a targeted search. We enriched the samples using DBCO-agarose, as described in the supplemental methods. Importantly, we reduced and alkylated all proteins prior to enrichment to prevent thiol-yne addition (26). The enriched peptides were digested using trypsin, desalted using stage tips, and run on a QTRAP using the same method previously used to characterize the 1:1 peptide mixture. Using Skyline software to integrate the area under the "light" and "heavy" peaks revealed that in the 25 µg peptide/mg lysate, we found >99.99% heavy peptide, and in the 2.5 µg

peptide/mg lysate, we found 99.95% heavy peptide (Fig 6.3C). Starting with a 44.44% "heavy" peptide mixture and bringing it to these levels indicate 8000-fold and 1600-fold enrichment, respectively.

MS characterization of enrichment using data-dependent acquisition. Shotgun proteomics experiments often do not use targeted searches, but instead rely on data-dependent acquisition methods to analyze peptides without prior knowledge of their sequences. We injected 100 ng of the enriched 2.5 μg peptide/mg lysate onto an Orbitrap Elite MS (Thermo Fisher Scientific) coupled to a Proxeon Easy-nLC 1000 (Thermo Fisher Scientific) and collected the top ions for fragmentation over a 1 hour run (See Experimental Procedures for more detailed information). We then used Skyline to search for the SPIQE peptide within this run and integrating the area under both curves revealed an enrichment factor of >10,000 (Fig 6.3D). This number may represent the limit of detection in this type of experiment; due to the lack of any signal for the light SPIQE peptide, we integrated the background noise to achieve this number so the confidence in this actual value is low.



D

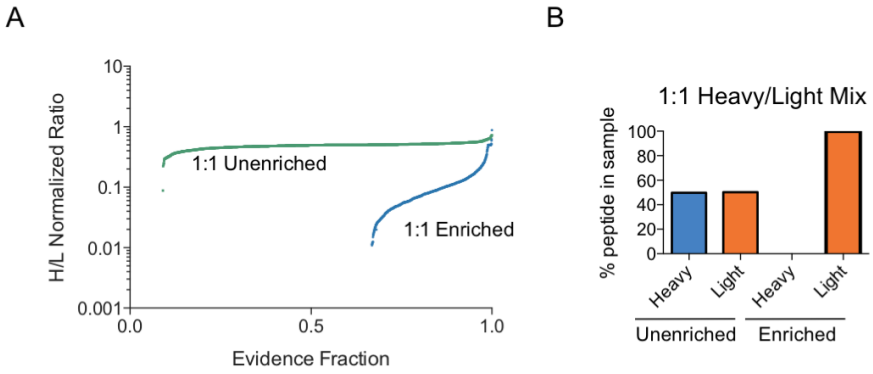
--	SPIQE peptide / mg lysate	Enrichment Method	Instrument	Peak Area (light)	Peak Area (heavy)	%Heavy	Enrichment
Unenriched	100µg	None	Q-Trap	1.15E+05	9.20E+04	44.39%	-
25ug Enriched	25µg	DBCO-agarose	Q-Trap	1.02E+04	1.17E+08	99.99%	8000
2.5ug Enriched	2.5µg	DBCO-agarose	Q-Trap	4.76E+03	8.66E+06	99.95%	1600
2.5ug Enriched	2.5µg	DBCO-agarose	Orbitrap Elite	3.63E05	7.11E+09	99.99%	15000
2.5ug Enriched	2.5µg	Dde-alkyne-agarose	Orbitrap Elite	9.29E06	5.02E+09	99.85%	540

Figure 6.3: A) *E. coli* lysate labeled with either Met+Lys0 or 30:1 AHA:Met+Lys8 and clicked to TAMRA-alkyne. Note: these lysates are not the ones used for the enrichment experiments with the SPIQE peptide as we used the label-swapped version of Met+Lys8 and 30:1 AHA:Met+Lys0 B) Traces of SPIQE peptide before and after enrichment shows the removal of the light version of the peptide. C) Table quantifying the area under the curve (AUC) from the trace in B) as well as AUCs from additional experiments described in the text.

Comparing the enrichment value of the spike-in peptide to SILAC values.

To characterize the enrichment factor using SILAC, we used MaxQuant to search the *E. coli* proteome for peptides that include light or heavy lysine. We took the light to heavy ratio of all the peptides identified by MaxQuant. After normalizing the data, we verified that while the 1:1 mix had a median value of 0.494, the enriched sample had a median value of 0 (Fig 6.4A). Because 99.99% of the peptides had SILAC values of <0.494, we

concluded that our enrichment factor is >1000, which is in agreement with the enrichment value obtained by the SPIQE peptide (Figure 6.4C).



C

	#Peptides Found	Median	%Heavy	Enrichment
Unenriched	1298	0.494	49.7%	-
DBCO Enriched	3190	<0.001	.01%	9900

Figure 6.4: A) *E. coli* lysate that had AHA + Lys0 or Met + Lys8 mixed 1:1 can be enriched for light peptides by using DBCO-agarose and quantified using SILAC. B) The enrichment factor using the SILAC method is similar to the enrichment factor quantified by SPIQE. C) The same E. coli lysate mixed 50:1 does not enrich as well using the same method.

Overall, the spike-in peptide confirmed that these azide-tagged proteins can be enriched >1000-fold following these protocols with DBCO-agarose. Furthermore, the method uses a very small (0.25% of protein lysate) amount of the spike-in peptide added to the sample that does not take up too much of the LC-MS/MS run or appreciably affect the available DBCO for reaction (Fig S6.2).

We also used MaxQuant to search for the sequence of the spike-in peptide (3). By adding the FASTA file of this spike-in peptide directly to the MaxQuant search, we hoped to be

able to quantify enrichment during any standard proteomic analysis without additional MS runs. While this search found four counts of the heavy-labeled SPIQE peptide, it did not find any of the light-labeled version so we could not quantify an enrichment gain using this method.

Characterizing a cleavable resin to find the site of incorporation. For some azide-tagged proteomic experiments, the site of the azide tag must be identified. The DBCO-agarose resin used in the prior experiments retains the azide-tagged peptide after digestion because there is no cleavable moiety; thus the azide-tagged peptide will not be found in subsequent MS (Fig S6.2C). Cleavable resins and linker with groups that are photo-cleavable, acid-labile, protease sites, or able to be reduced have been developed to release the entire protein from the resin or enrich for tagged sites (12, 27, 28). We used our spike-in peptide to quantify enrichment gain using an alkyne-agarose resin with a *N*-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl) (Dde) linker (Dde-alkyne-agarose, Click Chemistry Tools), which releases the clicked peptide upon addition of hydrazine (Fig 6.5A). This released peptide reveals a charged alkyl amine that increases ionization efficiency, but is only a small addition to the molecular mass of the peptide so it does not exceed the ideal mass range of MS. We then used Skyline to search for the SPIQE peptide within this run and integrating the area under both curves revealed an enrichment gain of 540 (Fig 6.3D, 6.5B). Skyline also revealed the presence of the released SPIQE peptide with the alkyl amine adduct, indicating this resin would be useful to identify sites of labeling (Fig 6.5C). While the released cleaved SPIQE peptide was found 20-fold

enriched over the untagged version, this number cannot be used for comparison since they are not identical peptides and do not ionize identically.

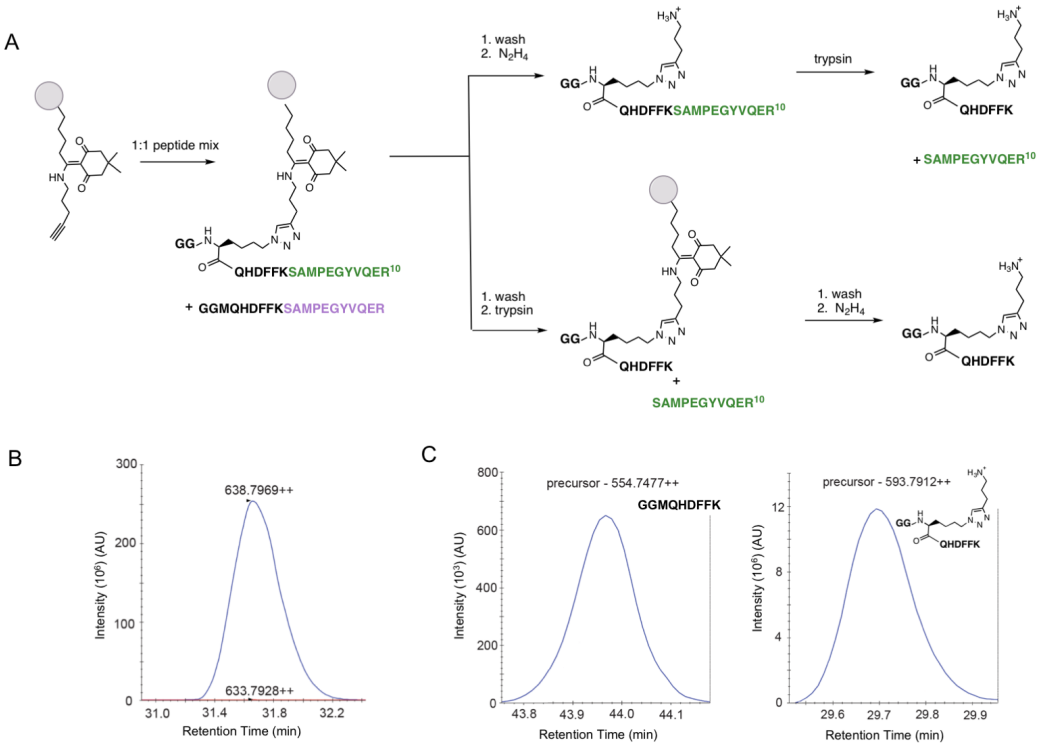


Figure 5: A) Dde resin can be cleaved with hydrazine (N_2H_4) to release the azide-tagged peptide/protein either before trypsin digest to find all peptides from azide-tagged proteins or after to find only the site of azide-incorporation into the peptide respectively. B) Trace of cleaved SPIQE peptide indicates enrichment (quantified in Fig 6.3).C) Traces of peptide after enrichment using cleavable Dde resin reveals the site of azide-tagging

Enrichment of 50:1 mix of SILAC+BONCAT labeled lysates. We also mixed the quantitatively labeled lysates in a 50:1 heavy:light ratio to test the ability to enrich with lightly labeled samples. We verified that the samples were 98% heavy labeled prior to enrichment. After enrichment, 84% of the sample was heavy labeled, indicating an

enrichment gain of only 8.7 (Fig S7.3). This enrichment gain is much lower than the values calculated for the 1:1 mixture.

6.5 Discussion

We have developed SPIQE, a method to quantify enrichment gain in the absence of other labels. The values obtained by SPIQE matched values calculated by SILAC quantification. Thus, we envision this peptide could be "spiked-in" to samples without SILAC labels as a method of quantifying enrichment without additional MS runs or other wet-lab procedures. Using heavy:light labeled lysates, we also quantified and compared the ability of various resins and linkers to enrich azide-tagged proteins, concluding that DBCO-based linkers worked better than terminal alkynes. We also found that Dde-alkyne-agarose resin can be used to find sites of azide-tagging. Finally, we determined that samples with a low abundance of tagged proteins (50:1 labeled/unlabeled lysate) have much lower enrichment gains. We envision this method could be used to quickly optimize specificity and gain of enrichment techniques.

Enrichment of azide-tagged proteins is always a battle between specificity and reactivity. Click reactions that are 100- to 1000- fold more selective for azides over other moieties (27, 29) will still show high levels of background in samples that have few azide-tagged molecules. Various strategies have been employed to decrease background reactivity in lightly labeled samples: For example, many research groups have discovered the importance of blocking reactive cysteine groups within proteins prior to the SPAAC

reaction (26, 30). In order to continue pushing the limits of azide-tagged protein detection, increasing the specificity of click reactions while maintaining high rates of reactivity will be necessary.

Limitations of the method

Although the SPIQE method indicates whether the conditions of the enrichment enabled successful click and pull-down of azide-tagged molecules, it cannot directly report on whether any particular enriched peptide came from an azide-tagged protein as SILAC does. The simple model peptide may not behave exactly as full-length proteins do in solution. It also cannot be used in samples that have GFP expressed as it is based on a peptide from GFP, which would obfuscate enrichment values. Finally, SPIQE enrichment values >1000 may not have much meaning, since integration of the light peptide at this point is often in the noise regime.

6.6 Future Work

While this peptide is a useful reporter for enrichment, other peptides may be synthesized to compare their enrichment factors. Adding amino acids such as cysteine may be useful to see if peptides containing certain amino acids are enriched more within the system chosen. Furthermore, one could ascertain if the presence of the azide within the peptide chain affects cleavage of the reaction depending on how close it is to the site of cleavage. Using additional peptides in a similar manner would allow investigators to determine the error associated with these enrichment gains.

6.7 Experimental Procedures.

Preparation of SILAC+BONCAT labeled lysates. *E. coli* strain KY14, that is auxotrophic for methionine and lysine **(**KY14**)** was grown in M9 minimal media with 18 amino acids (-Met, -Lys) and either 30:1 azidohomoalanine/methionine (Aha/Met) + Lys0 or Met + Lys8 ($^{13}C_6{}^{15}N_2{}^{1}H_{14}{}^{16}O_2$) (21, 24). The cells were refreshed three times 1:100 after growth to stationary phase into the respective media to ensure quantitative replacement of lysine with either its light or heavy version. After growth to OD600 = 1.0, cultures were spun down for 10 minutes at 3500 rcf at 4 °C. The supernatant was removed and the cells were resuspended in 500 µL 1% SDS in PBS. Benzonase (0.5 µL, 250 units/µL, Sigma) was added for 5 minutes and the cells were boiled at 95 °C for 10 minutes. After determination of protein concentration using Pierce BCA protein quantification kit, 10 µg of this lysate was conjugated to alkyne-TAMRA following the instructions of the Click-It Kit (Thermo Scientific, previously Life Technologies). The clicked samples were mixed in Laemmli sample buffer, run on an SDS-PAGE gel, washed, and visualized using a Typhoon gel scanner to assay tagging (Fig 6.3). After confirming Aha incorporation, all samples were digested with trypsin, then desalted using C18 StageTips as previously described (31) before being subjected to respective LC-MS/MS (detailed below).

MS analysis of unenriched SPIQE peptide mix in complex lysate. The 1:1 SPIQE peptide mix was added to 400 µL of the 1:1 "light + heavy" *E. coli* lysate at either 25 µg peptide/mg of lysate or 2.5 µg peptide/mg of lysate. This unenriched mixture (10 µL) was diluted with 90 µL AmBi to bring the final SDS concentration to 0.1%. These samples

were digested with trypsin overnight at 37 °C and the peptides were extracted using 10% ACN in 10 0mM AmBi. Leftover SDS was removed using HiPPR detergent removal resin (Thermo Scientific). Samples were then desalted using C18 StageTips (31), before being resuspended in 0.1% formic acid (FA). This mixture was analyzed in a targeted manner on a QTRAP 6500 LC-MS/MS system (ABSciex) coupled to an Eksigent ekspert nanoLC 425 pump, ekspert nanoLC400 autosampler and ekspert cHiPLC. Samples were separated using a 45 min linear gradient of acetonitrile in 0.2% FA from 0% to 40% at a flow rate of 300 nL/min on a cHiPLC Chrom XP C18-CL 3 μm trap column, 120 Å (200 μm × 0.5 mm), in line with a cHiPLC Chrom XP C18-CL 3 μm column, 120 Å (75 μm × 150 mm) . The transitions used to monitor the SPIQE peptide are in Table S7.4. The initial generation of multiple reaction monitoring (MRM) transitions and subsequent analysis was performed using Skyline. Up to five singly charged y-ions with an m/z greater than the precursor were selected for monitoring. Collision energies and declustering potentials were calculated using Skyline models for the QTRAP.

Enrichment of peptide from labeled lysate and trypsin digestion. Each sample was diluted with 400 μL of 8 M urea, reduced with 10 mM dithiothreitol (DTT) for 30 min, alkylated with 100 mM iodoacetamide for 45 min, then rotated overnight with 20 μL of washed DBCO-agarose (Click Chemistry Tools). The resin was placed in filter columns and washed with 10 x 5 mL SDS wash buffer (0.8% SDS, 150 mM NaCl, 100 mM Tris, pH 8.0), 10 x 5 mL 8 M urea in 100 mM Tris (pH 8.0), and 10 x 5 mL 20% ACN in water.

Washed resin was resuspended in 100 µL 10% ACN in 100 mM AmBi and 100 ng of trypsin was added for overnight digest at 37 °C.

Dde-alkyne-agarose. We followed the manufacturer's instructions to use Dde-alkyne-agarose (Click Chemistry Tools). After adding the 2x copper catalyst solution to the lysates, SPIQE peptides, and resin, we allowed the click reaction to proceed overnight while rotating end over end. The resin was washed with the same wash buffers used for the DBCO-agarose, then whole proteins were eluted using hydrazine, according to the manufacturer's protocol. These whole proteins were then digested using trypsin, and desalted using C18 StageTips before being loaded on to an Orbitrap Elite MS. We used Skyline to search the raw files and found both the clicked and released peptide and the unclicked version as well as its corresponding methionine version. These values cannot be quantitatively compared since the methionine version does not necessary ionize similarly to the clicked version.

Isolation and preparation of peptides for MS. After removing the supernatant from the overnight digest, the resin was washed three times with 100 µL 10% ACN in 100 mM AmBi to extract peptides and each wash was collected and combined. These extracted peptides were dried on a speed-vac, then desalted using C18 StageTips. The desalted samples were resuspended in 0.1% formic acid prior to injection on the MS.

MS Analysis. We injected 100 ng of the enriched 2.5 µg peptide/mg lysate onto a Proxeon Easy-nLC 1000 (Thermo Fisher Scientific) coupled to an Orbitrap Elite MS (Thermo

Fisher Scientific), equipped with a nano-electrospray ion source (Thermo Fisher Scientific). For LC, solvent A consisted of 97.8% H_2O, 2% ACN, and 0.2% formic acid and solvent B consisted of 19.8% H_2O, 80% ACN, and 0.2% formic acid. The chromatographic separation was achieved using a 60 min elution gradient from 2% to 30% Solvent B at a flow rate of 220 nL/minute. The mass spectrometer was operated in data-dependent mode to switch automatically between MS and MS/MS scans as described by Kalli and Hess (32). Survey full scan mass spectra were acquired in the Orbitrap (400–1600 m/z) with a resolution of 60,000 at 400 m/z. The top 20 most intense ions from the survey scan were isolated and fragmented in the linear ion trap by collisionally-induced dissociation (CID). Precursor ion charge state screening rejected all singly charged and unassigned charge states. The dynamic exclusion list was set with a max retention time of 90 sec and a relative mass window of 15 ppm.

MaxQuant. MaxQuant (v. 1.5.3.30) was used to process the Thermo RAW files and used to search the *E. coli* proteome (UniProt), and an in-house contaminant database (259 sequences) including human keratins and proteases. All default parameters were used, except the multiplicity was set to search for peptides with heavy Lysine (+8.0142), and up to two missed cleavages were allowed. Aha (-4.9863) and L-2,4-diaminobutanoate (-30.9768), a product of reduction of Aha, were specified as variable modifications for Met.

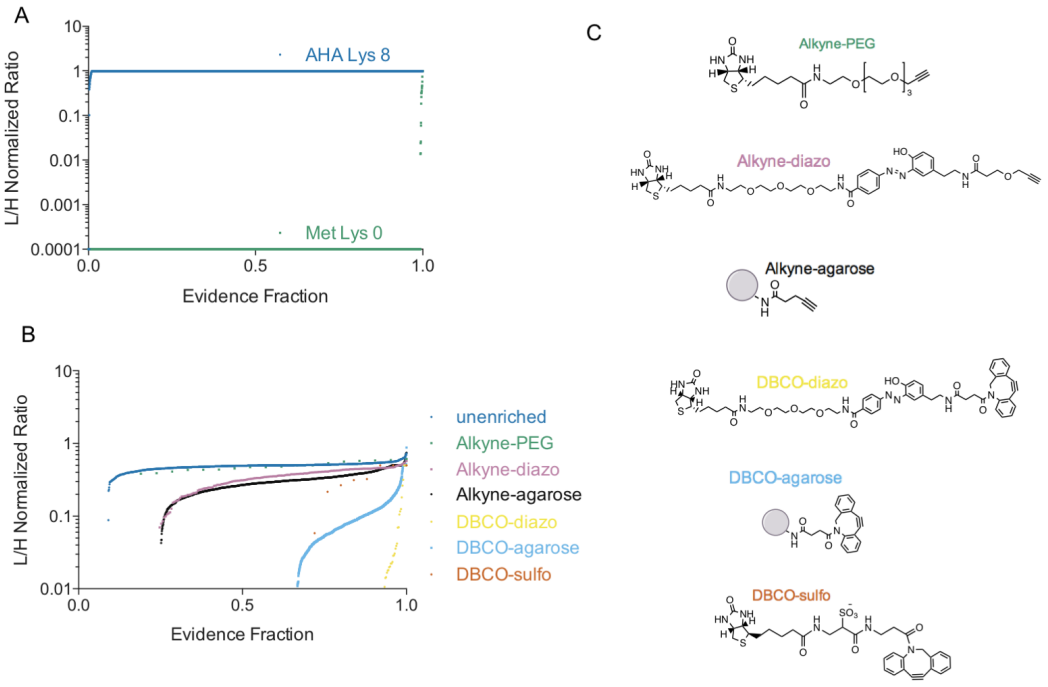


Figure S6.1. A) Prior to mixing, the samples were labeled quantitatively with either 30:1 Met/AHA:Lys8 or Met:Lys0. B) After mixing, the evidences had a median SILAC ratio of 0.5, indicated 1:1 mixing. Decreasing this value indicates successful enrichment. C) Chemical structures of each of the resins or linkers tested using these lysates. If the linker has a biotin handle, then Pierce Streptavidin Agarose resin was incubated with the clicked lysates.

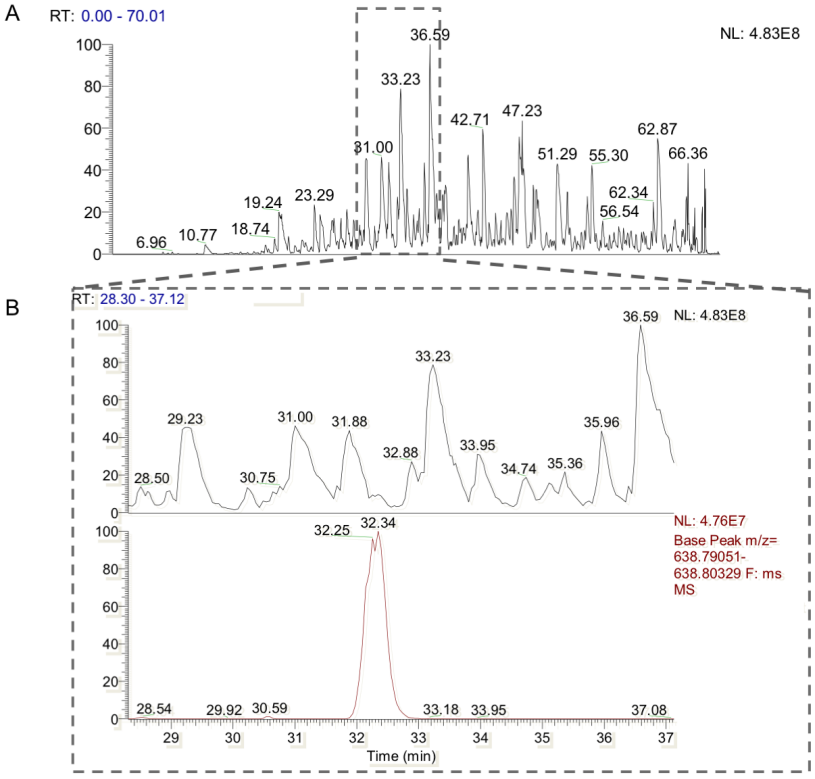
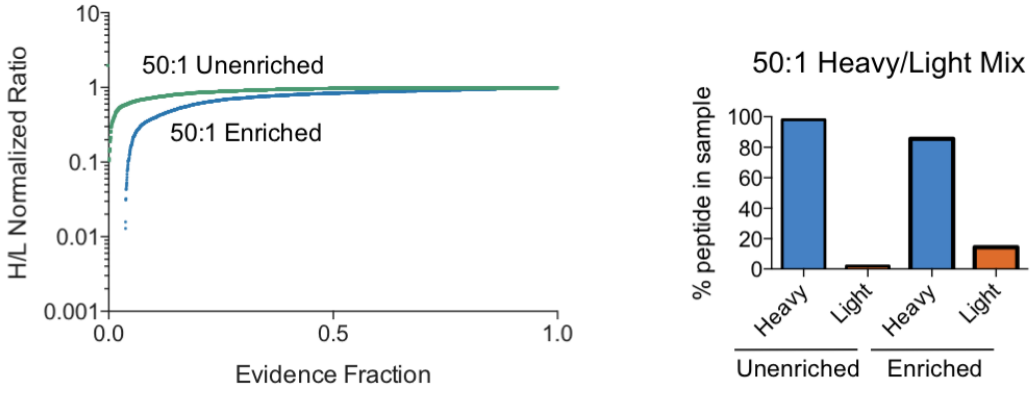


Figure S6.2 A) LC trace of *E. coli* labeled lysates with SPIQE peptide. B) Expanding the LC trace at elution of the SPIQE peptide reveals a small peak in comparison to the rest of the peptides from the lysate.



	#Peptides Found	Median	%Heavy	Enrichment
50:1 Unenriched	3061	0.971	98.1%	-
50:1 Enriched	3840	0.838	85.6%	8.7

Figure S6.3. *E. coli* lysates labeled with Met+Lys8 (heavy) or Aha+Lys0 (light) were mixed 50:1. They were enriched using DBCO-agarose beads as described for the 1:1 lysates in the methods section. After starting with 98.1% heavy labeled lysates, they were brought to 85.6% heavy labeled, which is an enrichment factor of 8.7.

Table S6.4. MRM transitions used to monitor SPIQE peptide on the QTRAP

Q1	Q3	Transition name	CE	DP
786.02	1266.58	Undigested.G[1Ac]GMQHDFFKSAMPEGYVQER.+3y11.light	88.4	40.3
786.02	1179.55	Undigested.G[1Ac]GMQHDFFKSAMPEGYVQER.+3y10.light	88.4	40.3
786.02	1108.51	Undigested.G[1Ac]GMQHDFFKSAMPEGYVQER.+3y9.light	88.4	40.3
786.02	977.47	Undigested.G[1Ac]GMQHDFFKSAMPEGYVQER.+3y8.light	88.4	40.3
786.02	880.42	Undigested.G[1Ac]GMQHDFFKSAMPEGYVQER.+3y7.light	88.4	40.3
793.70	1266.58	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y11.light	89	40.8
793.70	1179.55	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y10.light	89	40.8
793.70	1108.51	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y9.light	89	40.8
793.70	977.47	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y8.light	89	40.8
793.70	880.42	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y7.light	89	40.8
797.04	1276.59	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y11.heavy	89	40.8
797.04	1189.55	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y10.heavy	89	40.8
797.04	1118.52	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y9.heavy	89	40.8
797.04	987.48	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y8.heavy	89	40.8
797.04	890.42	Undigested.G[1Ac]GM[+23]QHDFFKSAMPEGYVQER.+3y7.heavy	89	40.8
554.75	1009.46	Pep1.G[1Ac]GMQHDFFK.+2y8.light	71.6	28.8
554.75	952.43	Pep1.G[1Ac]GMQHDFFK.+2y7.light	71.6	28.8
554.75	821.39	Pep1.G[1Ac]GMQHDFFK.+2y6.light	71.6	28.8
554.75	693.34	Pep1.G[1Ac]GMQHDFFK.+2y5.light	71.6	28.8
554.75	556.28	Pep1.G[1Ac]GMQHDFFK.+2y4.light	71.6	28.8
370.17	952.43	Pep1.G[1Ac]GMQHDFFK.+3y7.light	58.1	17.7
370.17	821.39	Pep1.G[1Ac]GMQHDFFK.+3y6.light	58.1	17.7
370.17	693.34	Pep1.G[1Ac]GMQHDFFK.+3y5.light	58.1	17.7
370.17	556.28	Pep1.G[1Ac]GMQHDFFK.+3y4.light	58.1	17.7
370.17	441.25	Pep1.G[1Ac]GMQHDFFK.+3y3.light	58.1	17.7
566.27	1032.50	Pep1.G[1Ac]GM[+23]QHDFFK.+2y8.light	72.4	29.2
566.27	975.48	Pep1.G[1Ac]GM[+23]QHDFFK.+2y7.light	72.4	29.2
566.27	821.39	Pep1.G[1Ac]GM[+23]QHDFFK.+2y6.light	72.4	29.2
566.27	693.34	Pep1.G[1Ac]GM[+23]QHDFFK.+2y5.light	72.4	29.2
377.85	975.48	Pep1.G[1Ac]GM[+23]QHDFFK.+3y7.light	58.7	18.1
377.85	821.39	Pep1.G[1Ac]GM[+23]QHDFFK.+3y6.light	58.7	18.1
377.85	693.34	Pep1.G[1Ac]GM[+23]QHDFFK.+3y5.light	58.7	18.1
377.85	556.28	Pep1.G[1Ac]GM[+23]QHDFFK.+3y4.light	58.7	18.1
377.85	441.25	Pep1.G[1Ac]GM[+23]QHDFFK.+3y3.light	58.7	18.1
641.79	1124.50	Pep2.SAM[Oxi]PEGYVQER.+2y9.light	77.9	32
641.79	977.47	Pep2.SAM[Oxi]PEGYVQER.+2y8.light	77.9	32
641.79	880.42	Pep2.SAM[Oxi]PEGYVQER.+2y7.light	77.9	32
641.79	751.37	Pep2.SAM[Oxi]PEGYVQER.+2y6.light	77.9	32
641.79	694.35	Pep2.SAM[Oxi]PEGYVQER.+2y5.light	77.9	32

646.79	1134.51	Pep2.SAM[Oxi]PEGYVQER.+2y9.heavy	77.9	32
646.79	987.48	Pep2.SAM[Oxi]PEGYVQER.+2y8.heavy	77.9	32
646.79	890.42	Pep2.SAM[Oxi]PEGYVQER.+2y7.heavy	77.9	32
646.79	761.38	Pep2.SAM[Oxi]PEGYVQER.+2y6.heavy	77.9	32
646.79	704.36	Pep2.SAM[Oxi]PEGYVQER.+2y5.heavy	77.9	32
633.79	1108.51	Pep2.SAMPEGYVQER.+2y9.light	77.3	31.7
633.79	977.47	Pep2.SAMPEGYVQER.+2y8.light	77.3	31.7
633.79	880.42	Pep2.SAMPEGYVQER.+2y7.light	77.3	31.7
633.79	751.37	Pep2.SAMPEGYVQER.+2y6.light	77.3	31.7
633.79	694.35	Pep2.SAMPEGYVQER.+2y5.light	77.3	31.7
638.80	1118.52	Pep2.SAMPEGYVQER.+2y9.heavy	77.3	31.7
638.80	987.48	Pep2.SAMPEGYVQER.+2y8.heavy	77.3	31.7
638.80	890.42	Pep2.SAMPEGYVQER.+2y7.heavy	77.3	31.7
638.80	761.38	Pep2.SAMPEGYVQER.+2y6.heavy	77.3	31.7
638.80	704.36	Pep2.SAMPEGYVQER.+2y5.heavy	77.3	31.7

References

1. Egertson JD, MacLean B, Johnson R, Xuan Y, & MacCoss MJ (2015) Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nat Prot* 10(6):887-903.
2. Duncan MW, Aebersold R, & Caprioli RM (2010) The pros and cons of peptide-centric proteomics. *Nat Biotech* 28(7):659-664.
3. Tyanova S, Temu T, & Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Prot* 11(12):2301-2319.
4. Nilsson T, *et al*. (2010) Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Methods* 7(9):681-685.
5. Yang Y, Fonovic M, & Verhelst SH (2017) Cleavable Linkers in Chemical Proteomics Applications. *Methods Mol Biol* 1491:185-203.
6. Prescher JA & Bertozzi CR (2005) Chemistry in living systems. *Nat Chem Biol* 1(1):13-21.
7. McKay CS & Finn MG (2014) Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chem Biol* 21(9):1075-1101.
8. Tanrikulu IC, Schmitt E, Mechulam Y, Goddard WA, & Tirrell DA (2009) Discovery of *Escherichia coli* methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine *in vivo*. *Proc Natl Acad Sci* 106(36):15285-15290.
9. Dieterich DC, Link AJ, Graumann J, Tirrell DA, & Schuman EM (2006) Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci* 103(25):9482-9487.

10. Yuet KP, *et al*. (2015) Cell-specific proteomic analysis in *Caenorhabditis elegans*. *Proc Natl Acad Sci* 112(9):2705-2710.
11. Eichelbaum K, Winter M, Diaz MB, Herzig S, & Krijgsveld J (2012) Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nat Biotech* 30(10):984-90.
12. Griffin ME, *et al*. (2016) Comprehensive mapping of *O*-GlcNAc modification sites using a chemically cleavable tag. *Mol Biosyst* 12(6):1756-1759.
13. Hahne H, *et al*. (2013) Proteome Wide Purification and Identification of *O*-GlcNAc-Modified Proteins Using Click Chemistry and Mass Spectrometry. *J Proteome Res* 12(2):927-936.
14. Kho Y, *et al*. (2004) A tagging-via-substrate technology for detection and proteomics of farnesylated proteins. *Proc Natl Acad Sci* 101(34):12479-12484.
15. Chuh KN & Pratt MR (2015) Chemical methods for the proteome-wide identification of posttranslationally modified proteins. *Curr Opin Chem Biol* 24:27-37.
16. Speers AE, Adam GC, & Cravatt BF (2003) Activity-based protein profiling *in vivo* using a copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. *J Am Chem Soc* 125(16):4686-4687.
17. Sadaghiani AM, Verhelst SHL, & Bogyo M (2007) Tagging and detection strategies for activity-based proteomics. *Curr Opin Chem Biol* 11(1):20-28.
18. Sibbersen C, Lykke L, Gregersen N, Jorgensen KA, & Johannsen M (2014) A cleavable azide resin for direct click chemistry mediated enrichment of alkyne-labeled proteins. *Chem Commun* 50(81):12098-12100.
19. Mann M (2006) Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Bio* 7(12):952-958.
20. Howden AJM, *et al*. (2013) QuaNCAT: quantitating proteome dynamics in primary cells. *Nat Meth* 10(4):343-6.
21. Bagert JD, *et al*. (2014) Quantitative, Time-Resolved Proteomic Analysis by Combining Bioorthogonal Noncanonical Amino Acid Tagging and Pulsed Stable Isotope Labeling by Amino Acids in Cell Culture. *Mol Cell Prot* 13(5):1352-1358.
22. Adibekian A, *et al*. (2012) Click-generated triazole ureas as ultrapotent *in vivo*-active serine hydrolase inhibitors. *Nat Chem Biol* 8(3):318-318.
23. Ma Y, McClatchy DB, Barkallah S, Wood WW, & Yates JR, 3rd (2017) HILAQ: A Novel Strategy for Newly Synthesized Protein Quantification. *J Proteome Res* 16(6):2213-2220.
24. Link AJ, *et al*. (2006) Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proc Natl Acad Sci* 103(27):10180-10185.
25. MacLean B, *et al*. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26(7):966-968.
26. van Geel R, Pruijn GJM, van Delft FL, & Boelens WC (2012) Preventing Thiol-Yne Addition Improves the Specificity of Strain-Promoted Azide-Alkyne Cycloaddition. *Bioconj Chem* 23(3):392-398.

27. Szychowski J, *et al*. (2010) Cleavable Biotin Probes for Labeling of Biomolecules via Azide-Alkyne Cycloaddition. *J Am Chem Soc* 132(51):18351-18360.
28. Shen WH, *et al*. (2014) Acute Synthesis of CPEB Is Required for Plasticity of Visual Avoidance Behavior in Xenopus. *Cell Rep* 6(4):737-747.
29. Sletten EM & Bertozzi CR (2011) From Mechanism to Mouse: A Tale of Two Bioorthogonal Reactions. *Acc Chem Res* 44(9):666-676.
30. Stockmann H, *et al*. (2011) Development and evaluation of new cyclooctynes for cell surface glycan imaging in cancer cells. *Chem Sci* 2(5):932-936.
31. Rappsilber J, Mann M, & Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Prot* 2(8):1896-1906.
32. Kalli A & Hess S (2012) Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer. *Proteomics* 12(1):21-31.

Appendix A

SUPPORTING INFORMATION FOR CHAPTER III

This Appendix includes:

Supplemental Materials & Methods
Supplementary Figures 1-16
Supplementary Tables 1-5

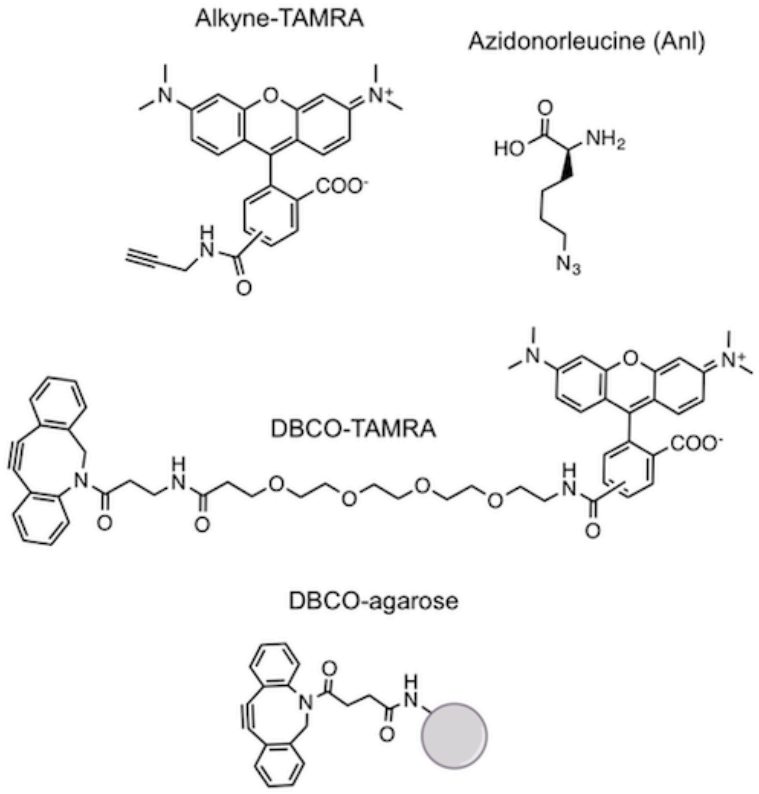
Supplemental Materials & Methods

Histological examination of mouse tissues. Mouse skin was harvested 2 days after inoculation and fixed in 10% neutral-buffered formalin for 48 hours. Fixed tissues were embedded in paraffin, sectioned, mounted on slides, and stained with hematoxylin and eosin (Pacific Pathology).

Keratinocyte culture and infection. Human epidermal keratinocytes from neonatal donors (HEKn) cells (ThermoFisher) were cultured in EpiLife Medium with calcium and human keratinocyte growth supplement (HKGS) with no antibiotics added. At 80% confluence, they were washed with PBS and *S. aureus* strains were added to fresh media at a 10:1 multiplicity of infection (MOI) (MRSA:HEKn). After 90 min, the HEKn cells were washed and lysed with 0.1% SDS. MRSA is resistant to lysis in these conditions and was subsequently diluted and plated on TSA to count the number of adherent and internalized CFUs.

Supplemental Figures

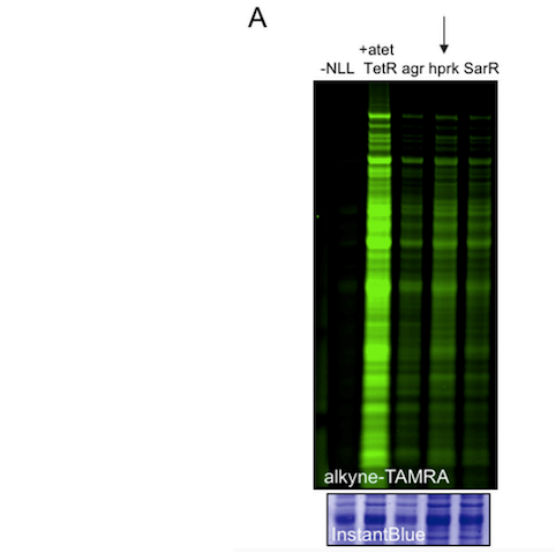
Supplementary Figure S1: Structures of reagents used in the text

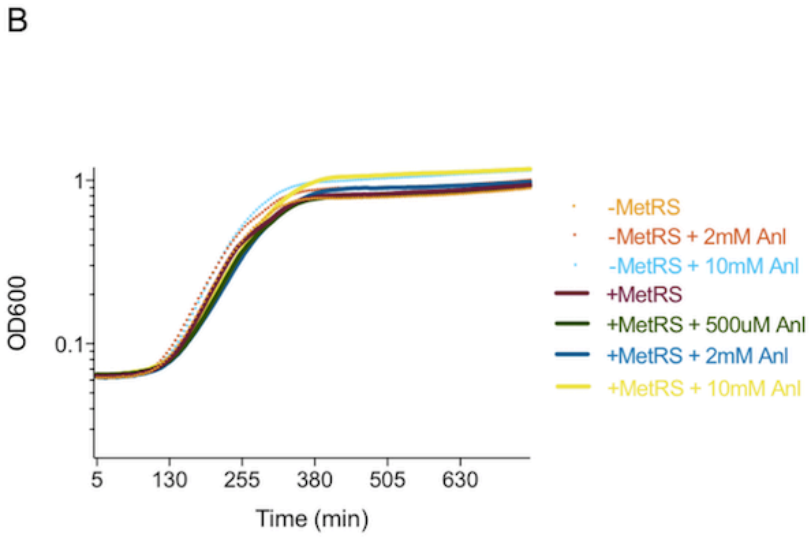


Supplementary Figure S2:

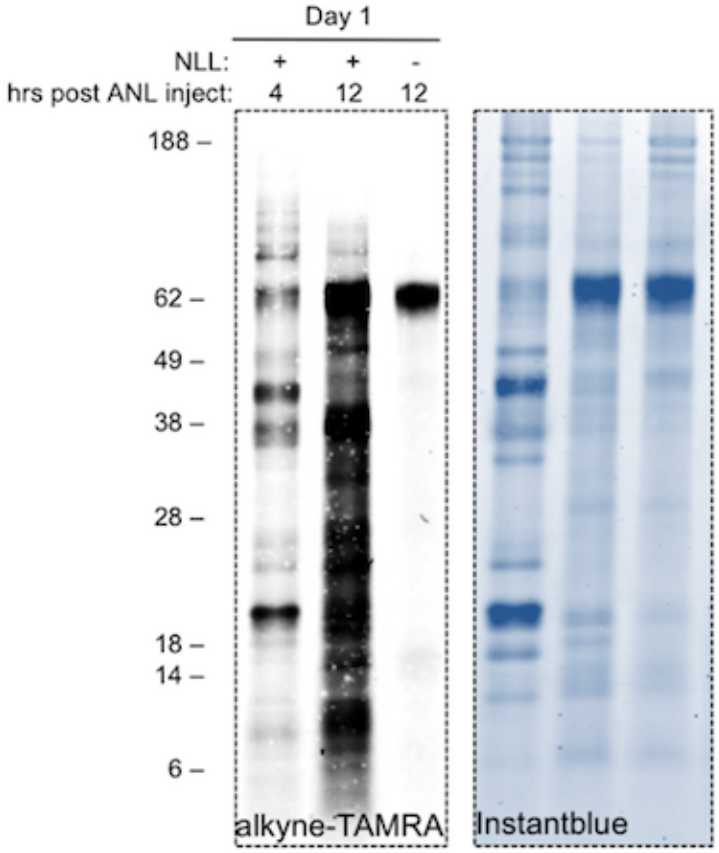
A) In-gel fluorescence reveals the strength of *S. aureus* promoters driving expression of the NLL-MetRS followed by Anl labeling. Selected promoters (~1000bp upstream) were inserted in front of the NLL-MetRS and confirmed via sequencing. A tetracycline-inducible promoter was also tested and induced with 500 ng/μl for 2 hours prior to labeling. The strains were grown in TSB and labeled for 1 hour with 2mM Anl during mid-exponential growth, clicked to alkyne-TAMRA, and ran on a gel. The gel was stained for total protein using Instantblue to confirm equal protein loading. The arrow indicates the selected strain used for all subsequent experiments.

B) Strains (+NLL or –NLL) were diluted 1:500 then grown aerobically in the presence of indicated amounts of Anl in TSB. Growth rate was not significantly affected by addition of Anl or expression of the NLL-MetRS.

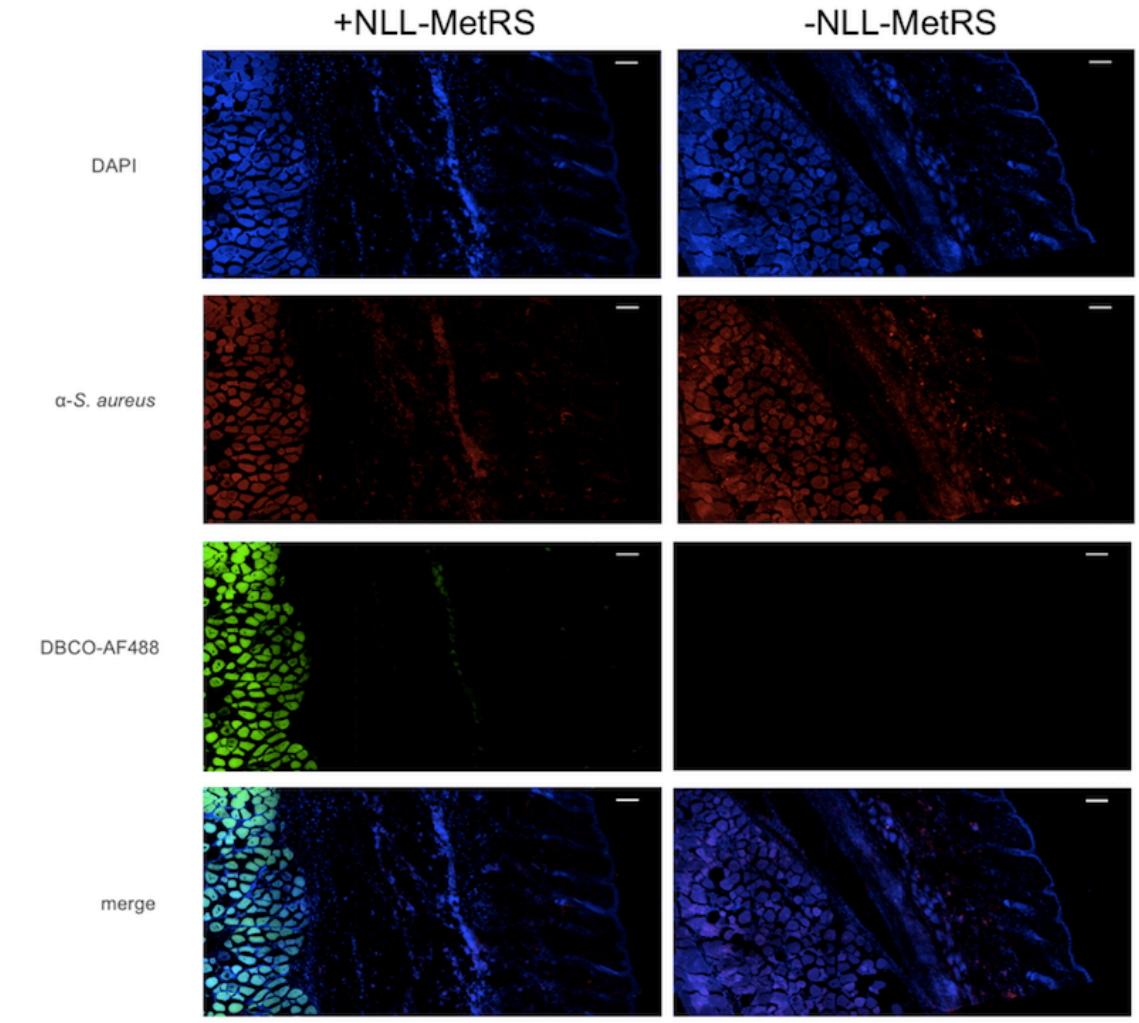




Supplementary Figure S3:
We infected mice subcutaneously with 5e7 CFU of +NLL or –NLL MRSA strains and allowed the infection to proceed for 12 hours. Then, we injected 25mM Anl into the lesion and and allowed cell-specific labeling for the indicated amount of time. Abscesses were excised and homogenized in 1 mL PBS with 1% SDS. The homogenized lysates were clicked to alkyne-TAMRA, ran on a gel, and imaged. The gel was stained with Instantblue to assess total protein loading.



Supplementary Figure S4:
Mice were infected subcutaneously with 5e7 CFU of +NLL or –NLL strains of MRSA and 25 μl of 25 mM Anl was injected at the same time. Abscesses were excised and fixed in 4% PFA before being cryosliced and placed on glass slides. Slices were treated with 50 μg/mL lysostaphin for 45 min at 37 ºC. Samples were blocked with 100 mM iodoacetamide for 15 min then DBCO-AF488 for 10 min. The slides were washed for 1 hour in PBS, blocked with 1% mouse serum in PBS, and treated with 1:3000 dilution of α-*S. aureus* antibody (rabbit polyclonal) in 1% mouse serum in PBS for 1 hour at room temperature. After 3 x 15 min washes in PBS, a goat anti-rabbit secondary antibody conjugated to AF555 was added for 1 hour, followed by 3 more PBS washes. Vectashield with DAPI was added prior to imaging.



Supplemental Figure S5: BONCAT enrichment selects for MRSA proteins

A) Table of spectral counts before and after enrichment with DBCO-agarose. After enrichment we found more spectra matched to MRSA proteins and fewer spectra matched with mouse proteins.

B) Enrichment increases the number of MRSA proteins found with at least 1 peptide.

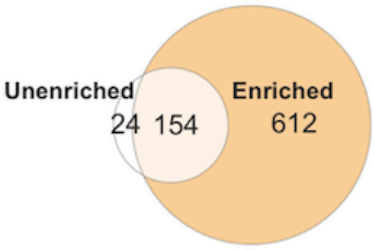
C) After enrichment, we found more spectra per individual protein found.

D) Of the 154 MRSA proteins found both before and after infection, LFQ intensities of MRSA proteins were higher after enrichment. Mouse proteins found in samples before and after enrichment showed a slight decrease in their LFQ intensities.

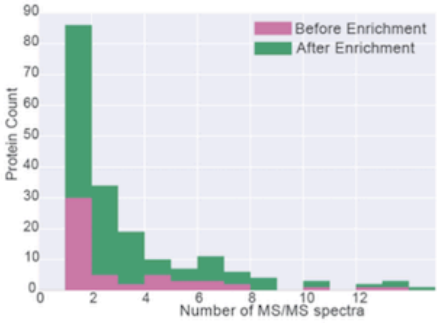
A

	MRSA counts	Mouse counts	%MRSA counts
Unenriched	1142	32627	3.4
Enriched	5095	16832	23.2

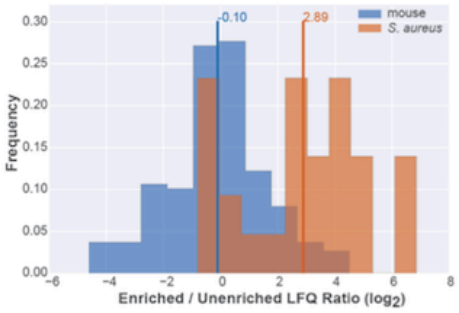
B



C



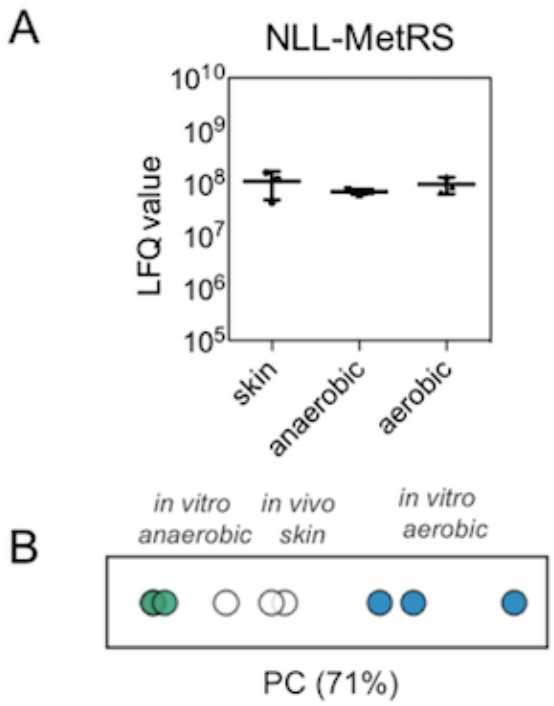
D



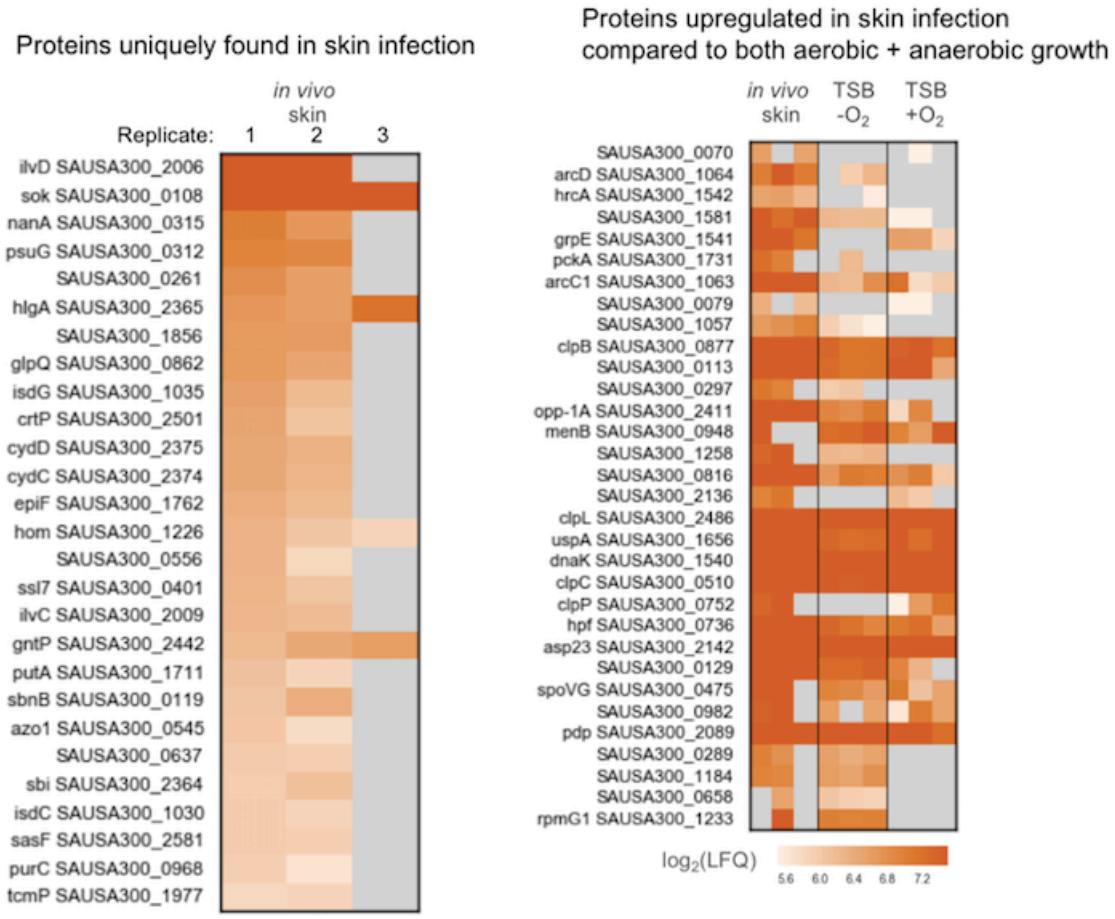
Supplemental Figure S6

A) The NLL-MetRS is expressed in similar levels in all samples as shown by LFQ values.

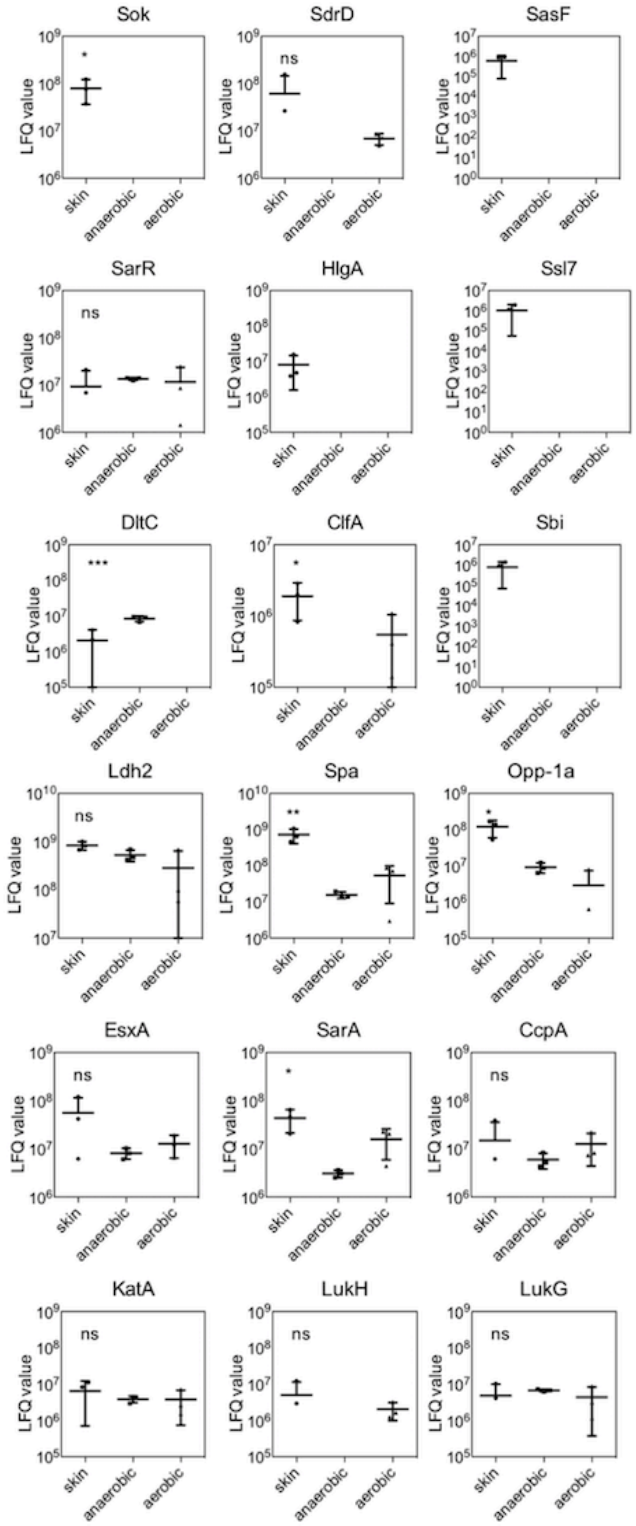
B) One-dimensional principal component analysis (PCA) of the proteins found in aerobically and anaerobically grown samples reveals that skin infection samples do not cluster with either.



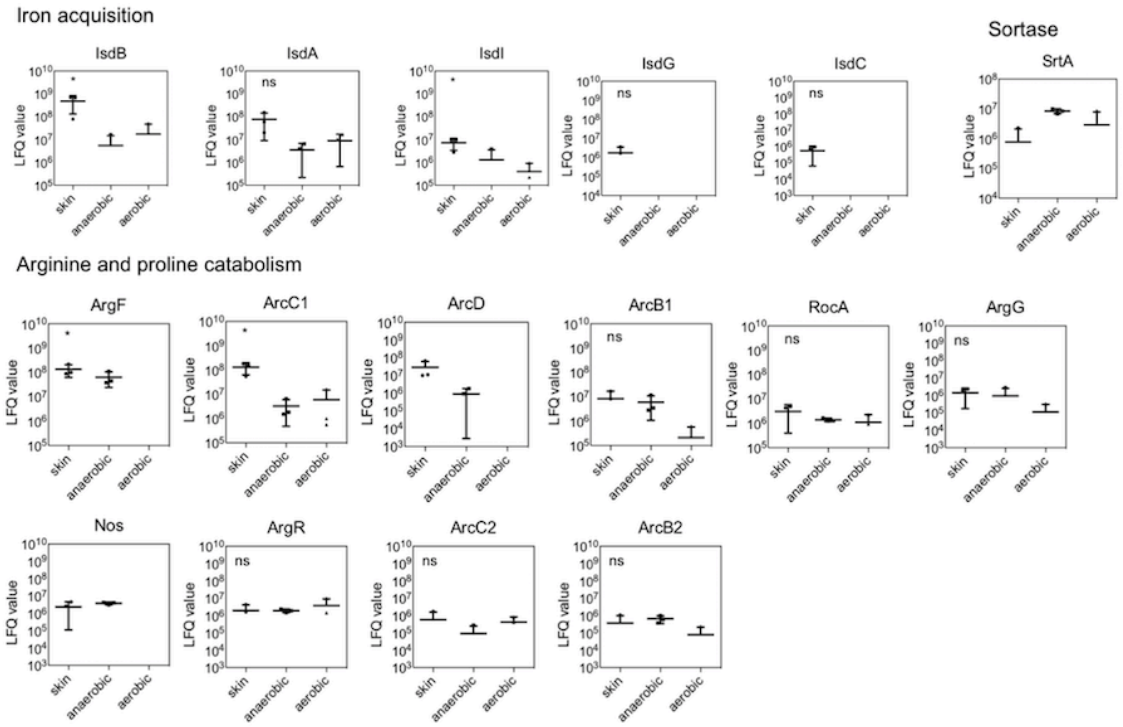
Supplemental Figure S7: Heatmaps of $\log_2$(LFQ) values.



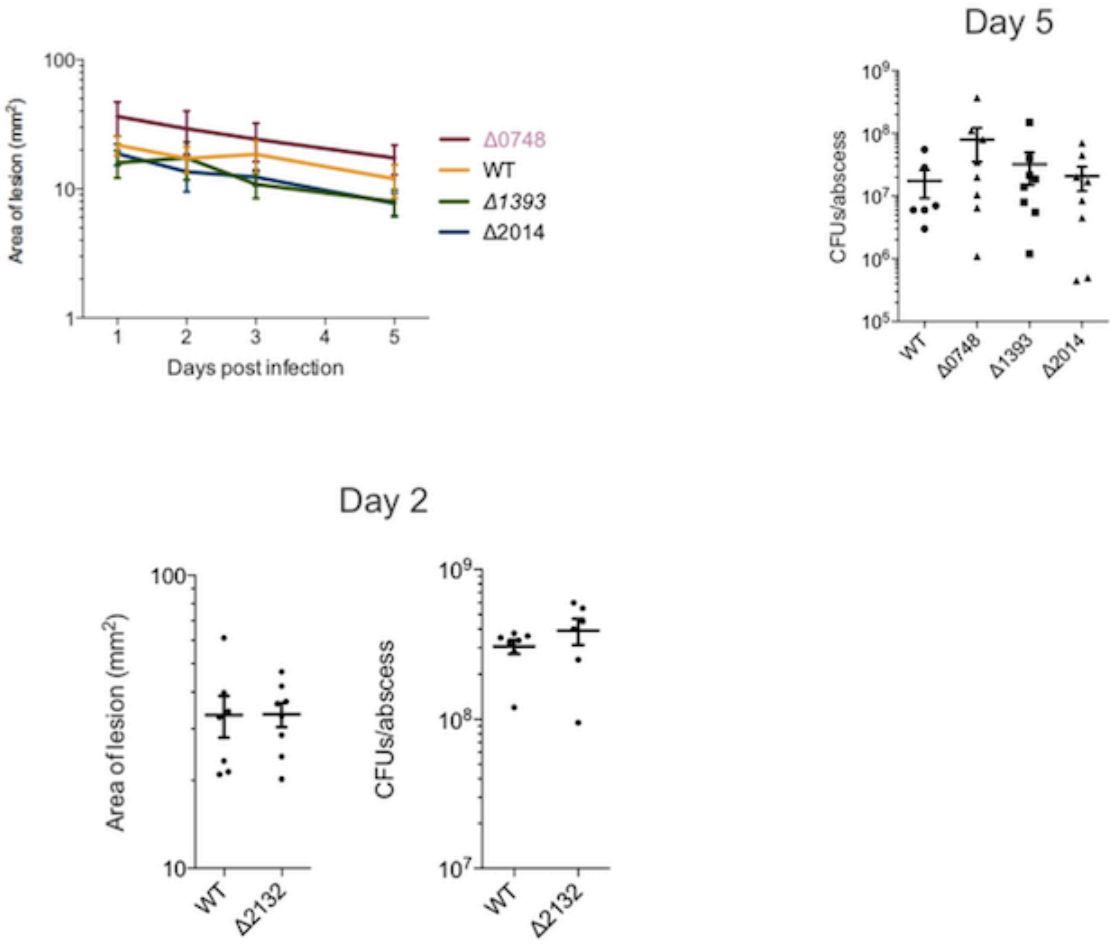
Supplemental Figure S8: LFQ values of proteins annotated as important for pathogenesis/virulence. Some proteins were not found in all samples.



Supplemental Figure S9: LFQ values of proteins annotated as important for iron acquisition or arginine/proline catabolism. Some proteins were not found in all samples.



Supplemental Figure S10: In addition to the mutants mentioned in the main text, we also tested other insertional mutants from the NARSA transposon library in a subcutaneous model of skin infection. There was no statistical differences between wild-type (WT) and the mutants tested in the area of the lesion or the number of CFUs.

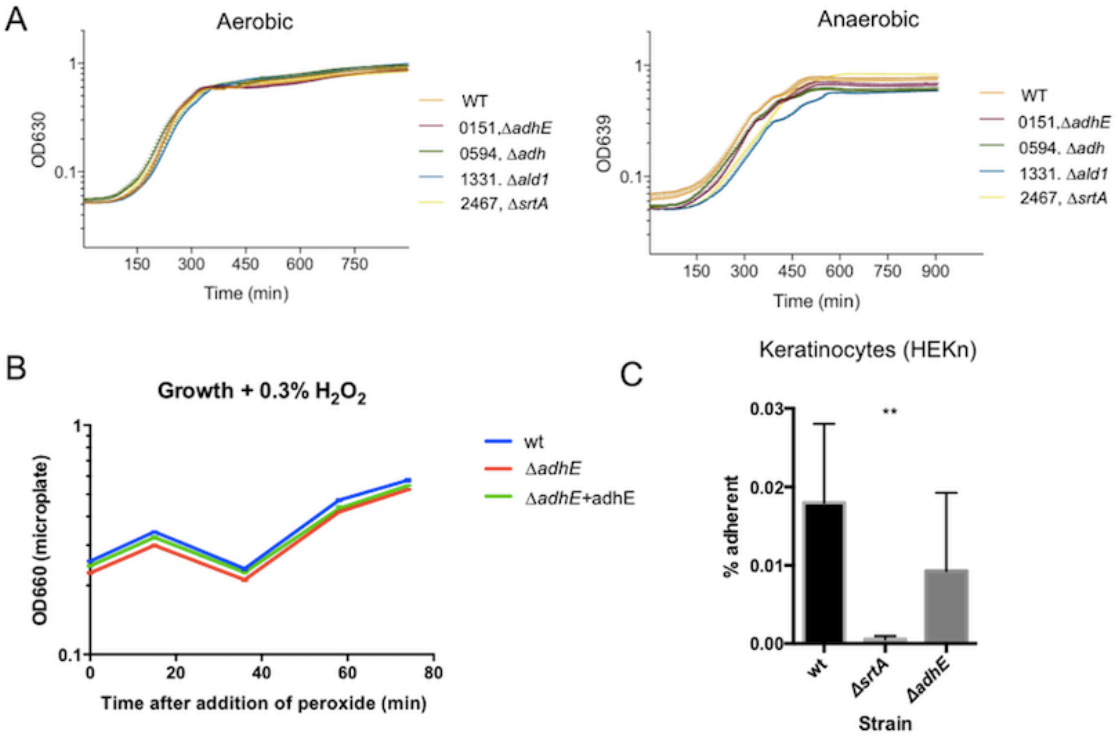


Supplementary Figure S11;

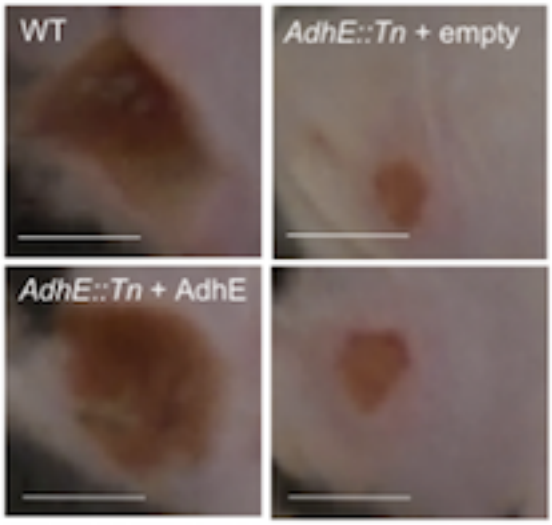
A) No difference in growth of MRSA wild-type, adhE mutant, or other tested mutants during aerobic growth. Some other mtuant strains show slight differences during anaerobic growth.

B) No difference in growth between wild-type, AdhE mutant, or complemented AdhE after exposure to peroxide

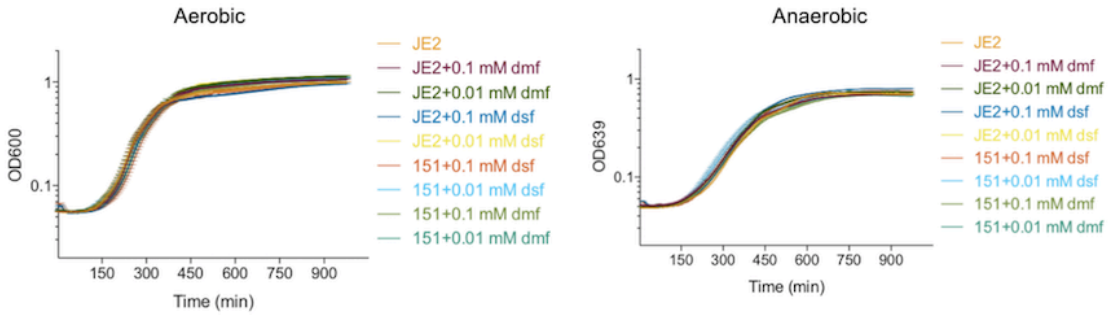
C) No difference in adhesion to HEKn cells, although a sortase mutant (2467, Δ srtA) displayed significant growth defects.



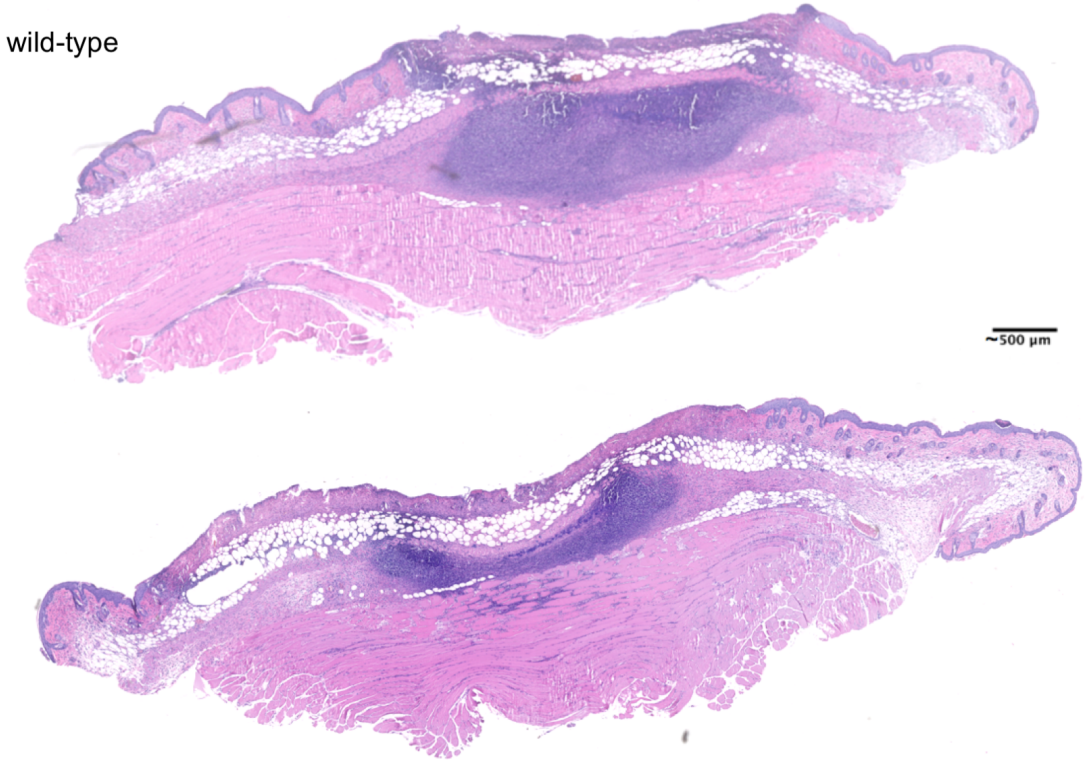
Supplementary Figure S12:
Representative lesions formed over time after subcutaneous injection with 2e7 CFUs of wild-type USA300 compared to Δ*AdhE* mutant complemented with empty vector, full-length AdhE, or AdhE with a single point mutation C258A (Δ*AdhE*+ C258A). Lesions on day 2 of infection. Scale bar is 5 mm.



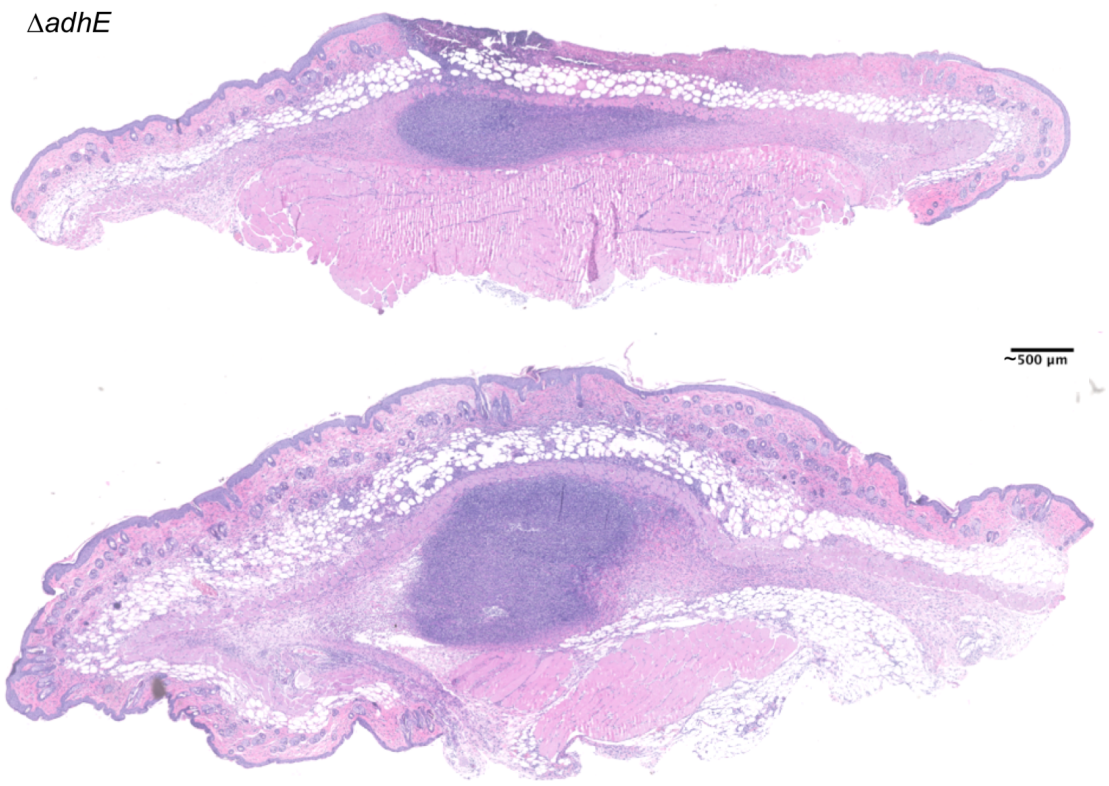
Supplementary Figure S13:
We saw no differences in growth between MRSA USA300 wild-type or Δ*adhE* with disulfiram (DSF) or vehicle (DMF) in TSB.



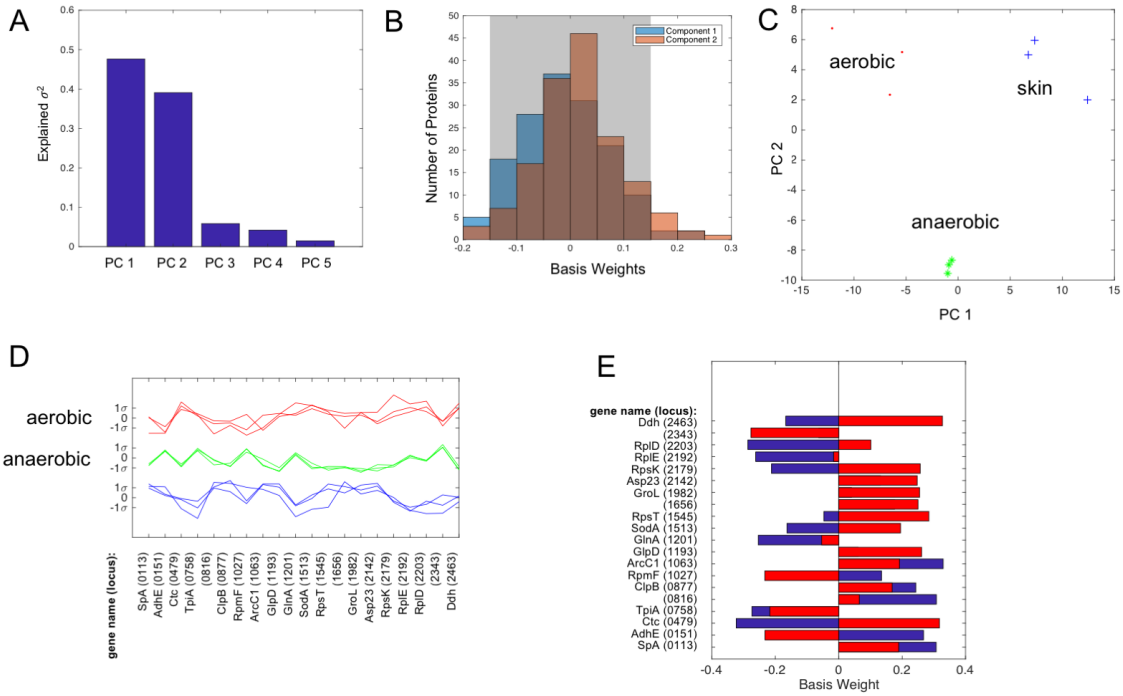
Supplementary Figure S14: H & E stained MRSA abscesses (wild-type) on Day 2 of infection



Supplementary Figure S15: H & E of MRSA abscesses (Δ*adhE*) on Day 2 of infection



Supplementary Figure S16: Further insight into infection using principal component analysis (PCA). A) The first two components of this 20-protein-PCA explains most of the variance. B) Only 20 proteins in principal components 1 and 2 have magnitudes above 0.15 (unshaded part of the histogram). C) PCA on only these 20 proteins reveals they cluster similarly to PCA on the entire data set. D) Expression profiles of these 20 proteins in each of the 3 conditions. E) Relative weight contributed by each protein to the first two principle components in the 20-protein-PCA. Blue = PC1, Red = PC2.



Supplemental Table 1. Strains used in Chapter 3.

Name	Organism	Resist-ance	Notes	Reference / Source
DH10B	*E. coli*	-	Strain for cloning and plasmid maintenance	NEB
RN4220	*S. aureus*	-	Restriction-deficient cloning strain	Olaf Schneewind
USA300 (JE2)	*S. aureus*	-	Base strain for Tn insertions	NARSA
CYL12349	*S. aureus* RN4220	Cm	RN4220 with engineered attB2 site,+ plasmid pYL112Δ19	31
CYL12376	*S. aureus* RN4220	Tet	CYL12349, pLL102	31
R_LLadhE	*S. aureus* RN4220	Tet/Cm	CYL12349, integrated pLL102_adhE	This work
R_LLadhE_C258A	*S. aureus* RN4220	Tet/Cm	CYL12349, integrated pLL102_adhE_C258A	This work
R_LL102	*S. aureus* RN4220	Tet/Cm	CYL12349, integrated pLL102_empty	This work
151_LLadhE	*S. aureus* JE2	Erm/Tet	NE114, integrated pLL102_adhE	This work
151_LLadhE_C258A	*S. aureus* JE2	Erm/Tet	NE114, integrated pLL102_adhE_C258A	This work
151_LL102	*S. aureus* JE2	Erm/Tet	NE114, integrated pLL102_empty	This work
JE2_LL102	*S. aureus* JE2	Tet	JE2, integrated pLL102_empty	This work
R_20_empty	*S. aureus* RN4220	Cm	pSS20_noMetRS	This work
R_20_tetR	*S. aureus* RN4220	Cm	pSS20_tetR+NLL_MetRS	This work
R_20_hprK	*S. aureus* RN4220	Cm	pSS20_hprk+NLL_MetRS	This work
R_20_agr	*S. aureus* RN4220	Cm	pSS20_agr+NLL_MetRS	This work

R_20_sarR	*S. aureus* RN4220	Cm	pSS20_sarR+NLL_MetRS	This work
JE2_20_empty	*S. aureus* JE2	Cm	pSS20_noMetRS	This work
JE2_20_tetR	*S. aureus* JE2	Cm	pSS20_tetR+NLL_MetRS	This work
JE2_20_hprk	*S. aureus* JE2	Cm	pSS20_hprk+NLL_MetRS	This work
JE2_20_agr	*S. aureus* JE2	Cm	pSS20_agr+NLL_MetRS	This work
JE2_20_sarR	*S. aureus* JE2	Cm	pSS20_sarR+NLL_MetRS	This work
NE1382	*S. aureus* JE2	Erm	Tn: SAUSA300_0594	NARSA (BEI Resources)
NE114	*S. aureus* JE2	Erm	Tn: SAUSA300_0151	NARSA (BEI Resources)
NE515	*S. aureus* JE2	Erm	Tn: SAUSA300_1330	NARSA (BEI Resources)
NE1136	*S. aureus* JE2	Erm	Tn: SAUSA300_1331	NARSA (BEI Resources)
NE1673	*S. aureus* JE2	Erm	Tn: SAUSA300_2132	NARSA (BEI Resources)
NE619	*S. aureus* JE2	Erm	Tn: SAUSA300_0307	NARSA (BEI Resources)
NE1039	*S. aureus* JE2	Erm	Tn: SAUSA300_0312	NARSA (BEI Resources)
NE1835	*S. aureus* JE2	Erm	Tn: SAUSA300_0315	NARSA (BEI Resources)
NE54	*S. aureus* JE2	Erm	Tn: SAUSA300_2375	NARSA (BEI Resources)
NE1787	*S. aureus* JE2	Erm	Tn: SAUSA300_2467	NARSA (BEI Resources)
NE47	*S. aureus* JE2	Erm	Tn: SAUSA300_1393	NARSA (BEI Resources)
Ec_20_tetR	*E. coli* DH10B	Amp	pSS20_tetR	This work
Ec_20_hprK	*E. coli* DH10B	Amp	pSS20_hprk	This work
Ec_20_agr	*E. coli* DH10B	Amp	pSS20_agr	This work

Ec_20_sarR	*E. coli* DH10B	Amp	pSS20_sarR	This work
Ec_102_adhE	*E. coli* DH10B	Spec	pLL102_adhE	This work
Ec_102_adhE_C258A	*E. coli DH10B*	Spec	pLL102_adhE_C258A	This work
Ec_102_empty	*E. coli DH10B*	Spec	pLL102_empty	This work

Supplemental Table 2. Plasmids used in Chapter 3.

Name	Organism	Resistance	Notes	Reference Lab
pmRFPmars	*E coli* / *S aureus*	Amp, Cm	tetracycline-inducible, mRFP, pALC2084 backbone, Addgene	Addgene plasmid # 26252
pWW412	*E coli* / *S aureus*	Cm	Empty complementation plasmid	Schneewind O
pSS20_empty	*E coli* / *S aureus*	Cm	pmRFPmars with mRFP removed	this work
pSS20_tet	*E coli* / *S aureus*	Amp, Cm	pmRFPmars with codon-optimized Ec NLL-MetRS inserted	this work
pSS20_hprk	*E coli* / *S aureus*	Amp, Cm	pmRFPmars with codon-optimized *Ec* NLL-MetRS inserted and hprK promoter instead of tetracycline	this work
pLL102	*E coli* / *S aureus*	Spec/Tet	chromosomal integration vector for use with pYL112Δ19	Lee CY
pYL112Δ19	*S. aureus*	Cm	integrase strain	Lee CY

Supplemental Table 3. Primers used in Chapter 3.

Name	Sequence
Cloning MetRS	
pSS20_insert_F	CCAGATCCTAGGAAAGGAGGATGATAAAATGAGAGGATCGCATCACCATCAC
pSS20_insert_R	GCAGAGAAACCTTGTTTTACAAGGTTTATTTAACTTGATGACCTGGTTTAGCAC
pSS20_vector_F	GTGCTAAACCAGGTCATCAAGTTAAATAAACCTTGTAAAACAAGGTTTCTCTGC
pSS20_vector_R	GTGATGGTGATGCGATCCTCTCATTTTATCATCCTCCTTTCCTAGGATCTGG
EcoRI_metrs_R	TAATGTAGAATTCGGCTTAACTGCAGAGA
SalI_hprk_F	ATTACAGTCGACGTGAGGCTCAACGTGTGAAA
KpnI_hprk_R	TGCTTGGTACCGTCCTTCTTCACCAGCGATT
Complementing AdhE and making mutants	
pLL_adhE_R	TGTCGATACCTATCAACAGATGCCGACAGTAAGACGGGTAAG
pLL_adhE_F	TGCCACATAAACTTTAAGTCGATAGAATTCGAGCTCGGTACCCG
adhE_F_pLL	CTTACCCGTCTTACTGTCGGCATCTGTTGATAGGTATCGACA
adhE_R_pLL	CGGGTACCGAGCTCGAATTCTATCGACTTAAAGTTTATGTGGCA
AdhE_C258A_F	TGCTTCTGAACAAGTTGTAGTC
AdhE_C258A_R	GCAATCATACCATTATCAAATGTTTTTG
Testing Tn insertions (with corresponding "Buster" or "Upstream" primer)(96)	
Tn_nanA_seq	GCATCAATGCCACGATGACGTA
Tn_2374_seq	TCTAAACGTTCTTTCGCACTCA
Tn_2375_seq	AGGGAGACGCATTAATTTTTGA
Tn_psuG_seq	GAACAATCAACCGATTGTAGCA
Tn_tdcB_seq	GGTGTTCGACGAATAAATGGTT
Tn_ald1_seq	TGGCGCTAATAATTCTGCTTTT
Tn_adh_seq	TTGGTGATGTTACAGGCGTTAC
Tn_1393_seq	CTGCTTCTGTACTGCGTCA
Tn_0748_seq	GGTTTATCTGGCGCAGGTAA
Tn_2132_seq	CATGCAGTTCCTAAATTAGACAGC
Tn_0307_seq	AGCAGAACTTTTGGCAAACG

Tn_adhE_seq	CAGCACACATTAAACGTGCAG
Tn_srtA_seq	GAACTCCTACATCTGTAGGCT

Testing chromosomal insertion of pLL102

OU9R10	CATACTACATATCAACGAAATCAG
scv4	ACCCAGTTTGTAATTCCAGGAG
OU9R7	ATGGGTGGTAAAACACAAATTTC
scv8	GCACATAATTGCTCACAGCCA

Supplemental Table 4: Log2 LFQ Values from skin

gene	lfq_skin 1	lfq_skin 2	lfq_skin 3	lfq_anae r1	lfq_anae r2	lfq_anae r3	lfq_aer 1	lfq_aer 2	lfq_aer 3
SAUSA300_pUSA010 001	NA	NA	NA	21.0	21.1	22.4	NA	NA	NA
NLL-MetRS	26.8	25.3	27.2	25.8	26.0	26.1	26.4	27.0	26.0
dnaA SAUSA300_0001	21.1	20.9	NA	24.2	23.2	24.2	21.6	24.0	NA
dnaN SAUSA300_0002	27.0	26.5	25.8	25.9	25.6	26.7	24.9	24.6	NA
SAUSA300_0003	23.2	23.4	NA	22.1	21.8	20.7	21.2	22.2	21.4
recF SAUSA300_0004	NA	NA	NA	18.9	18.6	19.2	NA	18.3	NA
gyrB SAUSA300_0005	22.6	22.9	22.5	25.0	24.7	24.6	21.7	24.8	22.6
gyrA SAUSA300_0006	NA	NA	NA	25.3	25.4	25.3	26.1	25.5	25.1
nnrD SAUSA300_0007	NA	NA	NA	23.3	23.4	23.4	NA	19.4	NA
serS SAUSA300_0009	25.6	28.4	26.7	27.3	27.8	28.4	25.3	26.8	25.9
SAUSA300_0014	NA	NA	NA	22.2	21.7	22.0	20.0	23.5	NA
rplI SAUSA300_0015	23.7	NA	22.2	25.8	26.1	26.5	26.4	25.1	24.4
dnaB SAUSA300_0016	NA	NA	NA	23.7	23.5	23.7	21.9	21.6	NA
purA SAUSA300_0017	NA	NA	NA	20.1	NA	20.7	19.5	23.4	19.9
walR SAUSA300_0020	NA	NA	NA	23.2	23.5	24.8	NA	NA	NA
walK SAUSA300_0021	NA	NA	NA	20.6	21.0	19.5	NA	NA	NA
SAUSA300_0022	NA	NA	NA	18.6	18.0	NA	NA	NA	NA
SAUSA300_0023	NA	NA	NA	19.6	19.1	19.8	NA	19.5	NA
adsA SAUSA300_0025	22.1	NA	NA	22.4	22.8	22.5	21.9	22.8	21.2
rlmH SAUSA300_0026	NA	NA	NA	20.5	20.1	19.9	NA	NA	NA
SAUSA300_0027	NA	NA	NA	20.4	20.9	21.1	NA	NA	NA
mecA SAUSA300_0032	23.7	24.4	NA	24.2	24.2	24.9	22.4	24.9	22.5
SAUSA300_0042	NA	NA	NA	21.7	21.4	21.9	NA	21.7	NA
SAUSA300_0044	NA	23.2	NA	NA	NA	NA	NA	NA	NA
arcB argF SAUSA300_0062	20.1	NA	NA	19.2	18.5	20.0	17.8	NA	NA
SAUSA300_0067	22.0	25.4	NA	22.5	22.5	21.8	NA	23.6	NA
SAUSA300_0070	21.9	NA	21.7	NA	NA	NA	NA	17.6	NA
SAUSA300_0073	NA	NA	NA	NA	NA	NA	NA	20.7	NA
copB SAUSA300_0078	NA	NA	NA	21.6	21.1	21.9	NA	NA	NA
SAUSA300_0079	21.2	NA	20.6	NA	NA	NA	17.2	17.0	NA
SAUSA300_0089	NA	NA	19.9	21.2	20.9	20.6	21.0	23.7	NA
SAUSA300_0093	NA	NA	NA	NA	NA	NA	NA	19.2	NA
SAUSA300_0095	NA	NA	NA	17.7	17.6	NA	16.9	NA	NA
plc SAUSA300_0099	NA	NA	NA	NA	NA	16.4	NA	NA	NA
SAUSA300_0105	NA	NA	NA	NA	18.3	19.9	NA	NA	NA
SAUSA300_0107	NA	NA	NA	NA	NA	NA	NA	17.1	NA
SAUSA300_0108	26.9	26.2	25.2	NA	NA	NA	NA	NA	NA
SAUSA300_0113	30.0	28.8	29.3	24.2	23.7	23.7	26.4	26.1	21.5
sarS SAUSA300_0114	NA	NA	NA	21.7	22.0	22.4	21.7	24.4	NA
sirA SAUSA300_0117	24.6	23.9	23.2	NA	NA	NA	22.4	22.1	NA
SAUSA300_0119	20.2	21.2	NA	NA	NA	NA	NA	NA	NA
SAUSA300_0128	NA	NA	NA	20.6	20.4	NA	NA	21.1	NA
SAUSA300_0129	26.4	25.9	26.2	24.2	24.2	24.6	23.1	21.0	NA
sodM SAUSA300_0135	19.9	20.6	NA	21.8	22.6	21.9	19.2	21.7	NA
sasD SAUSA300_0136	22.6	23.5	21.6	21.1	21.3	NA	22.2	18.5	NA
SAUSA300_0137	NA	NA	NA	NA	NA	19.4	NA	NA	NA

deoD SAUSA300_0138	NA	NA	NA	19.4	19.1	20.3	NA	NA	NA
deoC SAUSA300_0140	NA	21.5	NA	24.0	24.0	24.4	NA	NA	NA
deoB SAUSA300_0141	24.0	24.5	26.0	25.1	25.3	25.5	23.7	21.5	26.0
SAUSA300_0145	NA	NA	NA	NA	NA	NA	NA	15.7	NA
adhE SAUSA300_0151	29.7	29.4	31.7	30.6	30.9	30.6	25.0	26.5	24.7
isdI SAUSA300_0168	23.2	23.2	21.4	NA	NA	21.9	19.9	17.8	NA
SAUSA300_0172	NA	NA	NA	NA	NA	19.4	NA	NA	NA
SAUSA300_0173	NA	NA	NA	24.1	23.7	24.3	25.2	26.1	24.5
SAUSA300_0175	NA	NA	NA	NA	NA	NA	NA	18.2	NA
SAUSA300_0177	NA	NA	NA	19.3	19.1	19.7	21.7	23.2	21.9
SAUSA300_0181	NA	23.0	NA	NA	NA	NA	NA	NA	NA
ipdC SAUSA300_0190	21.9	20.7	NA	22.1	21.4	22.1	NA	20.9	NA
ptsG glcA SAUSA300_0191	24.0	23.9	NA	23.8	23.5	23.6	22.8	24.1	NA
hsdR SAUSA300_0196	NA	NA	NA	21.4	21.5	20.8	21.0	22.4	NA
azoR SAUSA300_0206	NA	NA	NA	18.8	19.3	19.2	NA	NA	NA
SAUSA300_0212	NA	20.9	19.3	21.6	21.4	21.3	19.8	20.5	18.2
SAUSA300_0214	NA	NA	NA	20.4	NA	NA	NA	NA	NA
SAUSA300_0215	NA	NA	NA	NA	NA	NA	19.4	NA	NA
SAUSA300_0217	NA	NA	NA	NA	NA	NA	NA	18.4	NA
pflB SAUSA300_0220	29.4	32.1	32.4	31.4	31.5	31.6	29.7	27.9	31.5
pflA SAUSA300_0221	25.4	24.2	NA	24.0	24.1	24.3	NA	23.5	NA
SAUSA300_0230	NA	NA	NA	21.4	21.2	20.7	NA	21.6	NA
SAUSA300_0231	NA	NA	NA	21.1	21.2	21.8	NA	20.0	NA
SAUSA300_0234	NA	21.4	NA	26.0	26.5	26.4	21.4	26.4	NA
ldh1 SAUSA300_0235	26.4	26.4	25.9	30.2	29.7	29.3	29.2	29.2	30.4
SAUSA300_0236	NA	NA	NA	19.2	19.5	19.0	NA	NA	NA
tarI SAUSA300_0245	NA	NA	NA	NA	NA	19.7	NA	NA	NA
SAUSA300_0248	NA	NA	NA	20.9	20.6	21.8	NA	17.2	NA
tarI1 SAUSA300_0249	24.6	24.3	NA	25.2	25.5	25.6	23.4	26.2	24.1
tarJ SAUSA300_0250	NA	NA	NA	23.1	23.0	22.8	23.8	22.4	NA
SAUSA300_0251	NA	NA	23.5	23.7	24.4	24.7	NA	24.5	22.3
SAUSA300_0252	23.7	23.8	23.8	24.1	24.2	24.3	23.5	23.8	NA
scdA SAUSA300_0253	24.5	23.4	21.6	24.2	23.6	25.7	22.2	21.4	22.2
lytR SAUSA300_0255	NA	NA	NA	21.0	21.4	21.3	NA	NA	NA
lrgA SAUSA300_0256	NA	NA	NA	19.8	19.1	19.2	19.6	NA	NA
SAUSA300_0261	22.6	21.8	NA	NA	NA	NA	NA	NA	NA
rbsK SAUSA300_0262	NA	NA	NA	19.8	20.6	20.7	20.0	20.8	NA
SAUSA300_0266	NA	NA	NA	21.9	21.3	22.2	NA	NA	NA
SAUSA300_0271	NA	NA	NA	NA	NA	NA	NA	18.1	NA
SAUSA300_0274	22.0	21.3	NA	21.5	21.2	21.5	24.3	24.1	26.0
esxA SAUSA300_0278	26.9	25.3	22.6	23.0	22.5	23.3	23.6	22.6	24.2
SAUSA300_0279	NA	NA	NA	21.7	21.6	21.1	21.7	22.6	NA
essB SAUSA300_0282	19.7	NA	NA	17.9	NA	17.6	19.0	19.7	NA
SAUSA300_0283	NA	NA	19.1	NA	NA	NA	NA	18.6	NA
esxB SAUSA300_0285	NA	NA	NA	NA	NA	NA	NA	18.1	NA
essD esaD SAUSA300_0288	20.6	NA	NA	NA	NA	NA	NA	18.1	NA
SAUSA300_0289	23.2	22.5	NA	21.8	21.3	21.8	NA	NA	NA
SAUSA300_0297	23.6	23.1	NA	19.9	20.2	NA	NA	NA	NA
SAUSA300_0303	NA	NA	NA	NA	NA	NA	NA	18.0	NA
SAUSA300_0307	24.0	23.8	23.3	22.8	21.5	22.9	23.1	23.4	20.6
SAUSA300_0308	NA	NA	NA	20.1	19.4	NA	NA	NA	NA
psuG SAUSA300_0312	23.1	22.9	NA	NA	NA	NA	NA	NA	NA
nanA	23.3	22.2	NA	NA	NA	NA	NA	NA	NA

SAUSA300_0315									
SAUSA300_0316	18.5	NA	NA	NA	NA	NA	NA	NA	NA
SAUSA300_0329	NA	NA	NA	21.5	21.2	20.3	NA	19.9	NA
SAUSA300_0331	NA	NA	NA	NA	NA	21.6	NA	NA	NA
SAUSA300_0345	NA	NA	NA	NA	NA	NA	17.3	20.2	NA
SAUSA300_0352	18.6	NA	NA	NA	NA	NA	NA	NA	NA
SAUSA300_0355	23.9	23.1	NA	22.9	23.1	23.0	26.4	24.2	24.8
SAUSA300_0362	NA	NA	NA	17.0	NA	NA	NA	19.8	17.8
SAUSA300_0363	NA	NA	NA	18.4	18.0	18.3	NA	18.7	NA
ychF SAUSA300_0364	20.4	20.0	NA	23.7	24.4	23.8	26.4	24.5	26.1
rpsF SAUSA300_0366	NA	NA	NA	29.2	29.1	28.8	27.1	27.8	27.9
ssb SAUSA300_0367	NA	22.6	NA	24.3	24.2	25.1	25.0	25.4	23.4
rpsR SAUSA300_0368	NA	NA	NA	19.3	19.1	19.3	18.4	23.9	NA
SAUSA300_0372	24.2	24.1	20.9	22.0	21.0	21.1	22.2	24.3	21.3
SAUSA300_0373	NA	NA	NA	21.0	20.1	21.0	NA	20.0	NA
SAUSA300_0377	21.2	21.1	NA	21.0	20.9	21.7	21.3	22.5	20.6
ahpF SAUSA300_0379	23.6	25.6	NA	24.3	24.0	24.7	22.8	23.8	22.9
ahpC SAUSA300_0380	26.4	27.0	29.7	26.5	26.1	26.2	28.4	23.2	25.2
SAUSA300_0382	NA	NA	NA	NA	NA	NA	19.8	18.3	NA
SAUSA300_0383	25.3	24.6	28.6	24.9	24.9	24.4	25.7	25.3	23.4
SAUSA300_0385	25.0	24.2	NA	NA	NA	21.1	NA	22.2	NA
xpt SAUSA300_0386	21.7	NA	NA	26.1	25.5	25.7	NA	22.4	NA
guaB SAUSA300_0388	26.8	26.7	27.4	28.2	27.8	27.8	28.8	28.7	27.1
guaA SAUSA300_0389	25.5	25.0	24.6	26.5	25.4	25.2	26.4	25.8	23.8
SAUSA300_0394	NA	NA	NA	21.6	21.9	22.8	NA	NA	NA
SAUSA300_0401	20.9	20.2	NA	NA	NA	NA	NA	NA	NA
SAUSA300_0419	NA	NA	NA	21.2	20.5	21.0	20.9	20.6	21.2
SAUSA300_0420	19.5	NA	NA	20.1	20.7	20.4	NA	18.9	NA
SAUSA300_0421	NA	NA	NA	19.0	18.9	20.4	NA	NA	NA
SAUSA300_0427	NA	NA	NA	22.8	22.8	22.9	NA	NA	NA
cysM SAUSA300_0433	NA	NA	NA	NA	NA	NA	NA	17.1	NA
metB SAUSA300_0434	NA	NA	NA	21.9	21.7	19.8	NA	NA	NA
SAUSA300_0437	NA	NA	NA	NA	NA	NA	18.1	19.3	20.3
sle1 aaa SAUSA300_0438	NA	NA	NA	NA	16.8	NA	20.5	18.7	16.6
SAUSA300_0443	NA	NA	NA	18.6	18.8	18.3	NA	NA	NA
gltB SAUSA300_0445	22.0	NA	NA	NA	NA	NA	NA	NA	NA
treP SAUSA300_0448	NA	NA	NA	20.4	21.1	21.1	20.0	21.2	NA
treC SAUSA300_0449	NA	NA	NA	20.7	21.0	22.7	NA	20.2	20.8
dnaX SAUSA300_0452	23.2	22.3	23.4	21.0	21.2	21.5	21.8	21.4	NA
SAUSA300_0453	NA	NA	NA	25.3	25.7	25.4	25.1	24.6	23.3
recR SAUSA300_0454	NA	NA	NA	18.3	18.7	18.4	NA	NA	NA
SAUSA300_0458	NA	NA	NA	NA	18.7	18.9	NA	NA	NA
tmk SAUSA300_0459	NA	NA	NA	19.2	20.8	NA	NA	NA	NA
SAUSA300_0460	21.7	21.3	NA	21.0	21.1	23.0	NA	20.1	NA
holB SAUSA300_0461	NA	NA	NA	NA	17.8	18.6	NA	19.3	NA
SAUSA300_0462	17.1	NA	NA	17.0	17.4	NA	NA	19.6	NA
SAUSA300_0463	NA	NA	NA	NA	NA	17.3	NA	NA	NA
SAUSA300_0464	NA	NA	NA	20.0	20.2	19.9	NA	18.5	NA
metS metG SAUSA300_0467	23.8	NA	NA	25.7	26.2	26.3	25.2	26.5	24.8
SAUSA300_0468	NA	NA	NA	22.1	22.0	21.8	NA	22.5	NA
rnmV SAUSA300_0469	NA	NA	NA	19.4	19.6	19.4	NA	17.5	NA
rsmA ksgA SAUSA300_0470	NA	NA	NA	22.3	22.3	22.8	NA	NA	NA
ispE SAUSA300_0472	NA	NA	NA	21.4	20.9	21.0	NA	NA	NA
purR SAUSA300_0473	20.4	NA	NA	24.0	24.1	24.3	23.6	22.2	23.9

SAUSA300_0474	19.9	20.1	NA	19.7	20.3	20.3	NA	NA	NA
spoVG SAUSA300_0475	25.9	25.5	NA	23.1	22.8	22.0	23.5	20.3	21.7
glmU SAUSA300_0477	21.7	21.5	NA	23.6	23.5	23.5	22.9	23.3	NA
prs SAUSA300_0478	25.1	24.3	NA	25.3	24.9	25.6	24.2	25.7	24.1
rplY ctc SAUSA300_0479	25.5	25.2	22.1	23.7	22.8	23.1	29.9	28.7	31.3
pth SAUSA300_0480	NA	NA	NA	19.6	19.5	20.1	NA	19.6	NA
mfd SAUSA300_0481	NA	NA	NA	21.1	20.9	20.7	NA	20.3	NA
SAUSA300_0483	NA	NA	NA	NA	NA	NA	20.3	19.7	NA
SAUSA300_0484	NA	NA	NA	22.3	22.4	22.3	20.0	21.6	22.9
SAUSA300_0485	NA	NA	NA	20.0	20.4	20.8	NA	20.8	NA
SAUSA300_0486	18.8	NA	21.6	23.2	22.9	21.8	23.1	24.6	NA
hpt SAUSA300_0488	23.7	24.0	NA	25.1	25.4	22.8	23.0	24.4	24.7
ftsH SAUSA300_0489	23.8	23.9	23.5	25.8	26.0	26.0	26.5	27.1	28.2
hslO SAUSA300_0490	NA	NA	NA	23.3	22.9	23.4	22.4	22.6	21.1
cysK SAUSA300_0491	24.5	24.7	NA	26.1	26.0	26.2	28.3	29.1	29.0
folP SAUSA300_0492	NA	NA	NA	20.1	20.4	21.5	20.9	21.8	NA
lysS SAUSA300_0496	27.1	26.2	25.7	26.5	27.0	26.8	27.1	26.1	25.4
SAUSA300_0503	NA	NA	NA	NA	19.5	20.1	NA	NA	NA
pdxS SAUSA300_0504	24.3	NA	25.4	27.3	27.7	28.0	24.5	25.7	21.1
pdxT SAUSA300_0505	NA	NA	NA	24.3	24.6	24.7	23.5	22.5	NA
nupC SAUSA300_0506	21.6	21.7	NA	16.2	17.2	16.5	16.9	20.5	NA
ctsR SAUSA300_0507	NA	NA	NA	21.4	21.0	20.9	NA	18.2	NA
SAUSA300_0508	NA	NA	NA	NA	NA	NA	NA	18.6	NA
mcsB SAUSA300_0509	NA	NA	NA	20.0	21.1	NA	20.8	18.8	NA
clpC SAUSA300_0510	28.8	28.8	28.0	25.0	24.7	25.0	26.1	25.1	25.3
radA SAUSA300_0511	NA	NA	NA	NA	NA	NA	NA	21.9	NA
SAUSA300_0512	NA	NA	NA	21.9	21.7	22.3	NA	NA	NA
gltX SAUSA300_0513	25.9	25.1	NA	25.8	25.9	26.1	25.4	26.4	NA
cysE SAUSA300_0514	20.6	20.0	NA	21.0	20.1	18.8	NA	NA	NA
cysS SAUSA300_0515	23.1	23.5	NA	25.5	25.2	25.7	23.5	22.9	24.9
mrnC SAUSA300_0516	NA	NA	NA	NA	NA	16.8	NA	NA	NA
SAUSA300_0517	20.7	NA	NA	23.0	23.9	24.2	21.3	24.3	22.6
SAUSA300_0518	NA	NA	NA	20.8	21.0	21.1	NA	21.5	NA
nusG SAUSA300_0521	NA	NA	NA	23.6	24.2	25.3	25.2	23.1	25.4
rplK SAUSA300_0522	24.6	26.0	NA	24.6	24.4	24.6	26.7	26.8	28.3
rplA SAUSA300_0523	26.3	26.0	24.8	26.7	26.8	27.4	27.9	28.1	29.8
rplJ SAUSA300_0524	27.1	26.8	NA	25.7	25.8	25.9	28.9	29.3	29.6
rplL SAUSA300_0525	26.3	25.2	26.6	25.8	25.1	27.4	30.9	29.1	29.8
SAUSA300_0526	NA	NA	NA	18.1	18.5	19.2	NA	NA	NA
rpoB SAUSA300_0527	27.5	26.6	23.8	28.1	28.2	28.4	27.9	28.4	27.7
rpoC SAUSA300_0528	26.4	26.2	27.0	28.6	29.0	29.1	28.2	28.5	28.0
rpsL SAUSA300_0530	27.2	25.5	25.6	24.0	24.1	23.5	26.4	27.7	25.4
rpsG SAUSA300_0531	22.6	24.9	NA	25.8	25.4	25.7	28.2	24.9	29.5
fusA SAUSA300_0532	28.7	28.6	30.0	30.4	30.1	30.2	30.9	31.0	31.2
tuf SAUSA300_0533	29.4	30.9	31.0	31.7	31.7	31.3	33.2	29.8	32.7
SAUSA300_0534	NA	NA	NA	21.6	21.6	21.6	NA	NA	NA
SAUSA300_0535	25.4	23.1	NA	24.8	24.6	24.4	23.9	25.4	22.1
hchA SAUSA300_0536	24.8	23.8	23.9	23.9	24.2	24.5	26.4	25.8	26.6
araB SAUSA300_0537	NA	NA	NA	NA	18.1	NA	NA	NA	NA
SAUSA300_0538	NA	NA	NA	21.0	20.7	21.2	NA	19.3	23.8
ilvE SAUSA300_0539	21.9	22.1	NA	23.2	22.8	23.1	23.5	22.7	22.5
SAUSA300_0540	NA	NA	NA	21.3	20.9	21.4	NA	NA	NA
SAUSA300_0541	NA	NA	NA	22.7	22.0	22.2	NA	NA	NA

SAUSA300_0542	NA	NA	NA	20.5	20.4	19.6	NA	NA	NA
SAUSA300_0544	22.2	22.3	22.8	24.7	24.4	23.8	21.7	23.7	21.8
azo1 SAUSA300_0545	20.1	19.1	NA	NA	NA	NA	NA	NA	NA
sdrD SAUSA300_0547	19.4	24.7	27.2	NA	NA	NA	22.8	22.2	23.0
sdrE SAUSA300_0548	NA	19.8	NA	19.7	NA	19.3	19.2	20.2	NA
SAUSA300_0549	18.7	NA	18.3	NA	NA	NA	17.3	18.2	NA
SAUSA300_0550	NA	NA	NA	22.1	22.5	22.9	21.4	20.7	NA
folE2 SAUSA300_0551	19.6	22.1	19.6	19.3	19.4	19.8	24.1	23.1	21.0
SAUSA300_0552	NA	NA	NA	19.6	19.2	19.3	NA	NA	NA
nagB SAUSA300_0554	NA	NA	NA	21.3	21.7	21.9	NA	21.1	NA
SAUSA300_0555	23.6	23.5	23.8	23.4	23.6	22.9	24.3	24.3	27.6
SAUSA300_0556	21.0	19.4	NA	NA	NA	NA	NA	NA	NA
SAUSA300_0557	20.1	20.0	NA	21.0	21.3	21.9	23.1	18.2	NA
thiD SAUSA300_0562	23.5	22.6	25.7	23.7	23.8	24.6	24.6	25.0	25.9
ung SAUSA300_0563	NA	NA	NA	19.4	19.7	20.7	NA	20.4	NA
SAUSA300_0564	NA	NA	NA	19.2	19.1	18.2	17.8	17.3	NA
SAUSA300_0566	NA	NA	NA	NA	NA	NA	17.6	18.7	NA
SAUSA300_0568	NA	NA	NA	18.3	18.2	NA	NA	NA	NA
SAUSA300_0569	23.4	23.2	23.1	22.4	22.8	23.4	23.8	22.9	25.6
pta SAUSA300_0570	23.7	23.8	NA	25.4	25.9	26.3	26.9	26.2	26.5
mvk SAUSA300_0572	NA	NA	NA	21.2	21.2	21.8	NA	NA	NA
mvaD SAUSA300_0573	NA	NA	NA	NA	NA	19.7	NA	20.9	NA
SAUSA300_0574	NA	NA	NA	23.2	22.4	22.5	NA	21.1	NA
SAUSA300_0576	NA	NA	NA	20.4	20.9	21.3	NA	20.3	NA
SAUSA300_0589	22.2	21.9	NA	20.7	20.0	20.5	NA	20.8	NA
SAUSA300_0593	NA	NA	NA	16.5	16.5	16.9	15.6	20.3	NA
adh SAUSA300_0594	31.8	31.1	30.6	30.2	30.3	30.0	28.2	27.4	31.0
argS SAUSA300_0596	23.7	23.5	NA	25.6	25.5	25.8	26.8	25.6	NA
SAUSA300_0598	NA	NA	NA	NA	NA	NA	NA	18.4	NA
SAUSA300_0600	NA	NA	NA	21.5	21.3	21.6	NA	20.7	NA
SAUSA300_0601	23.2	23.2	NA	22.1	22.6	23.3	20.5	22.4	20.8
SAUSA300_0602	23.4	24.0	21.5	NA	20.4	NA	20.5	20.6	18.8
sarA SAUSA300_0605	26.0	24.3	25.4	21.3	21.5	21.8	24.4	22.1	24.3
SAUSA300_0617	NA	NA	NA	20.3	19.4	19.8	NA	18.0	NA
SAUSA300_0618	27.0	27.3	25.5	25.5	26.2	26.0	25.4	26.7	27.0
SAUSA300_0620	21.1	NA	NA	19.8	19.8	21.0	20.4	18.6	19.3
SAUSA300_0621	NA	NA	NA	22.4	22.2	22.7	NA	NA	NA
tagH SAUSA300_0624	22.6	22.0	NA	23.4	23.4	23.7	NA	NA	NA
tagG SAUSA300_0625	NA	NA	NA	18.8	18.3	NA	NA	NA	NA
tagB SAUSA300_0626	NA	NA	NA	20.3	20.8	20.4	18.9	18.0	NA
tagX SAUSA300_0627	NA	NA	NA	22.2	22.3	22.8	NA	NA	NA
SAUSA300_0628	NA	NA	NA	21.4	21.5	20.6	NA	17.8	NA
pbp4 SAUSA300_0629	NA	NA	NA	20.2	20.2	19.7	NA	21.0	NA
SAUSA300_0630	22.8	22.5	NA	22.6	22.1	21.7	23.1	23.5	NA
SAUSA300_0632	NA	NA	NA	NA	NA	NA	NA	19.6	NA
fhuA SAUSA300_0633	21.1	23.5	NA	NA	NA	24.0	NA	NA	NA
SAUSA300_0636	NA	NA	NA	20.9	20.9	21.2	NA	20.2	18.8
SAUSA300_0637	19.9	19.7	NA	NA	NA	NA	NA	NA	NA
SAUSA300_0641	NA	NA	NA	NA	NA	NA	NA	17.1	NA
SAUSA300_0648	NA	NA	NA	NA	NA	NA	NA	17.6	NA
SAUSA300_0649	22.2	22.6	NA	25.9	25.0	24.8	NA	NA	NA
SAUSA300_0655	22.8	23.5	NA	23.1	23.4	24.2	22.2	23.1	21.7
SAUSA300_0658	NA	21.7	NA	20.1	19.7	19.6	NA	NA	NA
SAUSA300_0662	NA	NA	NA	NA	18.1	18.2	NA	21.2	NA
SAUSA300_0664	NA	NA	NA	24.9	24.6	25.0	NA	NA	NA
SAUSA300_0666	22.5	NA	NA	22.5	23.2	22.0	21.7	20.6	NA
SAUSA300_0668	26.3	NA	23.0	21.7	22.0	22.3	20.5	24.3	22.0
uppP SAUSA300_0669	18.3	NA	NA	19.5	20.5	20.5	21.8	23.1	19.3

SAUSA300_0672	27.5	27.0	27.4	26.5	26.6	27.1	26.9	24.8	25.4
SAUSA300_0674	NA	NA	NA	19.1	NA	20.2	NA	NA	NA
SAUSA300_0676	NA	NA	NA	18.1	18.6	17.8	NA	NA	NA
SAUSA300_0683	NA	NA	NA	23.2	23.1	23.5	NA	21.5	NA
fruB SAUSA300_0684	20.6	20.0	NA	26.6	26.1	26.2	NA	NA	24.4
fruA SAUSA300_0685	24.4	24.4	26.2	25.1	25.6	25.4	24.1	25.1	26.0
nagA SAUSA300_0686	NA	NA	NA	22.3	22.4	22.9	NA	NA	NA
SAUSA300_0688	23.0	23.6	NA	22.7	23.3	24.0	20.0	22.3	22.4
saeS SAUSA300_0690	NA	NA	NA	22.1	21.8	21.7	22.1	21.2	NA
saeR SAUSA300_0691	NA	NA	NA	21.8	21.8	21.9	20.8	22.1	NA
SAUSA300_0692	21.5	22.2	NA	22.3	22.5	23.0	NA	21.0	NA
SAUSA300_0693	22.7	24.8	NA	23.1	22.9	23.2	20.4	24.5	NA
queE SAUSA300_0695	NA	NA	NA	22.2	22.5	22.9	NA	NA	NA
SAUSA300_0696	22.6	23.2	NA	23.8	23.9	24.2	NA	21.7	NA
queC SAUSA300_0697	21.2	NA	NA	25.1	26.2	25.2	22.7	23.3	NA
SAUSA300_0699	NA	NA	NA	17.3	17.2	16.1	NA	16.0	NA
SAUSA300_0701	NA	NA	NA	23.1	23.0	22.3	19.3	NA	NA
SAUSA300_0702	NA	NA	NA	22.1	21.2	22.0	22.3	NA	NA
ltaS SAUSA300_0703	22.2	NA	NA	21.8	21.3	22.6	20.9	20.1	21.5
SAUSA300_0704	NA	NA	21.2	20.8	20.2	20.9	NA	20.9	NA
SAUSA300_0707	NA	NA	NA	20.3	21.2	21.7	NA	18.4	NA
SAUSA300_0709	NA	NA	NA	NA	NA	NA	NA	16.7	NA
queF SAUSA300_0713	NA	NA	NA	20.8	20.3	21.2	NA	NA	NA
SAUSA300_0716	23.9	23.5	NA	25.5	25.8	26.2	25.2	24.0	24.6
SAUSA300_0717	21.2	23.6	NA	25.0	25.2	25.3	22.6	25.1	23.7
SAUSA300_0721	22.8	23.5	NA	NA	NA	NA	22.5	23.8	22.0
murB SAUSA300_0722	22.1	22.5	NA	23.0	22.9	23.2	NA	22.4	NA
SAUSA300_0724	NA	NA	NA	NA	NA	NA	NA	18.2	NA
SAUSA300_0726	NA	NA	NA	21.5	21.9	22.7	NA	20.4	NA
pepT SAUSA300_0727	NA	NA	NA	18.9	18.7	NA	18.0	17.3	NA
SAUSA300_0730	NA	NA	NA	20.9	20.2	20.5	NA	20.1	NA
SAUSA300_0733	22.3	22.3	NA	22.5	22.4	22.5	NA	21.4	NA
hpf SAUSA300_0736	27.1	25.7	27.1	24.4	23.9	23.0	23.5	24.0	21.8
secA1 SAUSA300_0737	25.2	25.1	23.9	27.1	27.0	27.4	26.5	25.6	23.6
SAUSA300_0740	NA	NA	NA	NA	NA	NA	NA	18.9	NA
uvrB SAUSA300_0741	NA	NA	NA	21.9	22.3	23.1	20.2	22.0	NA
uvrA SAUSA300_0742	23.8	25.2	NA	25.3	25.3	25.0	23.0	23.2	NA
hprK SAUSA300_0743	21.0	22.0	NA	21.3	21.5	21.4	NA	22.6	21.3
trxB SAUSA300_0747	24.5	24.7	NA	26.3	26.5	27.4	24.8	25.0	24.7
SAUSA300_0748	NA	NA	NA	22.6	22.0	22.9	NA	NA	NA
SAUSA300_0749	NA	NA	NA	19.3	19.7	20.8	NA	NA	NA
whiA SAUSA300_0750	NA	NA	NA	NA	NA	NA	NA	19.9	NA
clpP SAUSA300_0752	24.5	25.9	NA	NA	NA	NA	17.8	22.1	23.8
SAUSA300_0753	NA	NA	NA	20.6	20.7	21.3	NA	NA	NA
SAUSA300_0754	NA	NA	NA	NA	NA	NA	NA	16.2	NA
SAUSA300_0755	22.3	NA	NA	21.8	22.3	22.8	20.8	23.0	20.3
gap SAUSA300_0756	28.0	28.0	27.0	29.3	29.6	28.7	31.4	30.4	31.8
pgk SAUSA300_0757	25.4	25.5	NA	26.7	26.7	27.7	28.2	27.7	30.4
tpiA SAUSA300_0758	26.5	24.6	21.6	29.5	29.9	30.3	28.1	28.8	28.4
gpmI SAUSA300_0759	27.3	27.3	NA	27.3	28.0	27.0	26.1	28.0	26.4
eno SAUSA300_0760	24.4	25.6	NA	29.8	29.7	30.0	31.2	30.7	32.8
SAUSA300_0761	22.3	NA	NA	NA	NA	NA	NA	NA	NA
secG SAUSA300_0762	NA	19.8	NA	21.1	20.2	20.8	NA	NA	NA
est SAUSA300_0763	NA	NA	NA	21.7	21.3	21.3	NA	NA	NA

rnr SAUSA300_0764	NA	NA	23.7	23.7	23.3	23.5	23.0	23.1	NA
smpB SAUSA300_0765	NA	NA	NA	18.2	18.9	NA	NA	NA	NA
SAUSA300_0767	NA	NA	NA	20.4	20.5	21.0	NA	NA	NA
SAUSA300_0769	NA	NA	NA	22.2	22.9	22.5	22.1	24.2	16.8
SAUSA300_0770	NA	NA	NA	21.5	21.6	23.0	NA	NA	NA
clfA SAUSA300_0772	20.9	19.7	21.5	NA	NA	NA	20.1	18.7	17.1
nuc SAUSA300_0776	NA	NA	NA	18.8	NA	17.6	NA	18.3	NA
SAUSA300_0777	23.3	26.4	NA	24.4	24.6	23.3	NA	25.1	NA
SAUSA300_0781	NA	19.3	NA	NA	NA	NA	NA	NA	NA
SAUSA300_0789	23.6	23.1	NA	24.1	23.8	23.8	23.6	21.8	NA
SAUSA300_0790	23.3	23.7	22.5	23.2	22.7	24.3	22.7	21.7	NA
gcvH SAUSA300_0791	22.5	22.2	21.1	21.4	20.9	21.8	28.3	26.1	NA
SAUSA300_0793	NA	NA	NA	22.5	22.7	23.8	NA	NA	NA
SAUSA300_0798	NA	NA	NA	NA	NA	NA	20.0	21.1	20.7
int SAUSA300_0799	NA	NA	NA	NA	NA	NA	NA	18.8	NA
SAUSA300_0812	18.4	NA	NA	19.7	19.9	20.3	NA	22.5	NA
SAUSA300_0814	NA	NA	NA	NA	NA	NA	NA	19.7	NA
SAUSA300_0816	26.6	26.4	27.8	22.2	23.5	23.3	22.5	23.3	20.0
sufC SAUSA300_0818	NA	20.9	NA	24.6	23.9	24.5	22.9	25.8	25.1
sufD SAUSA300_0819	22.5	NA	NA	23.3	23.2	23.0	24.1	25.2	NA
sufS SAUSA300_0820	NA	NA	NA	22.4	22.2	21.4	NA	NA	NA
SAUSA300_0821	NA	NA	NA	19.6	20.2	19.2	19.9	19.3	NA
SAUSA300_0822	24.7	24.4	NA	23.4	23.1	23.4	24.9	24.2	23.9
SAUSA300_0826	NA	NA	NA	NA	NA	19.3	NA	NA	NA
lipA SAUSA300_0829	NA	NA	NA	23.1	23.4	23.1	23.5	24.5	22.0
SAUSA300_0832	21.2	19.2	NA	17.2	17.5	NA	19.7	18.8	NA
nagD SAUSA300_0833	24.4	24.6	NA	24.2	24.2	24.2	24.8	23.8	22.2
SAUSA300_0834	NA	NA	NA	22.1	21.9	21.5	NA	NA	NA
dltA SAUSA300_0835	23.5	24.0	22.0	23.3	23.4	23.0	23.7	23.6	NA
dltC SAUSA300_0837	21.1	22.0	NA	23.1	23.2	22.7	NA	NA	NA
dltD SAUSA300_0838	21.8	NA	NA	22.7	22.5	22.4	22.8	24.2	21.3
SAUSA300_0839	NA	NA	NA	20.5	19.9	20.6	NA	NA	NA
SAUSA300_0841	NA	NA	NA	21.8	20.4	20.9	NA	NA	NA
SAUSA300_0843	NA	NA	NA	NA	NA	NA	NA	20.4	NA
SAUSA300_0844	26.9	26.2	26.1	24.9	25.0	25.1	26.7	24.6	26.1
ampA SAUSA300_0845	NA	NA	NA	23.3	23.3	23.8	21.7	20.6	NA
mnhG1 SAUSA300_0849	NA	NA	NA	21.5	21.1	20.9	NA	NA	NA
mnhD1 SAUSA300_0852	NA	NA	NA	21.7	21.4	21.2	NA	NA	NA
mnhC1 SAUSA300_0853	NA	NA	NA	17.9	19.5	NA	NA	NA	NA
SAUSA300_0856	NA	NA	NA	20.9	20.9	21.6	NA	NA	NA
SAUSA300_0857	22.5	22.6	26.8	21.9	22.1	23.7	25.1	23.2	23.1
SAUSA300_0859	NA	19.5	NA	21.1	21.1	21.4	21.0	20.0	25.3
rocD argD SAUSA300_0860	19.9	20.4	NA	NA	19.9	20.0	NA	NA	NA
gudB SAUSA300_0861	22.6	22.1	23.1	23.9	24.4	24.5	23.7	23.2	23.8
glpQ SAUSA300_0862	22.1	21.6	NA	NA	NA	NA	NA	NA	NA
argG SAUSA300_0864	20.9	21.2	NA	21.4	NA	NA	NA	18.4	NA
pgi SAUSA300_0865	27.4	27.3	27.3	27.3	27.3	27.7	28.4	27.7	28.7
spsB SAUSA300_0868	23.0	NA	NA	23.9	24.4	25.0	23.5	23.7	21.6
addB rexB SAUSA300_0869	NA	NA	NA	NA	NA	NA	NA	17.7	NA
addA rexA SAUSA300_0870	NA	NA	NA	18.3	18.5	18.9	NA	17.8	NA
SAUSA300_0871	26.7	26.4	26.2	25.6	25.1	25.9	24.6	26.0	26.2
cdr SAUSA300_0873	24.1	23.1	NA	25.1	24.3	24.4	24.3	24.7	NA
SAUSA300_0874	NA	NA	NA	21.9	22.0	22.4	NA	22.7	NA

clpB SAUSA300_0877	29.6	27.1	29.0	24.4	23.6	23.8	24.6	24.9	24.0
SAUSA300_0882	NA	NA	NA	18.0	18.7	17.8	NA	NA	NA
fabH SAUSA300_0885	23.4	24.0	NA	25.7	25.9	26.0	24.9	26.7	24.0
fabF SAUSA300_0886	24.6	24.1	NA	25.9	25.2	26.3	25.6	24.9	26.1
trpS SAUSA300_0897	22.9	22.3	NA	22.6	22.6	23.9	22.1	25.5	22.9
spxA SAUSA300_0898	NA	NA	NA	NA	17.4	18.5	NA	18.3	20.0
mecA SAUSA300_0899	NA	NA	NA	22.5	22.5	23.0	NA	18.8	NA
pepF SAUSA300_0902	24.9	25.1	NA	25.2	25.2	25.0	25.1	23.4	NA
SAUSA300_0905	NA	21.3	NA	20.0	20.0	20.1	21.4	17.9	NA
SAUSA300_0907	NA	NA	NA	20.3	NA	20.3	NA	NA	NA
nadK SAUSA300_0908	NA	NA	NA	20.9	21.5	21.6	19.6	21.2	NA
SAUSA300_0912	22.9	22.3	22.2	26.5	26.7	26.8	25.9	27.2	23.2
SAUSA300_0913	NA	NA	NA	21.8	21.4	21.2	18.0	21.6	NA
SAUSA300_0916	23.5	25.0	NA	23.8	23.7	24.2	NA	24.7	NA
murE SAUSA300_0919	22.3	NA	NA	24.2	23.9	23.5	24.4	23.4	NA
prfC SAUSA300_0921	21.6	21.2	NA	25.4	25.5	26.8	23.0	24.0	24.4
SAUSA300_0923	24.4	NA	NA	24.4	24.4	24.4	23.3	23.9	22.6
SAUSA300_0925	NA	20.3	NA	20.0	19.3	NA	NA	NA	NA
SAUSA300_0930	NA	NA	NA	23.5	23.1	23.5	24.7	24.8	NA
SAUSA300_0938	23.3	NA	NA	NA	NA	NA	20.4	19.6	20.7
SAUSA300_0939	22.9	23.0	NA	22.4	21.8	22.2	24.6	23.4	21.1
menA SAUSA300_0944	NA	NA	NA	21.2	20.8	20.8	NA	20.3	NA
menD SAUSA300_0946	NA	NA	NA	19.9	20.8	21.3	18.2	21.8	NA
menB SAUSA300_0948	28.2	NA	NA	24.1	24.3	25.0	23.1	22.0	26.0
atl SAUSA300_0955	23.9	24.3	25.2	24.0	23.8	23.7	24.0	24.2	21.9
SAUSA300_0958	NA	NA	NA	22.9	23.0	23.7	NA	23.1	NA
fmt SAUSA300_0959	22.7	NA	NA	24.0	24.3	24.5	23.7	21.8	NA
qoxC SAUSA300_0961	NA	NA	NA	22.3	22.3	21.9	NA	22.6	14.9
qoxB SAUSA300_0962	NA	NA	NA	23.3	22.5	22.8	23.2	23.2	27.7
qoxA SAUSA300_0963	28.2	28.2	24.8	24.6	24.8	24.7	24.3	25.7	24.8
folD SAUSA300_0965	24.2	23.9	23.5	23.2	23.2	23.3	22.4	23.4	NA
purC SAUSA300_0968	19.7	18.8	NA	NA	NA	NA	NA	NA	NA
purS SAUSA300_0969	NA	NA	NA	NA	16.8	16.4	19.2	17.2	NA
purL SAUSA300_0971	20.6	20.8	NA	NA	NA	19.0	NA	18.3	NA
purH SAUSA300_0975	19.9	NA	NA	20.5	20.0	21.8	19.0	18.4	NA
purD SAUSA300_0976	NA	NA	NA	NA	NA	NA	NA	18.1	19.1
SAUSA300_0980	NA	NA	NA	21.7	22.0	21.9	NA	20.5	NA
SAUSA300_0981	NA	NA	NA	21.5	21.3	22.1	NA	NA	NA
SAUSA300_0982	24.6	25.0	NA	21.9	NA	21.7	18.7	23.4	21.7
ptsH SAUSA300_0983	26.5	27.2	NA	24.6	24.8	25.1	23.1	22.7	26.2
ptsI SAUSA300_0984	26.0	25.8	24.3	24.8	25.0	25.7	25.3	26.3	25.6
SAUSA300_0986	NA	NA	19.0	NA	NA	NA	NA	19.7	NA
trkA SAUSA300_0988	NA	NA	NA	NA	NA	20.0	NA	NA	NA
rnj1 SAUSA300_0989	23.6	22.7	NA	24.6	25.1	25.1	25.4	26.2	24.3
def SAUSA300_0991	NA	NA	NA	24.7	24.6	25.0	23.7	22.6	23.5
SAUSA300_0992	22.2	NA	NA	NA	NA	NA	21.1	22.6	21.5
pdhA SAUSA300_0993	25.8	25.1	28.3	26.8	27.0	26.7	28.5	29.9	29.4
pdhB SAUSA300_0994	NA	21.9	NA	28.1	27.8	26.9	31.6	29.6	30.0
SAUSA300_0995	26.6	26.0	NA	27.7	27.8	27.7	28.9	28.6	25.3
lpdA SAUSA300_0996	25.5	25.9	25.0	27.3	27.2	27.0	29.3	28.8	29.5

SAUSA300_0997	NA	NA	NA	19.7	19.6	18.8	NA	NA	NA
SAUSA300_0998	NA	NA	NA	16.6	17.0	16.0	17.9	17.4	NA
potA SAUSA300_0999	NA	NA	NA	NA	NA	NA	19.5	19.4	NA
potD SAUSA300_1002	NA	NA	NA	NA	NA	19.1	NA	NA	NA
SAUSA300_1003	23.3	21.4	NA	21.8	21.5	22.3	22.3	21.0	22.2
SAUSA300_1004	NA	NA	NA	20.6	20.2	20.3	NA	19.6	NA
SAUSA300_1006	NA	NA	20.5	19.9	20.1	19.7	NA	21.1	21.1
SAUSA300_1007	21.2	NA	NA	NA	NA	NA	NA	20.9	NA
typA SAUSA300_1009	23.7	NA	25.3	25.3	25.4	25.0	25.5	25.9	24.6
SAUSA300_1011	NA	NA	NA	20.3	20.2	20.4	21.6	20.3	NA
pyc SAUSA300_1014	25.1	24.0	23.5	22.5	23.7	23.6	24.3	23.9	21.9
SAUSA300_1020	NA	NA	NA	20.6	19.9	20.3	NA	NA	NA
SAUSA300_1021	NA	NA	NA	19.7	19.5	19.1	NA	NA	NA
SAUSA300_1023	NA	NA	NA	NA	NA	NA	18.1	19.4	NA
coaD SAUSA300_1024	NA	NA	NA	22.3	21.9	22.3	NA	NA	NA
SAUSA300_1026	NA	NA	NA	21.8	21.9	22.3	19.8	23.1	21.1
rpmF SAUSA300_1027	20.6	20.0	23.6	23.3	23.3	22.6	18.7	22.1	17.5
isdB SAUSA300_1028	29.3	29.4	26.2	NA	NA	24.0	25.6	NA	NA
isdA frpA stbA SAUSA300_1029	27.2	25.8	24.3	22.6	NA	21.9	23.3	24.0	NA
isdC sirD SAUSA300_1030	19.8	19.5	NA	NA	NA	NA	NA	NA	NA
isdG SAUSA300_1035	21.8	20.7	NA	NA	NA	NA	NA	NA	NA
SAUSA300_1036	21.6	NA	NA	19.9	20.5	19.2	20.9	21.3	21.9
pheS SAUSA300_1037	24.8	NA	NA	25.0	24.3	25.1	24.5	25.0	25.1
pheT SAUSA300_1038	23.0	24.7	NA	25.7	25.5	26.5	25.1	22.8	NA
rnhC SAUSA300_1039	NA	NA	NA	NA	NA	NA	NA	18.3	NA
SAUSA300_1042	NA	NA	NA	18.3	18.9	19.5	18.0	21.1	NA
mutS2 SAUSA300_1043	NA	NA	21.3	21.9	21.7	22.0	20.5	21.7	NA
trxA SAUSA300_1044	25.2	24.7	24.2	19.8	19.8	19.6	23.1	21.3	NA
sdhA SAUSA300_1047	22.0	20.5	NA	23.6	23.5	23.1	22.8	22.2	NA
sdhB SAUSA300_1048	NA	NA	NA	19.9	19.3	20.8	21.0	22.2	NA
murI SAUSA300_1049	NA	NA	NA	22.4	22.3	22.7	22.4	21.7	NA
SAUSA300_1050	23.0	22.3	22.7	22.7	22.2	22.0	NA	20.7	NA
SAUSA300_1057	22.1	22.5	23.2	19.7	18.9	17.8	NA	NA	NA
hla SAUSA300_1058	20.7	20.0	NA	19.7	19.5	19.5	NA	18.3	NA
argF SAUSA300_1062	26.6	27.7	26.4	25.2	25.4	26.7	NA	NA	NA
arcC1 SAUSA300_1063	27.4	27.2	25.7	20.8	20.5	22.6	23.9	19.1	19.9
SAUSA300_1064	23.3	26.0	23.4	NA	19.8	20.8	NA	NA	NA
SAUSA300_1070	NA	20.1	NA	NA	NA	NA	NA	NA	NA
bshC SAUSA300_1071	19.0	NA	NA	22.9	21.6	23.6	NA	NA	NA
rsmH mraW SAUSA300_1073	NA	NA	NA	22.5	23.1	22.6	NA	NA	NA
ftsL SAUSA300_1074	NA	NA	NA	NA	NA	NA	NA	19.5	NA
pbpA SAUSA300_1075	22.6	NA	21.9	21.9	21.4	22.0	21.3	21.0	21.0
murD SAUSA300_1077	NA	NA	NA	22.7	23.5	24.1	20.9	23.5	21.7
divIB SAUSA300_1078	20.5	20.8	NA	21.0	21.5	22.2	NA	20.8	19.6
ftsA SAUSA300_1079	26.3	25.3	NA	23.6	23.6	23.7	25.9	24.8	NA
ftsZ SAUSA300_1080	29.8	29.4	26.9	29.0	28.7	29.0	28.0	28.6	27.1
SAUSA300_1082	NA	NA	NA	21.4	21.3	21.3	NA	NA	NA
sepF SAUSA300_1083	NA	NA	21.7	21.8	21.5	21.3	NA	NA	NA
SAUSA300_1086	22.8	22.7	NA	21.0	21.3	20.8	22.6	23.6	22.7
ileS SAUSA300_1087	25.9	26.5	NA	27.3	27.3	27.0	25.5	27.5	26.7

SAUSA300_1090	NA	NA	NA	24.9	24.6	24.9	23.2	24.5	NA
pyrR SAUSA300_1091	NA	NA	NA	NA	NA	NA	19.6	21.4	20.2
pyrP SAUSA300_1092	NA	NA	NA	NA	NA	NA	17.4	23.6	NA
pyrB SAUSA300_1093	NA	NA	NA	NA	NA	NA	26.1	26.5	24.9
pyrC SAUSA300_1094	23.0	23.9	20.3	NA	NA	22.2	27.7	27.1	29.0
carA SAUSA300_1095	NA	NA	NA	NA	NA	NA	23.9	22.7	21.6
carB SAUSA300_1096	24.1	NA	22.7	NA	23.9	24.1	28.0	28.9	27.4
pyrF SAUSA300_1097	NA	NA	NA	NA	NA	NA	22.6	22.7	21.8
pyrE SAUSA300_1098	NA	NA	NA	NA	NA	NA	23.9	23.3	17.1
SAUSA300_1101	NA	NA	NA	21.2	21.2	20.0	23.6	19.7	NA
gmk SAUSA300_1102	20.3	21.2	NA	25.5	25.6	25.6	22.8	24.3	17.7
rpoZ SAUSA300_1103	23.2	23.5	NA	25.0	25.5	26.1	23.0	25.5	24.3
coaBC SAUSA300_1104	NA	NA	25.6	22.9	22.3	23.2	NA	23.1	NA
priA SAUSA300_1105	NA	NA	NA	19.2	19.7	19.2	NA	NA	NA
def SAUSA300_1108	NA	NA	NA	NA	NA	NA	NA	17.8	NA
fmt SAUSA300_1109	NA	NA	NA	21.9	21.8	22.5	NA	20.9	NA
sun SAUSA300_1110	NA	NA	21.2	20.7	20.1	20.8	NA	NA	NA
rlmN SAUSA300_1111	NA	23.2	NA	23.1	23.2	23.2	24.2	22.4	20.2
SAUSA300_1112	NA	NA	NA	24.9	24.1	24.3	NA	22.8	NA
pknB SAUSA300_1113	23.0	21.9	NA	22.1	22.1	22.5	23.0	21.8	24.0
SAUSA300_1118	NA	NA	NA	21.7	22.0	21.7	21.3	22.4	21.4
SAUSA300_1119	24.9	23.2	24.1	26.2	26.5	26.6	26.0	26.2	21.2
plsX SAUSA300_1122	23.3	23.8	NA	23.3	24.1	22.8	27.1	25.2	22.7
fabD SAUSA300_1123	24.1	23.8	NA	23.5	23.4	23.7	23.3	24.1	NA
fabG SAUSA300_1124	26.3	25.9	NA	25.1	25.8	25.8	25.2	26.3	24.4
acpP SAUSA300_1125	27.4	28.7	24.5	25.5	25.6	26.6	25.2	25.6	24.4
rnc SAUSA300_1126	NA	NA	NA	20.5	20.6	20.8	NA	NA	NA
smc SAUSA300_1127	NA	22.4	NA	21.3	21.2	20.8	21.9	22.4	20.8
ftsY SAUSA300_1128	NA	NA	NA	23.1	23.0	23.0	NA	NA	NA
SAUSA300_1129	NA	NA	18.7	NA	NA	NA	NA	19.2	NA
ffh SAUSA300_1130	NA	NA	NA	24.6	24.7	24.6	24.4	25.0	21.8
rpsP SAUSA300_1131	NA	24.8	NA	22.9	23.5	23.0	24.9	23.0	27.2
rimM SAUSA300_1132	NA	NA	NA	20.2	19.8	NA	19.9	21.0	19.7
trmD SAUSA300_1133	NA	NA	NA	22.0	21.7	NA	NA	NA	NA
rplS SAUSA300_1134	NA	NA	NA	26.1	25.9	25.8	25.6	26.3	26.4
SAUSA300_1135	NA	NA	NA	18.7	18.2	18.9	NA	NA	NA
SAUSA300_1136	NA	NA	19.1	20.5	20.6	20.6	18.6	22.2	NA
sucC SAUSA300_1138	24.2	23.5	NA	22.4	21.6	22.3	22.9	23.9	24.7
sucD SAUSA300_1139	24.9	23.4	25.0	22.5	22.3	23.1	25.8	23.8	25.7
topA SAUSA300_1143	NA	NA	NA	22.9	22.2	23.3	22.8	21.2	23.0
trmFO gid SAUSA300_1144	NA	NA	NA	22.8	23.6	22.4	23.4	23.5	NA
hslV SAUSA300_1146	NA	NA	NA	NA	19.4	NA	20.9	20.5	NA
hslU SAUSA300_1147	25.6	23.9	NA	25.0	25.3	25.6	23.9	24.8	24.2
codY SAUSA300_1148	25.2	24.6	26.8	24.9	25.0	25.6	25.5	25.1	23.0
rpsB SAUSA300_1149	26.2	27.0	26.9	27.8	27.8	27.8	29.7	28.5	30.0
tsf SAUSA300_1150	29.0	28.3	28.9	25.5	25.6	25.6	27.5	28.8	29.1
pyrH SAUSA300_1151	24.5	22.4	21.4	24.6	25.6	25.2	24.7	26.8	27.1
frr SAUSA300_1152	27.7	28.8	NA	24.5	24.6	24.8	25.8	26.0	26.3
uppS SAUSA300_1153	20.5	NA	NA	21.5	22.2	21.8	20.5	23.0	NA
SAUSA300_1155	NA	NA	NA	21.1	21.0	19.9	22.0	20.9	18.7
proS SAUSA300_1156	22.9	23.2	NA	26.7	26.1	25.5	25.4	24.6	22.0
polC SAUSA300_1157	NA	NA	NA	22.6	22.6	23.1	NA	22.2	NA
rimP SAUSA300_1158	NA	NA	NA	20.4	20.5	20.9	NA	16.4	NA
nusA	23.7	23.9	NA	24.9	24.7	25.0	26.6	26.0	23.9

SAUSA300_1159									
SAUSA300_1160	NA	NA	NA	23.4	23.5	23.2	NA	22.7	NA
SAUSA300_1161	NA	NA	NA	19.9	21.0	21.0	NA	20.4	18.8
infB SAUSA300_1162	26.6	25.8	27.6	26.1	26.4	25.7	27.9	26.7	25.3
rbfA SAUSA300_1163	NA	NA	NA	22.7	23.6	23.8	NA	NA	NA
truB SAUSA300_1164	NA	NA	NA	22.0	22.4	21.9	21.0	22.3	19.5
ribF SAUSA300_1165	NA	21.8	NA	21.5	21.5	21.5	NA	NA	NA
rpsO SAUSA300_1166	NA	NA	NA	24.1	23.6	23.5	NA	22.8	NA
SAUSA300_1167	23.6	25.4	23.7	26.5	26.4	26.8	25.9	25.9	26.0
rnj2 SAUSA300_1168	23.5	24.7	21.1	26.0	25.9	25.7	23.5	25.6	23.7
ftsK SAUSA300_1169	20.5	NA	NA	22.5	22.4	22.6	20.6	20.6	20.7
SAUSA300_1170	NA	NA	NA	21.8	20.4	21.4	NA	NA	NA
SAUSA300_1171	19.5	NA	NA	20.0	19.8	20.4	18.4	20.8	NA
SAUSA300_1172	22.6	NA	NA	22.1	21.9	21.1	NA	21.1	20.6
SAUSA300_1174	NA	NA	NA	NA	18.8	18.3	NA	19.2	NA
cinA SAUSA300_1177	NA	NA	NA	NA	NA	NA	NA	19.6	NA
recA SAUSA300_1178	25.8	23.4	NA	24.3	23.8	24.7	25.1	25.0	25.6
rny cvfA SAUSA300_1179	NA	23.0	NA	22.1	22.8	23.0	19.8	23.9	NA
SAUSA300_1180	NA	NA	NA	20.2	20.0	20.7	NA	NA	NA
SAUSA300_1181	NA	NA	NA	25.2	24.9	24.9	22.2	21.3	NA
SAUSA300_1182	24.4	NA	NA	23.7	23.3	23.4	22.2	23.5	NA
SAUSA300_1183	NA	NA	NA	NA	NA	NA	NA	20.3	NA
SAUSA300_1184	23.2	23.0	NA	22.0	21.7	22.5	NA	NA	NA
miaB SAUSA300_1185	NA	NA	NA	23.5	23.8	23.9	NA	22.4	NA
mutS SAUSA300_1188	NA	NA	NA	23.9	23.6	24.1	NA	22.0	22.3
mutL SAUSA300_1189	NA	NA	NA	22.2	21.7	22.1	NA	21.7	NA
glpP SAUSA300_1190	20.2	NA	NA	23.1	23.6	23.7	NA	NA	NA
glpK SAUSA300_1192	22.9	23.3	19.7	19.8	19.6	20.9	22.4	21.6	23.8
glpD SAUSA300_1193	28.0	27.3	25.7	22.7	22.8	23.4	26.2	25.6	26.1
SAUSA300_1194	NA	NA	NA	21.3	20.3	19.6	NA	NA	NA
miaA SAUSA300_1195	NA	NA	NA	19.9	20.0	NA	NA	NA	NA
SAUSA300_1197	NA	NA	NA	24.8	25.0	25.0	NA	19.0	NA
SAUSA300_1199	NA	NA	NA	21.5	21.2	21.5	NA	NA	NA
glnA SAUSA300_1201	24.3	24.1	21.9	26.4	26.6	26.9	26.4	26.7	28.9
SAUSA300_1203	NA	NA	NA	NA	NA	NA	NA	16.9	NA
SAUSA300_1215	22.7	23.3	NA	22.4	21.9	21.7	NA	NA	NA
SAUSA300_1216	NA	NA	NA	21.4	21.4	22.1	NA	21.8	NA
nuc SAUSA300_1222	NA	NA	NA	19.9	20.4	20.0	NA	21.4	NA
SAUSA300_1223	23.2	24.6	NA	24.5	25.1	25.0	21.6	24.2	21.7
SAUSA300_1226	21.0	20.1	19.4	NA	NA	NA	NA	NA	NA
thrC SAUSA300_1227	21.7	21.2	NA	21.1	20.5	21.2	NA	NA	NA
SAUSA300_1229	19.7	21.3	NA	21.6	21.7	21.4	21.9	22.4	23.1
katA SAUSA300_1232	23.4	23.0	NA	22.1	21.5	22.0	21.3	22.8	20.5
rpmG1 SAUSA300_1233	NA	26.7	NA	23.3	23.2	23.2	NA	NA	NA
guaC SAUSA300_1235	24.4	20.0	NA	NA	NA	NA	21.5	20.6	NA
lexA SAUSA300_1237	NA	NA	NA	NA	NA	19.4	NA	NA	NA
SAUSA300_1238	NA	20.5	NA	NA	NA	NA	20.0	20.9	20.0
tkt SAUSA300_1239	28.0	27.2	28.1	29.0	29.2	28.7	28.2	28.4	29.7
SAUSA300_1240	NA	NA	22.6	20.9	21.6	22.6	24.6	21.2	25.0
sbcD SAUSA300_1242	NA	NA	NA	20.7	20.4	20.7	NA	NA	NA
sbcC SAUSA300_1243	NA	NA	NA	NA	NA	NA	21.2	18.7	NA
opuD SAUSA300_1245	22.8	23.0	NA	22.4	22.5	23.0	23.6	20.6	20.7
acnA SAUSA300_1246	24.1	25.6	21.4	26.1	25.2	26.0	26.1	26.5	25.3
SAUSA300_1247	NA	NA	NA	20.3	20.3	NA	19.2	18.0	NA

SAUSA300_1248	20.4	21.6	NA	20.9	20.9	22.5	NA	19.5	NA
plsY SAUSA300_1249	NA	NA	NA	18.0	19.1	18.1	NA	NA	NA
parE SAUSA300_1250	NA	20.9	NA	23.2	23.4	23.7	22.9	23.3	NA
parC SAUSA300_1251	21.8	23.3	NA	22.8	23.4	23.1	20.9	23.4	NA
SAUSA300_1252	NA	NA	NA	19.5	19.4	19.6	NA	NA	NA
glcT SAUSA300_1253	NA	NA	NA	17.6	17.8	17.6	NA	19.0	NA
fmtC mprF SAUSA300_1255	NA	NA	NA	19.1	19.7	20.1	NA	NA	NA
msrR SAUSA300_1257	21.9	23.0	NA	22.5	23.3	22.8	22.4	20.4	NA
SAUSA300_1258	24.4	24.9	NA	20.8	20.7	20.9	NA	NA	NA
SAUSA300_1261	20.9	19.9	19.7	21.0	20.9	21.9	20.8	21.5	21.3
femA SAUSA300_1269	23.6	25.0	NA	23.5	24.2	24.2	23.6	25.8	24.3
femB SAUSA300_1270	23.5	22.6	24.0	23.6	23.5	23.7	NA	22.8	NA
SAUSA300_1272	NA	NA	NA	NA	18.3	19.4	NA	NA	NA
SAUSA300_1277	NA	NA	NA	NA	20.5	21.5	NA	18.6	NA
pepF SAUSA300_1278	NA	NA	NA	21.1	20.8	22.0	NA	20.1	NA
cvfB SAUSA300_1284	NA	NA	22.4	NA	NA	NA	22.8	23.0	NA
SAUSA300_1285	24.3	21.2	NA	25.5	24.7	25.0	25.7	24.7	24.4
lysA SAUSA300_1293	NA	NA	NA	NA	NA	NA	NA	20.6	NA
cspA SAUSA300_1295	24.0	25.4	29.3	26.4	25.4	25.7	27.7	27.3	26.0
acyP SAUSA300_1297	NA	NA	NA	NA	20.4	21.7	NA	NA	NA
SAUSA300_1299	19.3	19.6	NA	20.7	21.1	21.1	NA	18.5	19.9
SAUSA300_1301	NA	NA	NA	NA	17.6	17.1	NA	NA	NA
SAUSA300_1302	NA	NA	NA	21.1	21.3	21.6	19.4	21.4	21.6
SAUSA300_1303	22.6	NA	NA	23.9	24.0	23.0	NA	21.7	NA
SAUSA300_1304	NA	NA	NA	22.1	22.5	22.7	21.6	23.2	21.3
odhB sucB SAUSA300_1305	22.7	21.6	NA	20.4	20.6	20.7	21.7	21.1	21.0
odhA SAUSA300_1306	NA	NA	NA	19.7	19.2	19.4	19.9	21.5	NA
arlR SAUSA300_1308	NA	NA	NA	19.3	18.8	18.7	NA	NA	NA
murG SAUSA300_1311	NA	19.4	NA	18.6	NA	18.6	NA	NA	NA
ctpA SAUSA300_1313	NA	NA	NA	18.6	18.5	20.6	18.2	21.1	NA
crr SAUSA300_1315	25.2	24.3	25.3	23.9	23.7	23.1	26.0	23.0	22.3
msrB SAUSA300_1316	22.8	23.3	NA	21.3	21.1	23.2	NA	NA	NA
msrA SAUSA300_1317	NA	NA	NA	21.9	21.0	21.2	NA	NA	NA
SAUSA300_1318	NA	NA	NA	NA	NA	NA	20.5	18.7	20.1
folA SAUSA300_1319	NA	NA	NA	22.5	22.8	23.3	NA	NA	NA
thyA SAUSA300_1320	NA	NA	NA	23.2	22.8	22.7	NA	NA	NA
SAUSA300_1321	NA	NA	NA	21.7	21.0	21.7	NA	NA	NA
cvfC SAUSA300_1322	NA	NA	NA	NA	20.9	NA	NA	20.4	NA
SAUSA300_1323	21.9	21.9	22.6	22.7	22.9	NA	21.0	22.0	NA
ebh SAUSA300_1327	NA	NA	NA	18.6	19.3	19.3	21.6	NA	NA
norB SAUSA300_1328	NA	NA	NA	NA	18.3	NA	NA	17.6	NA
SAUSA300_1329	NA	NA	NA	21.5	21.2	18.8	NA	23.2	20.3
tdcB SAUSA300_1330	30.2	29.7	31.1	29.1	29.4	29.3	29.3	27.5	28.2
ald1 SAUSA300_1331	27.7	28.7	29.3	30.5	30.2	29.8	27.2	26.0	28.4
SAUSA300_1333	NA	NA	NA	18.9	20.1	19.7	NA	NA	NA
SAUSA300_1334	NA	NA	NA	NA	17.7	NA	NA	19.0	NA
SAUSA300_1336	20.5	NA	NA	23.5	23.7	24.6	22.7	20.1	23.0
SAUSA300_1337	23.9	25.5	24.7	22.8	22.6	24.9	23.2	22.9	26.2
pbp2 SAUSA300_1341	24.4	24.0	NA	25.3	25.1	25.2	25.6	25.0	24.5
asnS SAUSA300_1345	23.5	24.7	NA	25.6	25.9	24.9	26.5	26.0	25.7
dinG SAUSA300_1346	NA	NA	NA	NA	NA	NA	18.9	NA	NA
SAUSA300_1348	NA	NA	NA	17.9	17.7	NA	NA	NA	NA
SAUSA300_1349	NA	NA	NA	21.3	21.1	22.0	NA	20.8	NA
SAUSA300_1350	NA	NA	NA	23.2	23.3	22.8	NA	21.0	NA

SAUSA300_1354	NA	NA	19.9	20.0	19.8	21.1	NA	NA	NA
aroA SAUSA300_1355	NA	NA	NA	21.2	21.4	21.7	NA	21.2	NA
aroC SAUSA300_1357	19.4	NA	NA	22.1	22.5	21.5	22.7	21.0	NA
ndk SAUSA300_1358	NA	NA	NA	22.8	22.1	22.1	NA	NA	NA
SAUSA300_1359	NA	NA	NA	19.8	20.0	19.4	NA	NA	NA
ubiE SAUSA300_1360	NA	20.1	NA	24.3	24.0	24.2	22.3	22.9	NA
SAUSA300_1361	20.3	19.7	NA	21.6	21.6	24.3	NA	NA	NA
hup SAUSA300_1362	28.3	29.5	28.9	30.9	30.4	30.6	28.5	28.0	30.8
gpsA SAUSA300_1363	NA	19.1	NA	NA	20.0	19.7	20.6	20.0	NA
der engA SAUSA300_1364	NA	21.9	NA	23.7	23.2	24.6	23.5	24.2	23.2
rpsA SAUSA300_1365	26.6	27.0	24.7	25.4	25.3	26.0	26.0	25.9	26.3
cmk SAUSA300_1367	23.9	25.3	NA	24.3	24.4	24.2	22.6	23.7	NA
ansA SAUSA300_1368	NA	NA	NA	21.7	21.1	22.3	NA	NA	NA
SAUSA300_1369	20.5	20.9	NA	22.6	22.2	22.1	21.2	20.0	NA
ebpS SAUSA300_1370	24.3	24.5	22.4	22.3	22.4	22.7	23.6	24.0	26.4
recQ SAUSA300_1371	NA	NA	NA	19.4	18.4	19.9	NA	NA	NA
SAUSA300_1373	NA	NA	NA	20.8	20.1	NA	NA	NA	NA
lukF-PV SAUSA300_1381	NA	NA	NA	20.8	21.7	22.4	NA	NA	NA
lukS-PV SAUSA300_1382	20.7	NA	NA	19.8	20.0	20.5	NA	NA	NA
SAUSA300_1393	20.2	20.7	NA	NA	NA	NA	19.6	23.6	NA
SAUSA300_1434	NA	18.5	NA	NA	NA	NA	NA	NA	NA
SAUSA300_1440	19.9	NA	NA	NA	NA	NA	NA	NA	NA
srrB SAUSA300_1441	20.7	NA	NA	22.4	21.7	21.5	NA	NA	NA
srrA SAUSA300_1442	24.1	24.2	NA	25.0	24.6	24.8	23.0	25.4	23.3
rluB SAUSA300_1443	NA	NA	NA	22.9	23.0	22.7	23.0	24.0	21.6
scpB SAUSA300_1444	NA	NA	NA	20.2	NA	NA	NA	NA	NA
scpA SAUSA300_1445	NA	NA	NA	18.3	18.2	18.1	NA	NA	NA
SAUSA300_1448	21.7	22.7	NA	22.2	22.1	21.6	20.6	22.5	NA
SAUSA300_1449	NA	NA	NA	21.3	21.4	20.5	20.5	20.3	20.4
SAUSA300_1450	NA	NA	NA	21.8	22.6	22.9	22.6	23.2	NA
SAUSA300_1451	NA	NA	NA	NA	NA	14.6	NA	NA	NA
proC SAUSA300_1452	NA	NA	NA	20.5	20.6	20.7	NA	18.3	NA
rnz SAUSA300_1453	NA	NA	NA	21.2	20.9	21.0	NA	NA	NA
zwf SAUSA300_1454	24.7	25.0	NA	25.4	25.4	25.7	22.6	25.4	23.6
SAUSA300_1456	20.7	20.0	NA	20.2	20.6	20.2	NA	20.0	NA
SAUSA300_1458	NA	NA	NA	20.2	19.8	20.5	NA	NA	NA
gnd SAUSA300_1459	25.6	23.6	24.8	26.0	26.2	26.9	27.7	26.9	28.7
SAUSA300_1460	21.5	NA	19.0	22.2	21.3	22.1	23.1	23.4	NA
SAUSA300_1463	NA	NA	NA	21.4	21.6	21.2	22.6	20.3	NA
SAUSA300_1464	22.9	NA	22.9	25.0	24.8	24.8	23.0	22.0	19.2
SAUSA300_1465	NA	NA	NA	23.8	24.1	24.2	NA	24.1	NA
SAUSA300_1466	23.6	24.1	23.1	23.1	23.3	22.7	24.7	23.8	NA
lpdA SAUSA300_1467	20.8	NA	NA	21.2	21.2	21.8	NA	22.0	NA
recN SAUSA300_1468	NA	NA	NA	22.9	22.2	21.7	NA	23.2	NA
argR SAUSA300_1469	20.6	22.1	NA	21.2	20.9	20.5	23.3	20.5	NA
SAUSA300_1470	NA	NA	NA	21.5	20.7	21.9	22.2	22.1	NA
xseB SAUSA300_1471	NA	NA	NA	21.7	21.7	22.0	NA	NA	NA
xseA SAUSA300_1472	NA	21.4	NA	NA	NA	NA	NA	20.2	NA
nusB SAUSA300_1473	NA	19.9	NA	20.3	20.5	20.5	NA	20.5	18.6
SAUSA300_1474	NA	NA	NA	19.0	18.4	18.3	19.3	23.0	NA
accC SAUSA300_1475	NA	18.6	NA	23.6	22.8	23.2	20.5	21.1	26.5
accB SAUSA300_1476	22.5	22.2	NA	NA	NA	NA	20.9	23.5	23.7
SAUSA300_1478	21.6	21.4	NA	NA	NA	NA	22.9	22.4	16.5
SAUSA300_1489	NA	NA	NA	18.9	18.7	19.1	NA	18.5	NA
efp SAUSA300_1490	NA	23.0	NA	25.4	26.1	25.7	24.6	25.1	27.6
SAUSA300_1491	22.6	NA	NA	22.6	22.3	22.6	23.5	23.1	NA

lipM SAUSA300_1494	NA	21.1	NA	21.7	21.4	20.7	21.2	22.1	NA
SAUSA300_1495	NA	NA	NA	23.9	24.6	24.1	20.9	22.7	NA
gcvPB SAUSA300_1496	22.0	23.4	NA	21.0	21.4	22.3	22.4	22.6	NA
gcvPA SAUSA300_1497	21.4	20.8	NA	21.0	21.2	19.2	NA	NA	NA
SAUSA300_1505	NA	NA	NA	NA	16.8	NA	NA	17.6	NA
SAUSA300_1506	20.9	20.2	NA	20.0	19.6	20.1	NA	NA	NA
glk SAUSA300_1507	19.5	NA	NA	20.3	19.9	19.5	22.9	22.0	NA
SAUSA300_1509	NA	NA	NA	18.2	18.6	19.3	NA	NA	NA
rpmG2 SAUSA300_1511	NA	NA	NA	23.4	22.9	22.9	NA	23.1	NA
pbp3 SAUSA300_1512	NA	NA	NA	21.2	21.1	21.9	NA	22.6	21.5
sodA SAUSA300_1513	23.4	22.7	20.6	20.7	21.0	21.8	24.7	25.4	25.3
fur SAUSA300_1514	NA	NA	16.5	NA	NA	NA	19.5	19.3	NA
SAUSA300_1516	NA	NA	NA	19.7	19.7	19.4	NA	20.2	NA
nfo SAUSA300_1517	22.8	23.4	NA	23.2	22.9	23.3	22.8	23.9	NA
cshB SAUSA300_1518	22.6	21.7	20.7	21.6	21.7	21.6	22.7	23.2	21.6
SAUSA300_1519	NA	NA	NA	19.1	19.0	NA	NA	20.2	NA
SAUSA300_1520	NA	NA	NA	22.7	22.5	22.2	NA	NA	NA
rpoD sigA SAUSA300_1521	22.5	23.4	NA	24.2	24.2	23.8	23.9	23.1	21.5
dnaG SAUSA300_1522	NA	NA	NA	NA	NA	NA	19.6	NA	NA
SAUSA300_1523	NA	21.2	NA	23.3	23.9	24.3	NA	23.0	21.1
SAUSA300_1524	NA	NA	NA	20.2	20.1	20.3	NA	19.6	NA
glyQS SAUSA300_1525	24.2	23.8	23.1	26.2	26.0	25.6	25.2	27.1	NA
era SAUSA300_1527	22.5	21.6	NA	23.6	23.7	22.9	22.0	21.9	NA
cdd SAUSA300_1528	NA	NA	NA	18.6	18.3	NA	NA	NA	NA
ybeY SAUSA300_1530	NA	NA	NA	20.8	20.5	NA	NA	NA	NA
phoH SAUSA300_1531	18.5	NA	NA	18.9	19.1	19.0	18.9	21.0	19.2
SAUSA300_1532	NA	NA	NA	21.8	21.5	22.2	NA	NA	NA
SAUSA300_1533	27.1	26.4	24.4	26.8	26.7	27.1	25.7	25.7	26.8
rpsU SAUSA300_1535	23.7	23.1	NA	20.8	19.6	NA	NA	19.7	NA
SAUSA300_1536	NA	NA	NA	21.0	20.9	21.9	22.4	21.6	21.7
SAUSA300_1537	NA	NA	NA	19.1	NA	19.1	NA	16.3	NA
dnaJ SAUSA300_1539	22.4	23.2	22.7	22.4	22.2	22.2	23.4	24.3	23.1
dnaK SAUSA300_1540	31.4	31.2	29.5	26.3	26.9	26.8	28.3	27.7	27.4
grpE SAUSA300_1541	26.2	26.1	23.7	NA	NA	NA	21.8	21.8	19.4
hrcA SAUSA300_1542	21.7	21.9	20.8	NA	NA	18.5	NA	NA	NA
lepA SAUSA300_1544	23.1	22.0	NA	23.4	23.1	22.4	24.7	23.0	NA
rpsT SAUSA300_1545	27.8	28.9	23.4	23.7	23.3	22.8	26.6	27.2	27.6
SAUSA300_1548	22.0	NA	NA	21.6	21.0	22.6	20.3	19.1	NA
rsfS SAUSA300_1551	20.7	23.4	NA	24.0	23.5	23.9	22.3	24.4	24.4
SAUSA300_1552	NA	NA	NA	20.6	19.8	20.0	19.9	NA	NA
SAUSA300_1554	21.1	NA	19.9	22.4	22.4	20.9	19.6	24.5	17.1
aroE SAUSA300_1555	NA	NA	21.0	NA	NA	NA	NA	20.3	NA
mtnN SAUSA300_1558	22.3	21.2	23.7	23.3	22.8	22.6	20.9	23.7	21.2
greA SAUSA300_1567	23.9	21.4	NA	24.6	24.6	24.1	26.7	25.6	20.6
udk SAUSA300_1568	NA	NA	21.8	23.2	23.1	23.6	21.7	24.5	22.6
SAUSA300_1569	NA	NA	NA	22.3	22.1	20.8	22.8	22.3	16.9
SAUSA300_1570	NA	NA	NA	19.6	19.4	20.8	20.3	22.4	NA
SAUSA300_1572	NA	20.7	NA	22.0	21.9	21.3	NA	NA	NA
SAUSA300_1573	NA	NA	NA	19.1	19.4	18.6	NA	NA	NA
SAUSA300_1574	NA	NA	NA	NA	22.2	22.7	NA	NA	NA
alaS SAUSA300_1575	23.5	22.8	24.7	24.6	24.8	24.5	24.6	25.3	25.8
recD2 SAUSA300_1576	NA	NA	NA	19.4	19.2	NA	NA	NA	NA

mnmA SAUSA300_1578	NA	NA	NA	22.9	23.4	23.3	NA	23.0	21.4
SAUSA300_1579	NA	NA	NA	19.6	20.8	21.1	NA	21.3	NA
SAUSA300_1580	NA	NA	NA	NA	17.5	NA	NA	16.5	NA
SAUSA300_1581	25.4	24.0	25.9	20.8	20.5	20.6	14.5	17.4	NA
SAUSA300_1582	NA	NA	NA	17.1	17.3	NA	18.9	NA	NA
SAUSA300_1583	NA	NA	21.2	23.1	22.8	23.2	23.2	22.7	NA
SAUSA300_1584	NA	NA	NA	21.1	21.6	22.3	NA	21.2	NA
SAUSA300_1585	NA	NA	NA	23.9	23.5	23.3	22.9	24.3	NA
aspS SAUSA300_1586	24.0	NA	NA	22.8	23.0	23.1	23.2	21.3	22.2
hisS SAUSA300_1587	NA	NA	21.5	22.5	22.6	22.4	24.5	22.7	22.7
lytH SAUSA300_1588	NA	NA	NA	21.1	20.3	21.7	NA	19.8	NA
dtd SAUSA300_1589	NA	NA	NA	21.9	21.7	21.9	NA	20.4	NA
SAUSA300_1590	23.3	22.8	NA	22.7	22.7	22.6	21.2	22.4	NA
apt SAUSA300_1591	19.8	NA	NA	22.2	22.1	23.7	NA	NA	NA
recJ SAUSA300_1592	NA	NA	NA	20.3	21.1	21.4	NA	17.7	NA
secF secD SAUSA300_1593	24.9	24.5	NA	25.9	25.9	26.0	24.9	24.5	24.7
yajC SAUSA300_1594	NA	NA	NA	25.0	25.4	25.8	23.6	25.5	24.0
tgt SAUSA300_1595	22.9	21.5	NA	20.3	19.8	22.5	NA	20.5	NA
queA SAUSA300_1596	NA	NA	NA	23.4	22.9	23.5	20.2	23.2	21.2
ruvB SAUSA300_1597	NA	NA	NA	NA	NA	20.7	NA	19.7	NA
ruvA SAUSA300_1598	NA	NA	NA	21.3	21.7	22.3	NA	23.5	22.5
obg SAUSA300_1600	22.6	22.0	NA	22.6	22.8	23.4	22.3	22.4	22.8
rpmA SAUSA300_1601	NA	NA	18.8	NA	17.0	NA	25.9	27.3	27.1
rplU SAUSA300_1603	26.3	27.7	25.4	27.9	27.4	27.4	28.1	26.7	28.3
mreC SAUSA300_1605	NA	NA	20.5	20.3	20.3	19.9	21.0	22.0	20.0
SAUSA300_1606	NA	NA	NA	17.6	NA	19.6	NA	NA	NA
folC SAUSA300_1610	19.4	NA	NA	23.1	22.4	22.1	NA	NA	NA
valS SAUSA300_1611	25.1	23.9	NA	27.8	28.0	27.0	25.5	26.1	25.3
tag SAUSA300_1612	NA	NA	NA	NA	NA	NA	18.3	NA	19.3
hemL1 SAUSA300_1614	21.9	21.1	NA	23.6	22.8	23.1	25.1	24.5	28.1
hemB SAUSA300_1615	NA	NA	NA	24.7	24.4	24.8	NA	23.7	NA
hemC SAUSA300_1617	24.9	NA	NA	23.6	23.5	24.0	23.1	22.0	NA
engB SAUSA300_1620	NA	NA	NA	20.9	19.9	20.1	NA	NA	NA
clpX SAUSA300_1621	23.7	22.7	NA	24.2	24.5	24.8	27.3	26.5	28.8
tig SAUSA300_1622	27.2	27.4	24.7	27.3	27.5	27.3	29.0	30.3	30.2
SAUSA300_1624	22.7	22.2	NA	22.5	22.0	21.9	22.7	24.3	22.8
rplT SAUSA300_1625	25.2	25.4	23.6	23.4	22.9	24.2	24.5	26.3	25.2
rpmI SAUSA300_1626	20.9	20.0	NA	21.0	20.6	21.4	23.1	20.2	NA
infC SAUSA300_1627	23.5	23.5	23.0	24.6	24.7	24.7	24.4	25.5	28.1
lysP SAUSA300_1628	NA	NA	NA	19.8	20.1	18.8	NA	18.4	NA
thrS SAUSA300_1629	29.0	28.7	29.9	27.8	27.5	27.4	26.7	28.9	27.8
dnaI SAUSA300_1630	NA	18.6	NA	21.1	21.4	21.5	NA	NA	NA
nrdR SAUSA300_1632	NA	NA	NA	19.7	19.9	NA	NA	18.9	NA
coaE SAUSA300_1634	NA	NA	NA	22.8	23.0	23.4	21.4	20.9	NA
polA SAUSA300_1636	22.8	21.8	NA	24.7	24.8	24.9	23.8	24.2	23.8
phoP SAUSA300_1639	19.0	19.4	NA	20.3	20.2	20.3	NA	NA	NA
icd SAUSA300_1640	22.2	20.2	NA	19.9	20.1	20.8	22.5	23.0	NA
gltA SAUSA300_1641	NA	NA	NA	19.8	18.8	19.4	NA	NA	NA
pyk SAUSA300_1644	27.6	26.3	26.8	30.7	30.8	30.8	29.2	29.2	30.6
pfkA SAUSA300_1645	23.5	26.0	NA	26.5	27.3	27.6	24.5	25.9	25.4
accA SAUSA300_1646	NA	NA	NA	22.4	22.5	22.1	21.6	23.2	NA

accD SAUSA300_1647	22.2	NA	NA	21.3	21.7	21.6	22.2	24.3	NA
SAUSA300_1648	24.5	23.6	NA	25.3	25.6	25.5	26.6	25.7	23.3
SAUSA300_1650	NA	NA	NA	23.5	23.3	23.5	22.7	21.6	NA
SAUSA300_1651	NA	NA	NA	19.0	19.8	20.4	19.5	NA	NA
SAUSA300_1652	NA	NA	NA	25.5	25.6	25.8	NA	25.7	24.1
SAUSA300_1653	23.4	23.7	NA	23.5	23.1	21.9	NA	22.4	24.9
SAUSA300_1654	NA	NA	NA	23.2	23.2	23.5	23.8	23.4	NA
ald2 SAUSA300_1655	22.1	21.5	NA	20.8	20.8	21.3	NA	21.2	21.0
SAUSA300_1656	27.8	29.3	30.6	24.4	24.1	24.2	27.0	24.0	27.8
ackA SAUSA300_1657	26.6	26.3	26.3	26.8	26.8	26.7	28.9	28.1	28.3
SAUSA300_1658	NA	NA	NA	17.6	18.1	18.0	20.0	22.9	20.6
tpx SAUSA300_1659	25.7	24.7	NA	24.5	23.9	23.8	NA	NA	NA
thiI SAUSA300_1661	NA	NA	NA	23.2	22.7	22.9	NA	21.9	NA
SAUSA300_1662	NA	NA	NA	NA	NA	NA	20.7	19.0	NA
ezrA SAUSA300_1664	24.1	23.5	NA	27.0	26.6	26.5	25.1	26.0	25.7
SAUSA300_1665	NA	NA	NA	NA	NA	NA	22.7	19.6	NA
rpsD SAUSA300_1666	22.9	24.9	NA	23.7	23.1	24.6	25.7	26.1	28.0
nagE SAUSA300_1672	NA	NA	NA	21.6	21.5	22.1	NA	22.7	NA
SAUSA300_1673	22.0	NA	NA	25.3	25.3	25.6	23.6	23.2	NA
SAUSA300_1674	22.1	20.9	NA	22.1	21.4	21.2	22.0	23.1	20.5
tyrS SAUSA300_1675	24.6	23.4	NA	24.4	24.6	24.5	25.7	25.0	25.3
fhs SAUSA300_1678	21.4	23.5	NA	23.1	23.6	24.4	25.1	25.8	23.5
ccpA SAUSA300_1682	22.6	25.2	NA	22.0	22.3	23.0	22.8	24.4	23.0
SAUSA300_1683	23.7	21.7	NA	22.7	22.8	22.9	NA	20.2	NA
SAUSA300_1684	23.7	23.9	22.9	21.4	20.6	21.4	23.7	25.2	20.9
SAUSA300_1685	25.5	27.5	NA	26.7	26.8	26.5	26.5	26.2	29.2
murC SAUSA300_1686	21.9	21.6	20.2	22.9	23.8	23.5	20.6	21.2	22.2
SAUSA300_1687	23.8	NA	NA	25.0	25.3	25.0	24.7	24.6	22.4
SAUSA300_1688	25.0	25.1	NA	25.5	25.5	25.8	26.3	24.8	26.0
SAUSA300_1689	NA	NA	NA	20.7	19.6	21.1	21.3	21.6	NA
SAUSA300_1690	23.4	NA	22.3	23.8	23.7	22.8	NA	19.2	NA
SAUSA300_1691	24.7	24.2	NA	23.5	23.9	24.7	NA	24.9	NA
SAUSA300_1692	NA	NA	NA	20.6	20.5	19.9	NA	NA	NA
trmB SAUSA300_1694	NA	NA	NA	22.8	22.5	22.8	NA	NA	NA
dat SAUSA300_1696	24.4	23.9	NA	22.9	22.8	22.4	24.0	24.8	22.3
SAUSA300_1697	25.0	24.4	24.9	24.0	24.5	24.1	25.8	25.1	26.8
SAUSA300_1698	23.3	22.5	22.0	18.7	NA	19.0	21.2	22.3	17.4
SAUSA300_1701	NA	NA	NA	19.1	19.3	19.8	NA	NA	NA
SAUSA300_1703	18.8	NA	NA	NA	NA	18.6	20.0	20.8	NA
leuS SAUSA300_1704	24.0	24.1	NA	25.7	26.2	26.3	26.4	25.4	24.6
SAUSA300_1706	NA	NA	NA	18.0	17.9	NA	17.5	18.5	NA
rot SAUSA300_1708	NA	NA	NA	NA	NA	NA	20.5	19.8	24.5
SAUSA300_1710	NA	NA	NA	21.1	21.9	22.0	18.8	22.0	NA
putA SAUSA300_1711	20.3	19.4	NA	NA	NA	NA	NA	NA	NA
ribH SAUSA300_1712	NA	NA	NA	20.1	20.9	21.1	NA	NA	NA
SAUSA300_1720	20.5	NA	NA	21.3	21.2	20.9	NA	21.1	19.6
SAUSA300_1725	22.1	22.0	NA	22.8	21.8	22.8	24.4	23.1	23.2
SAUSA300_1728	NA	NA	NA	20.9	21.8	21.5	NA	19.8	20.1
SAUSA300_1729	25.4	25.2	NA	22.9	23.3	23.6	23.3	21.6	23.7
metK SAUSA300_1730	NA	23.5	NA	24.9	25.0	25.6	24.4	23.4	NA
pckA SAUSA300_1731	24.1	23.2	NA	NA	20.7	NA	NA	NA	NA
SAUSA300_1750	NA	NA	NA	NA	NA	NA	NA	17.0	NA
epiF SAUSA300_1762	21.2	20.6	NA	NA	NA	NA	NA	NA	NA
epiC SAUSA300_1765	NA	NA	NA	NA	NA	23.0	NA	20.1	NA
SAUSA300_1780	NA	NA	NA	21.6	21.2	21.8	NA	NA	NA
hemG	20.5	19.7	NA	23.5	23.1	23.0	22.5	18.6	NA

SAUSA300_1781									
hemH SAUSA300_1782	NA	NA	NA	20.0	20.3	19.7	NA	19.4	NA
hemE SAUSA300_1783	21.2	22.1	NA	24.2	24.8	24.7	22.5	23.8	23.1
traP SAUSA300_1784	NA	NA	NA	20.9	20.6	21.2	23.5	21.5	23.8
SAUSA300_1785	20.4	NA	NA	21.4	20.2	20.9	NA	NA	NA
SAUSA300_1786	19.1	NA	NA	NA	NA	NA	17.6	19.9	NA
SAUSA300_1787	21.3	NA	NA	20.8	20.1	21.4	NA	19.7	NA
SAUSA300_1788	NA	NA	NA	23.8	23.4	23.6	18.5	23.9	NA
prsA SAUSA300_1790	24.2	23.9	24.6	21.9	22.8	22.0	23.1	24.0	22.6
cbf1 SAUSA300_1791	22.7	NA	24.3	21.2	21.3	19.3	21.3	21.8	NA
SAUSA300_1792	23.8	23.8	21.3	24.7	24.4	24.2	22.6	24.6	23.9
SAUSA300_1793	NA	20.4	NA	20.1	20.2	21.1	21.6	22.4	NA
SAUSA300_1795	24.7	25.1	28.2	24.5	24.3	24.9	23.3	22.1	20.8
SAUSA300_1797	NA	NA	21.7	23.7	23.6	23.3	25.0	24.5	23.5
SAUSA300_1800	NA	NA	NA	23.0	22.8	23.0	NA	23.9	21.2
fumC SAUSA300_1801	NA	NA	22.1	NA	NA	20.0	20.1	24.8	NA
SAUSA300_1804	24.9	24.6	23.1	23.4	23.1	23.4	24.8	24.8	18.6
SAUSA300_1805	NA	NA	NA	NA	NA	NA	NA	18.5	NA
SAUSA300_1808	NA	NA	NA	18.5	18.3	19.0	NA	NA	NA
SAUSA300_1844	24.5	23.5	NA	23.5	23.5	24.4	23.5	23.0	24.0
hemL2 gsaB SAUSA300_1845	19.6	19.7	NA	22.2	23.5	22.7	23.5	22.9	17.1
SAUSA300_1848	NA	NA	NA	18.5	18.6	20.1	NA	NA	NA
SAUSA300_1851	NA	NA	NA	18.6	19.4	18.8	17.2	16.6	NA
recX SAUSA300_1854	NA	NA	NA	NA	17.7	NA	NA	16.8	NA
mgt SAUSA300_1855	NA	NA	NA	20.9	20.7	19.4	NA	NA	NA
SAUSA300_1856	22.2	22.0	NA	NA	NA	NA	NA	NA	NA
SAUSA300_1857	NA	NA	NA	NA	NA	NA	19.2	17.9	19.9
SAUSA300_1858	NA	NA	NA	21.4	21.8	22.4	NA	NA	NA
pepS SAUSA300_1860	NA	NA	NA	22.3	21.2	21.7	23.1	23.3	22.6
SAUSA300_1861	NA	NA	NA	20.9	19.9	20.9	NA	NA	NA
ptpA SAUSA300_1862	NA	NA	NA	22.4	22.6	23.6	NA	23.4	NA
SAUSA300_1863	NA	NA	NA	26.2	25.9	26.0	NA	24.3	NA
SAUSA300_1864	20.7	21.2	NA	20.3	21.0	20.6	21.0	20.8	21.5
vraR SAUSA300_1865	21.6	21.9	NA	24.8	24.1	25.0	NA	21.4	NA
vraS SAUSA300_1866	NA	NA	NA	18.3	19.7	19.6	17.0	19.2	NA
SAUSA300_1867	NA	NA	NA	NA	19.8	NA	NA	NA	NA
map SAUSA300_1869	NA	22.6	24.9	25.0	24.8	25.0	24.5	25.0	24.9
SAUSA300_1872	21.8	20.8	NA	22.3	22.0	22.2	NA	24.0	NA
SAUSA300_1873	NA	26.4	NA	24.6	24.6	24.6	NA	23.6	22.0
ftnA SAUSA300_1874	23.2	26.6	22.2	26.4	26.8	28.9	26.8	27.6	NA
rumA SAUSA300_1878	NA	NA	NA	20.6	20.0	21.6	NA	NA	NA
dagK SAUSA300_1879	NA	20.3	18.7	22.3	22.8	22.7	20.0	22.8	20.3
gatB SAUSA300_1880	25.7	26.1	25.8	27.3	27.2	28.3	28.1	27.8	28.8
gatA SAUSA300_1881	23.6	24.9	27.3	25.3	25.2	26.3	26.1	27.4	27.3
gatC SAUSA300_1882	NA	NA	NA	23.6	23.6	24.3	24.4	23.6	NA
SAUSA300_1884	NA	NA	NA	20.7	20.7	21.2	NA	22.5	NA
ligA SAUSA300_1885	NA	NA	NA	21.8	21.8	21.0	21.9	21.5	20.4
pcrA SAUSA300_1886	20.4	NA	NA	24.2	24.1	23.6	22.6	22.1	22.4
SAUSA300_1888	NA	NA	NA	20.3	20.0	21.0	21.8	19.2	18.0
purB SAUSA300_1889	NA	NA	NA	23.9	24.1	24.5	23.6	24.4	NA
SAUSA300_1892	NA	NA	NA	21.7	21.3	21.7	NA	NA	NA
nadE SAUSA300_1893	23.0	22.6	NA	21.7	22.2	22.1	NA	23.7	22.9
SAUSA300_1894	23.6	23.5	NA	25.2	25.3	24.5	25.9	24.1	24.4
SAUSA300_1895	22.1	21.4	NA	22.0	21.6	22.1	NA	NA	NA
pheA SAUSA300_1896	NA	NA	NA	19.1	18.7	NA	NA	NA	NA
SAUSA300_1899	NA	19.4	NA	21.0	20.2	21.5	20.8	19.4	NA

ppaC SAUSA300_1900	25.1	23.4	26.0	24.9	25.1	24.6	27.1	25.5	26.4
aldA2 SAUSA300_1901	21.3	20.9	NA	21.5	21.8	21.7	NA	21.2	NA
SAUSA300_1902	NA	NA	NA	21.6	20.9	22.1	20.1	21.9	NA
SAUSA300_1904	25.2	24.7	25.2	22.8	22.4	25.2	22.5	24.2	22.9
SAUSA300_1906	NA	NA	NA	19.8	20.0	20.0	18.7	21.1	NA
SAUSA300_1909	21.8	23.1	21.6	22.5	21.2	21.4	24.7	23.4	NA
SAUSA300_1911	20.1	21.3	NA	NA	NA	NA	22.8	20.8	NA
SAUSA300_1916	21.5	24.0	NA	23.7	24.0	23.7	21.7	22.2	23.4
SAUSA300_1928	NA	23.9	NA	NA	NA	NA	NA	NA	NA
SAUSA300_1965	NA	20.4	NA	20.2	20.4	22.3	NA	21.4	NA
SAUSA300_1969	NA	NA	NA	23.2	23.0	21.9	18.5	21.9	NA
SAUSA300_1970	21.5	19.9	NA	21.3	21.2	21.4	22.5	20.7	20.9
SAUSA300_1971	NA	NA	NA	NA	19.9	20.5	NA	NA	NA
SAUSA300_1974	22.0	23.3	NA	22.8	22.6	22.7	21.6	23.1	20.1
SAUSA300_1975	21.5	23.6	NA	NA	NA	NA	20.7	20.2	21.6
SAUSA300_1977	19.3	19.5	NA	NA	NA	NA	NA	NA	NA
groL groEL SAUSA300_1982	30.0	29.1	28.3	25.2	25.7	24.9	28.2	28.7	28.7
groS groES SAUSA300_1983	26.7	27.3	23.7	24.3	24.3	24.9	23.7	24.8	23.7
SAUSA300_1987	NA	NA	NA	NA	NA	NA	NA	19.9	NA
SAUSA300_1988	NA	NA	NA	22.7	22.1	NA	NA	NA	NA
agrA SAUSA300_1992	NA	NA	NA	21.5	22.6	23.0	NA	18.9	NA
SAUSA300_1993	NA	NA	NA	18.1	19.1	19.9	NA	NA	NA
scrB SAUSA300_1994	NA	NA	NA	21.1	20.8	20.4	NA	20.6	NA
rex SAUSA300_1999	NA	NA	NA	21.1	21.2	21.3	23.7	20.2	NA
vga SAUSA300_2000	NA	NA	NA	22.3	22.1	22.3	NA	20.7	NA
tsaD gcp SAUSA300_2002	21.3	NA	NA	20.7	20.6	20.3	21.3	20.8	NA
SAUSA300_2004	NA	NA	NA	20.3	19.8	20.8	NA	NA	NA
ilvD SAUSA300_2006	27.4	27.2	NA	NA	NA	NA	NA	NA	NA
ilvC SAUSA300_2009	20.8	20.5	NA	NA	NA	NA	NA	NA	NA
SAUSA300_2021	22.9	NA	NA	23.7	23.3	23.5	24.3	23.2	23.2
rpoF SAUSA300_2022	NA	NA	NA	20.0	19.4	20.4	NA	18.9	NA
rsbW SAUSA300_2023	22.9	23.3	NA	21.3	21.7	21.1	21.7	21.1	22.5
rsbV SAUSA300_2024	NA	NA	NA	23.2	22.6	21.6	NA	NA	NA
rsbU SAUSA300_2025	21.0	20.6	NA	21.3	21.1	21.7	NA	21.2	NA
mazF SAUSA300_2026	NA	NA	NA	18.1	18.5	19.7	NA	NA	NA
alr SAUSA300_2027	NA	NA	NA	22.7	23.0	22.9	NA	21.8	NA
acpS SAUSA300_2028	NA	NA	NA	NA	NA	20.2	NA	NA	NA
SAUSA300_2030	NA	NA	NA	NA	NA	NA	NA	18.2	NA
cshA SAUSA300_2037	25.0	24.0	NA	28.5	27.8	28.4	24.6	28.0	26.8
murF SAUSA300_2038	NA	NA	NA	23.0	23.3	23.9	23.3	22.6	22.6
ddl SAUSA300_2039	23.7	23.8	NA	24.6	24.7	25.0	22.3	25.1	23.5
SAUSA300_2040	NA	NA	NA	20.3	NA	20.4	NA	NA	NA
SAUSA300_2043	NA	NA	NA	NA	NA	NA	NA	19.3	NA
SAUSA300_2045	NA	NA	NA	NA	NA	NA	16.1	16.6	NA
yidC SAUSA300_2046	NA	NA	NA	22.9	22.6	22.2	23.2	22.4	NA
fabZ SAUSA300_2054	21.3	NA	NA	23.5	22.3	23.3	22.8	24.0	22.1
murA SAUSA300_2055	21.8	20.4	NA	24.4	23.5	24.5	24.4	25.7	23.5
atpC SAUSA300_2057	NA	NA	NA	25.0	24.6	25.5	NA	25.3	NA
atpD SAUSA300_2058	26.0	25.6	25.0	27.9	28.1	27.2	28.1	28.7	28.5
atpG SAUSA300_2059	23.3	24.0	NA	23.8	23.6	23.9	25.0	24.8	24.2
atpA SAUSA300_2060	26.0	26.6	NA	25.5	25.3	25.4	29.1	26.7	30.2
atpH SAUSA300_2061	NA	25.3	NA	23.9	23.5	23.7	25.3	25.7	21.7
atpF SAUSA300_2062	NA	NA	21.4	23.0	22.7	22.6	24.5	25.2	24.5
SAUSA300_2065	NA	NA	NA	22.7	22.0	22.7	21.2	22.6	20.9

upp SAUSA300_2066	25.1	26.2	23.4	25.4	26.0	26.5	24.1	26.1	24.4
glyA SAUSA300_2067	24.8	22.4	23.3	22.2	22.3	22.4	23.7	24.2	23.5
SAUSA300_2070	NA	NA	NA	19.0	19.4	20.3	NA	NA	NA
prfA SAUSA300_2072	25.1	24.4	23.1	25.5	25.4	24.9	22.7	24.2	21.5
tdk SAUSA300_2073	NA	NA	NA	23.3	23.1	23.6	NA	NA	NA
rpmE2 SAUSA300_2074	NA	NA	NA	27.8	27.5	27.8	NA	25.8	NA
rho SAUSA300_2075	24.7	22.8	NA	24.3	24.3	24.5	NA	24.9	NA
SAUSA300_2076	25.1	25.2	23.6	23.6	23.5	24.3	23.9	23.3	24.6
SAUSA300_2077	NA	NA	NA	NA	18.6	19.7	NA	NA	NA
murA SAUSA300_2078	21.9	22.5	21.2	22.4	23.0	22.8	22.4	23.9	23.3
fba SAUSA300_2079	27.5	25.8	23.2	26.6	26.5	27.6	26.5	27.6	29.7
pyrG SAUSA300_2081	25.1	24.5	NA	26.0	26.5	26.9	27.6	27.2	25.7
rpoE SAUSA300_2082	21.1	22.4	NA	21.0	21.0	20.7	22.2	19.4	NA
SAUSA300_2083	NA	NA	NA	18.6	NA	18.0	NA	NA	NA
coaW coaA SAUSA300_2084	19.3	18.2	NA	20.6	21.2	20.7	21.6	20.9	17.8
SAUSA300_2085	NA	22.0	NA	24.2	23.8	24.8	NA	21.5	NA
luxS SAUSA300_2088	NA	NA	NA	23.3	23.0	24.3	NA	21.0	NA
pdp pyn SAUSA300_2089	28.2	27.7	28.6	26.4	26.1	26.7	25.7	25.2	24.1
deoC SAUSA300_2090	25.4	25.8	25.0	26.6	26.8	27.3	25.4	26.9	25.4
deoD SAUSA300_2091	23.7	24.7	NA	24.2	24.3	25.4	24.5	25.0	28.7
dps SAUSA300_2092	24.6	24.3	28.7	NA	NA	NA	24.2	23.1	20.6
SAUSA300_2093	20.6	NA	NA	NA	NA	NA	NA	19.5	NA
SAUSA300_2094	NA	NA	NA	21.3	21.2	21.5	21.8	22.7	20.8
manA SAUSA300_2096	NA	17.7	NA	NA	NA	21.0	NA	NA	NA
SAUSA300_2097	24.8	24.0	23.4	22.8	22.7	23.3	21.5	22.7	NA
arsR SAUSA300_2098	NA	NA	NA	NA	NA	NA	NA	17.4	18.7
SAUSA300_2100	26.1	NA	26.1	25.4	25.5	24.8	24.7	25.3	22.4
SAUSA300_2102	21.7	21.5	NA	22.9	23.0	23.3	NA	NA	NA
glmS SAUSA300_2104	25.9	25.8	NA	27.5	27.7	28.2	27.2	26.7	27.8
mtlF SAUSA300_2105	19.5	18.9	NA	NA	NA	20.3	NA	NA	NA
mtlD SAUSA300_2108	22.0	21.3	NA	21.2	20.2	20.3	22.3	18.6	NA
glmM SAUSA300_2111	25.2	24.8	NA	27.1	27.3	26.9	26.4	26.1	26.1
SAUSA300_2112	21.0	21.0	22.7	21.4	21.1	21.6	21.7	21.2	22.5
dacA SAUSA300_2113	19.4	22.0	NA	22.2	21.1	20.6	20.7	19.9	NA
SAUSA300_2125	20.8	NA	NA	23.2	23.5	23.2	23.4	24.5	22.8
SAUSA300_2130	24.2	23.3	NA	24.3	23.5	24.3	24.7	24.8	25.1
SAUSA300_2132	25.7	25.2	25.0	19.7	19.9	18.5	22.5	23.6	NA
SAUSA300_2133	NA	NA	NA	20.9	21.1	20.4	20.4	20.0	NA
SAUSA300_2136	23.2	23.6	NA	NA	NA	NA	20.6	19.9	NA
SAUSA300_2140	NA	NA	NA	NA	NA	NA	NA	17.1	NA
asp23 SAUSA300_2142	31.3	30.7	29.8	26.0	25.3	26.7	27.4	26.9	29.5
SAUSA300_2144	24.9	26.0	22.6	22.1	21.9	22.6	23.6	22.4	NA
SAUSA300_2146	21.6	19.8	NA	20.4	20.1	20.6	20.8	20.5	NA
SAUSA300_2147	23.9	21.9	NA	22.1	22.0	22.4	NA	23.9	NA
SAUSA300_2148	NA	NA	NA	18.7	18.6	NA	19.0	19.8	NA
SAUSA300_2156	NA	NA	NA	21.2	20.8	20.0	NA	NA	NA
SAUSA300_2159	NA	NA	NA	19.1	19.4	19.1	NA	NA	NA
hysA SAUSA300_2161	NA	NA	NA	NA	NA	NA	NA	16.7	NA
SAUSA300_2163	NA	NA	NA	19.3	18.9	18.8	18.3	20.6	NA
SAUSA300_2165	NA	NA	NA	23.8	24.0	24.2	23.6	23.3	23.4

alsS SAUSA300_2166	NA	NA	NA	27.3	28.0	28.0	24.6	25.9	NA
rpsI SAUSA300_2171	NA	21.6	NA	23.2	22.5	24.8	27.2	26.0	23.7
rplM SAUSA300_2172	25.0	26.7	23.5	28.1	27.9	27.5	28.8	28.0	29.0
truA SAUSA300_2173	NA	NA	NA	20.1	19.6	19.4	NA	NA	NA
ecfT SAUSA300_2174	NA	NA	NA	NA	NA	18.0	NA	17.7	NA
ecfA1 cbiO1 SAUSA300_2176	NA	NA	NA	23.5	24.1	24.3	NA	NA	NA
rplQ SAUSA300_2177	26.9	28.8	NA	25.7	25.4	25.4	28.7	26.9	29.5
rpoA SAUSA300_2178	27.4	NA	25.8	26.0	25.0	26.1	28.2	26.4	29.2
rpsK SAUSA300_2179	22.0	22.4	23.4	20.8	21.4	21.3	25.1	23.8	29.8
rpsM SAUSA300_2180	NA	NA	NA	23.3	23.6	23.7	27.9	27.7	28.4
rpmJ SAUSA300_2181	20.1	21.6	NA	25.1	25.0	24.8	NA	22.4	NA
adk SAUSA300_2183	25.3	23.7	22.4	26.9	26.4	25.8	25.4	27.1	24.9
rplO SAUSA300_2185	26.6	25.7	24.4	27.5	26.6	27.7	27.0	27.6	26.5
rpmD SAUSA300_2186	22.1	25.0	NA	25.1	24.9	24.7	25.4	24.8	25.0
rpsE SAUSA300_2187	27.9	26.8	23.7	27.0	26.8	27.0	28.4	29.8	29.3
rplR SAUSA300_2188	28.5	26.9	23.4	27.3	27.0	27.2	28.7	27.1	27.0
rplF SAUSA300_2189	28.0	28.9	NA	26.3	25.1	25.8	26.6	27.3	27.8
rpsH SAUSA300_2190	28.2	28.5	26.8	24.7	24.9	25.4	26.8	26.4	27.2
rpsZ rpsN1 SAUSA300_2191	NA	NA	NA	22.1	22.4	22.1	NA	NA	NA
rplE SAUSA300_2192	24.8	25.6	25.0	28.0	27.7	28.2	29.8	28.9	30.6
rplX SAUSA300_2193	22.6	24.8	28.9	25.5	25.9	25.7	26.2	23.8	26.2
rplN SAUSA300_2194	25.0	24.2	28.0	26.1	25.9	25.6	28.9	26.3	27.2
rpsQ SAUSA300_2195	26.3	25.3	NA	25.5	25.0	24.9	27.4	25.6	NA
rpmC SAUSA300_2196	23.9	28.4	25.3	26.9	26.7	26.0	25.6	25.0	29.9
rplP SAUSA300_2197	25.7	24.9	24.3	27.3	27.1	26.7	27.8	27.9	25.7
rpsC SAUSA300_2198	26.1	27.2	25.5	27.7	27.3	27.8	28.0	28.5	28.1
rplV SAUSA300_2199	27.6	27.1	25.8	27.4	27.1	27.5	28.7	27.6	27.8
rpsS SAUSA300_2200	23.3	25.1	NA	23.8	23.9	24.4	24.9	21.1	28.4
rplB SAUSA300_2201	25.4	24.0	30.0	27.4	27.4	27.2	29.1	28.7	29.7
rplW SAUSA300_2202	24.4	23.4	26.0	27.2	27.0	27.0	28.9	24.9	28.2
rplD SAUSA300_2203	24.8	24.1	21.9	25.1	24.9	24.9	27.2	28.3	29.6
rplC SAUSA300_2204	22.8	22.2	NA	26.2	26.4	26.1	22.5	26.1	28.4
rpsJ SAUSA300_2205	26.9	28.7	23.6	26.7	27.0	26.5	28.5	26.2	30.2
SAUSA300_2206	NA	NA	NA	21.3	20.1	20.3	NA	20.8	NA
topB SAUSA300_2208	NA	NA	NA	22.6	22.0	21.6	NA	NA	NA
SAUSA300_2213	21.7	23.7	NA	22.1	21.9	21.2	22.7	22.2	21.6
femX fmhB SAUSA300_2214	26.2	NA	NA	24.2	23.8	24.1	24.0	25.0	23.5
moaA SAUSA300_2219	NA	NA	NA	NA	NA	NA	18.4	18.0	NA
moaD SAUSA300_2221	NA	NA	NA	20.0	19.8	NA	NA	NA	NA
moaE SAUSA300_2222	NA	NA	NA	19.1	19.8	20.2	NA	21.2	NA
mobB SAUSA300_2223	NA	NA	NA	NA	NA	19.2	NA	NA	NA
moeA SAUSA300_2224	24.6	22.3	23.5	24.1	24.1	24.3	23.5	22.2	26.3
moaB SAUSA300_2226	NA	NA	NA	21.7	21.7	21.5	21.2	22.0	NA
moeB SAUSA300_2227	NA	NA	NA	NA	NA	NA	19.0	19.3	NA
modC SAUSA300_2228	NA	NA	NA	19.3	19.5	19.7	NA	18.1	NA
modA SAUSA300_2230	22.6	23.2	NA	22.4	22.6	23.1	21.4	22.9	NA
SAUSA300_2235	23.5	23.3	21.0	21.5	21.4	21.5	24.2	24.1	26.1

SAUSA300_2236	21.2	NA	NA	21.9	21.9	22.1	22.1	21.5	23.1
ureE SAUSA300_2241	NA	NA	NA	21.4	21.5	22.8	NA	18.4	NA
sarR SAUSA300_2245	24.3	22.7	NA	23.8	23.6	23.7	24.6	23.0	20.4
ssaA SAUSA300_2249	NA	NA	NA	23.9	23.9	22.9	25.6	25.9	25.4
SAUSA300_2251	22.6	23.1	NA	22.6	22.7	22.8	23.8	23.7	NA
SAUSA300_2252	NA	NA	NA	NA	NA	NA	NA	20.0	NA
SAUSA300_2254	24.7	24.5	25.7	25.8	25.5	25.7	24.3	25.9	23.0
SAUSA300_2255	21.8	NA	NA	21.2	21.4	21.0	21.8	22.6	NA
SAUSA300_2257	22.8	20.7	NA	NA	NA	20.1	NA	21.7	NA
SAUSA300_2258	NA	NA	NA	25.4	25.6	24.7	25.1	25.5	23.7
SAUSA300_2259	25.1	25.0	NA	23.6	24.3	24.7	23.0	24.9	NA
SAUSA300_2260	NA	NA	NA	NA	16.9	16.6	NA	NA	NA
SAUSA300_2261	NA	NA	NA	20.5	20.4	19.1	NA	NA	NA
SAUSA300_2264	NA	24.2	NA	20.6	19.6	21.9	NA	21.7	NA
SAUSA300_2265	NA	NA	NA	NA	NA	NA	19.2	NA	NA
SAUSA300_2267	NA	NA	NA	NA	NA	NA	19.8	17.8	NA
SAUSA300_2269	NA	NA	NA	NA	NA	20.1	NA	NA	NA
SAUSA300_2272	NA	NA	NA	22.0	21.5	21.6	NA	21.0	NA
SAUSA300_2273	NA	NA	NA	NA	NA	20.1	NA	16.6	NA
SAUSA300_2275	20.7	20.2	NA	NA	NA	NA	21.7	20.9	NA
hutG SAUSA300_2281	NA	NA	NA	17.1	17.0	NA	18.7	18.9	NA
lyrA SAUSA300_2282	NA	23.1	NA	22.2	22.3	22.6	20.3	23.3	21.0
rpiA SAUSA300_2283	23.2	22.6	NA	23.6	23.7	24.3	NA	22.7	NA
SAUSA300_2284	NA	NA	NA	21.2	21.0	21.5	NA	NA	NA
galM SAUSA300_2285	NA	NA	NA	18.9	19.0	20.1	19.8	20.6	NA
SAUSA300_2287	NA	NA	NA	19.4	20.0	20.1	NA	17.6	NA
SAUSA300_2289	23.6	21.8	NA	22.0	21.8	21.4	22.0	22.0	19.0
SAUSA300_2290	NA	NA	NA	NA	NA	NA	NA	17.8	NA
fni SAUSA300_2292	NA	NA	NA	21.7	21.8	22.0	21.4	20.4	NA
SAUSA300_2295	NA	NA	NA	NA	NA	NA	18.8	18.7	NA
SAUSA300_2296	19.7	NA	NA	22.5	22.2	21.6	22.3	21.8	18.0
SAUSA300_2298	NA	NA	NA	23.8	23.7	23.6	22.1	20.8	NA
SAUSA300_2299	NA	NA	NA	25.0	25.2	24.6	20.6	23.2	NA
tcaA SAUSA300_2302	NA	NA	NA	19.4	19.6	NA	NA	NA	NA
mqo SAUSA300_2312	NA	NA	NA	22.0	22.2	21.3	NA	NA	NA
SAUSA300_2313	25.4	24.4	NA	22.1	21.6	21.7	23.7	26.1	26.4
SAUSA300_2314	NA	NA	NA	22.2	22.5	22.0	NA	20.9	NA
SAUSA300_2315	23.1	23.7	NA	23.7	22.5	22.8	24.9	24.4	NA
SAUSA300_2317	21.4	21.7	NA	22.8	22.4	22.9	24.4	22.4	26.2
SAUSA300_2327	22.8	22.7	NA	21.8	20.0	22.4	NA	19.8	NA
SAUSA300_2328	26.3	23.5	27.6	23.7	23.7	22.8	22.4	20.2	22.9
gltT SAUSA300_2329	NA	NA	NA	NA	NA	NA	NA	19.8	NA
SAUSA300_2330	NA	NA	NA	23.1	22.9	22.9	NA	22.6	NA
sarZ SAUSA300_2331	21.2	24.0	NA	22.6	22.6	22.5	NA	23.4	NA
SAUSA300_2332	19.0	NA	NA	18.1	18.8	19.1	19.9	18.7	19.7
narT narK SAUSA300_2333	NA	NA	NA	21.2	22.2	22.3	NA	NA	NA
nreC SAUSA300_2337	NA	NA	NA	23.4	24.0	22.4	NA	21.8	19.7
nreB SAUSA300_2338	NA	NA	19.2	21.1	20.9	20.1	NA	NA	19.9
narI SAUSA300_2340	NA	NA	NA	23.3	23.0	22.6	18.0	NA	NA
narJ SAUSA300_2341	NA	NA	NA	25.7	25.6	25.6	NA	NA	NA
narH SAUSA300_2342	22.3	22.7	NA	27.4	27.0	27.5	22.6	24.4	22.7
SAUSA300_2343	24.3	26.8	22.4	28.7	29.3	28.6	25.4	25.4	24.3
SAUSA300_2344	NA	NA	NA	20.2	20.0	17.8	NA	NA	NA
nirD SAUSA300_2345	25.2	25.2	NA	25.4	25.4	24.6	NA	NA	NA
nirB SAUSA300_2346	23.0	23.4	25.2	27.5	27.4	28.2	23.9	23.8	NA
nirR SAUSA300_2347	NA	NA	NA	19.8	20.0	20.8	NA	NA	NA
SAUSA300_2348	NA	NA	NA	NA	NA	NA	16.1	19.0	NA
SAUSA300_2349	NA	NA	NA	21.6	21.2	21.3	NA	NA	NA
SAUSA300_2351	NA	NA	NA	20.9	20.6	19.7	21.6	21.9	NA
SAUSA300_2352	NA	NA	NA	NA	NA	18.8	NA	17.6	NA
SAUSA300_2357	NA	25.6	NA	24.3	24.4	24.7	NA	23.6	NA

SAUSA300_2358	NA	NA	NA	22.6	22.3	22.3	NA	NA	NA
SAUSA300_2359	NA	23.0	NA	25.0	25.1	25.0	25.4	27.1	26.9
gpmA SAUSA300_2362	26.1	24.9	NA	26.4	26.3	27.1	25.4	26.0	NA
sbi SAUSA300_2364	19.9	20.4	NA	NA	NA	NA	NA	NA	NA
hlgA SAUSA300_2365	22.2	21.9	23.9	NA	NA	NA	NA	NA	NA
SAUSA300_2370	NA	NA	NA	NA	NA	NA	NA	18.2	NA
SAUSA300_2374	21.4	20.9	NA	NA	NA	NA	NA	NA	NA
SAUSA300_2375	21.4	21.1	NA	NA	NA	NA	NA	NA	NA
SAUSA300_2377	NA	NA	NA	20.0	19.8	20.0	NA	NA	NA
SAUSA300_2378	20.6	NA	NA	20.6	20.4	20.1	19.5	19.1	NA
SAUSA300_2381	NA	NA	NA	20.7	20.1	20.2	NA	NA	NA
SAUSA300_2387	NA	NA	NA	NA	NA	NA	19.2	18.0	NA
panE SAUSA300_2388	NA	NA	NA	21.1	NA	19.6	16.4	20.7	NA
opuCa SAUSA300_2393	NA	NA	NA	23.0	23.4	23.7	NA	21.4	NA
SAUSA300_2394	NA	NA	NA	20.6	20.3	20.4	NA	NA	NA
SAUSA300_2397	NA	NA	NA	17.8	17.1	17.2	NA	NA	NA
SAUSA300_2398	NA	NA	NA	NA	NA	NA	NA	18.6	18.2
SAUSA300_2399	NA	NA	NA	NA	NA	19.7	NA	NA	NA
SAUSA300_2400	23.9	NA	NA	24.1	24.0	23.7	24.3	21.0	NA
SAUSA300_2403	21.7	NA	NA	17.5	18.0	NA	22.3	17.1	NA
SAUSA300_2404	NA	NA	NA	21.6	21.4	20.8	NA	NA	NA
opp-1A SAUSA300_2411	27.1	25.6	27.3	23.1	22.6	23.5	19.3	22.9	NA
SAUSA300_2422	21.4	21.0	NA	19.0	19.1	20.2	NA	20.3	21.4
SAUSA300_2431	NA	NA	NA	NA	NA	NA	20.2	16.1	NA
SAUSA300_2432	20.2	19.2	NA	19.6	20.0	20.5	NA	NA	NA
pgcA SAUSA300_2433	NA	NA	NA	22.1	22.8	23.3	22.7	21.5	NA
gtaB galU SAUSA300_2439	20.7	NA	NA	23.8	23.7	24.0	24.0	22.6	NA
fnbB SAUSA300_2440	NA	NA	NA	NA	NA	NA	17.2	20.0	NA
fnbA SAUSA300_2441	NA	NA	NA	NA	NA	NA	19.2	18.2	NA
gntP SAUSA300_2442	20.7	21.5	21.9	NA	NA	NA	NA	NA	NA
gntK SAUSA300_2443	NA	NA	NA	22.3	22.6	23.3	NA	NA	NA
SAUSA300_2446	NA	NA	NA	22.1	21.7	21.8	NA	NA	NA
SAUSA300_2448	20.0	NA	NA	NA	NA	NA	NA	NA	NA
SAUSA300_2452	NA	NA	NA	NA	NA	NA	NA	18.0	NA
SAUSA300_2453	21.2	19.2	NA	20.1	20.2	19.6	NA	NA	NA
fbp SAUSA300_2455	20.3	20.4	NA	19.7	19.6	20.7	NA	19.3	NA
SAUSA300_2459	NA	NA	NA	NA	NA	NA	NA	18.5	NA
SAUSA300_2460	NA	NA	NA	20.3	20.4	20.6	NA	NA	NA
SAUSA300_2461	NA	NA	NA	NA	NA	NA	NA	18.3	NA
SAUSA300_2462	NA	NA	NA	19.6	19.7	20.1	NA	NA	NA
ddh SAUSA300_2463	23.7	23.4	22.1	20.0	20.8	19.8	26.0	27.5	26.1
srtA SAUSA300_2467	NA	NA	21.1	23.0	22.7	23.3	NA	23.0	NA
sdaAA SAUSA300_2469	NA	NA	NA	18.8	18.7	NA	NA	NA	NA
SAUSA300_2473	NA	NA	NA	21.4	21.2	22.1	NA	20.1	19.7
glcB SAUSA300_2476	NA	22.8	NA	NA	NA	NA	23.1	23.6	NA
cidC SAUSA300_2477	23.1	23.0	20.4	21.8	21.9	22.3	21.9	24.6	22.9
SAUSA300_2480	NA	20.8	NA	NA	NA	NA	NA	NA	NA
SAUSA300_2483	NA	NA	NA	22.4	22.9	23.5	21.5	23.7	22.6
SAUSA300_2484	25.9	24.5	NA	26.3	26.3	26.3	25.6	26.3	28.0
clpL SAUSA300_2486	30.1	29.9	27.5	25.9	25.4	25.8	26.1	25.4	24.9
rocA SAUSA300_2491	22.2	22.0	NA	20.7	20.3	20.3	21.2	19.8	NA
SAUSA300_2492	NA	NA	NA	21.4	20.8	20.3	NA	NA	NA
copA SAUSA300_2494	NA	NA	NA	20.6	21.0	20.8	NA	NA	NA
copZ SAUSA300_2495	NA	NA	NA	19.4	19.3	19.0	NA	22.5	NA

SAUSA300_2496	NA	NA	NA	19.1	19.3	19.3	NA	NA	NA
crtP SAUSA300_2501	21.6	20.2	NA	NA	NA	NA	NA	NA	NA
SAUSA300_2504	NA	NA	NA	22.0	22.3	22.3	21.2	21.7	NA
isaA SAUSA300_2506	NA	NA	NA	22.0	21.9	22.1	25.6	22.1	26.6
SAUSA300_2513	22.1	21.8	NA	21.9	21.3	21.7	22.3	20.4	NA
SAUSA300_2516	NA	NA	NA	19.4	NA	NA	NA	NA	NA
SAUSA300_2517	NA	NA	NA	20.7	20.5	20.8	19.8	20.5	NA
SAUSA300_2518	23.7	24.0	27.7	22.0	21.9	22.2	23.8	25.7	24.0
SAUSA300_2523	NA	NA	NA	24.7	NA	24.3	NA	NA	NA
SAUSA300_2525	NA	NA	NA	19.3	19.3	20.3	NA	NA	NA
pyrD SAUSA300_2526	22.3	21.6	NA	23.8	23.6	22.5	25.4	24.6	27.5
SAUSA300_2527	NA	NA	NA	NA	NA	NA	NA	19.6	NA
SAUSA300_2528	NA	NA	NA	18.1	18.4	18.8	18.3	21.6	NA
SAUSA300_2529	20.6	20.6	NA	NA	NA	20.5	NA	NA	NA
panD SAUSA300_2532	NA	NA	NA	NA	NA	NA	20.4	21.7	19.8
panC SAUSA300_2533	NA	NA	NA	NA	NA	NA	NA	18.8	NA
panB SAUSA300_2534	NA	21.7	NA	23.2	22.8	23.2	24.1	23.9	22.3
ldh2 SAUSA300_2537	29.3	29.6	29.9	28.7	28.8	29.4	26.5	25.8	29.4
fda SAUSA300_2540	29.4	28.9	31.2	27.5	27.3	27.8	28.8	27.9	28.4
mqo SAUSA300_2541	26.2	25.8	26.1	28.6	28.7	28.6	28.3	28.8	30.0
SAUSA300_2542	22.6	NA	NA	22.7	22.9	22.6	23.6	22.7	23.4
betA SAUSA300_2545	NA	NA	NA	18.2	18.1	NA	NA	NA	NA
SAUSA300_2547	NA	NA	NA	NA	NA	NA	NA	NA	18.5
nrdD SAUSA300_2551	NA	21.8	NA	23.2	23.4	23.3	NA	21.4	NA
SAUSA300_2554	NA	NA	NA	21.3	20.8	21.7	NA	NA	NA
SAUSA300_2560	NA	NA	NA	20.5	21.0	20.3	NA	NA	NA
SAUSA300_2562	NA	NA	NA	22.3	22.4	23.1	NA	17.0	NA
estA SAUSA300_2564	NA	NA	NA	21.2	21.7	22.0	NA	NA	NA
clfB SAUSA300_2565	NA	NA	NA	NA	21.6	23.2	23.4	25.4	23.4
arcC2 SAUSA300_2567	20.8	NA	NA	NA	18.1	NA	18.7	19.7	NA
arcB argF SAUSA300_2569	23.0	24.0	NA	21.4	21.7	23.5	NA	19.3	NA
arcA SAUSA300_2570	NA	NA	NA	19.5	19.4	22.0	NA	NA	21.4
isaB SAUSA300_2573	NA	17.1	NA	NA	NA	NA	22.7	21.4	NA
SAUSA300_2576	NA	NA	NA	18.8	18.9	21.7	NA	17.5	NA
SAUSA300_2578	NA	NA	NA	21.1	21.0	21.2	20.6	22.7	20.3
SAUSA300_2580	NA	NA	NA	19.1	19.6	20.0	NA	NA	20.4
SAUSA300_2581	19.8	19.7	NA	NA	NA	NA	NA	NA	NA
sraP SAUSA300_2589	NA	NA	NA	16.0	NA	15.0	NA	NA	NA
tetR SAUSA300_2599	21.2	NA	NA	19.5	19.4	20.3	20.2	19.5	18.7
SAUSA300_2620	24.0	23.2	20.6	19.6	19.4	19.5	NA	19.8	NA
SAUSA300_2627	19.3	NA	NA	24.0	24.2	23.9	NA	20.8	NA
SAUSA300_2639	NA	NA	NA	24.1	22.6	21.6	NA	NA	NA
SAUSA300_2643	NA	NA	NA	20.8	20.2	20.0	21.8	22.9	NA
rsmG SAUSA300_2644	NA	NA	NA	21.4	21.3	21.4	NA	20.8	NA
mnmG gidA SAUSA300_2645	23.3	NA	NA	25.4	25.4	24.8	24.6	25.2	20.2
mnmE trmE SAUSA300_2646	22.1	20.7	NA	20.8	21.0	20.8	22.1	22.7	20.1
traI SAUSA300_pUSA030018	NA	NA	NA	25.3	25.3	NA	NA	NA	NA
psmA4 SAUSA300_0424.1	NA	NA	NA	20.4	20.8	21.4	NA	NA	NA
psmA1 SAUSA300_0424.4	NA	NA	NA	19.8	20.7	20.4	NA	NA	NA

Supplemental Plasmid Annotations

pSS20_tet insert (pmRFPmars backbone)
TetR +tet promoter
his tag
MetRS
cut site

pmRFPmars backbone ---
*GTCGAC*GGTATCGATAACTCGACATCTTGGTTACCGTGAAGTTACCATCACGGAAAAAGGTTA
TGCTGCTTTTAAGACCCACTTTCACATTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAG
GCCGAATAAGAAGGCTGGCTCTGCACCTTGGTGATCAAATAATTCGATAGCTTGTCGTAATAA
TGGCGGCATACTATCAGTAGTAGGTGTTTCCCTTTCTTCTTTAGCGACTTGATGCTCTTGATCTT
CCAATACGCAACCTAAAGTAAAATGCCCCACAGCGCTGAGTGCATATAATGCATTCTCTAGTG
AAAAACCTTGTTGGCATAAAAAGGNTAATTGATTTTCGAGAGTTTCATACTGTTTTTCTGTAGG
CCGTGTACCTAAATGTACTTTTGCTCCATCGCGATGACTTAGTAAAGCACATCTAAAACTTTTA
GCGTTATTACGTAAAAAATCTTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTATGGTGC
CTATCTAACATCTCAATGGCTAAGGCGTCGAGCAAAGCCCGCTTATTTTTTACATGCCAATACA
ATGTAGGCTGCTCTACACCTAGCTTCTGGGCGAGTTTACGGGTTGTTAAACCTTCGATTCCGAC
CTCATTAAGCAGCTCTAATGCGCTGTTAATCACTTTACTTTTATCTAATCTAGACATCATTAATT
CCTCCTTTTTGTTGACACTCTATCATTGATAGAGTTATTTGTCAAACTAGTTTTTTATTTGGATC
CCCTCGAGTTCATGAAAAACTAAAAAAAATATTGACACTCTATCATTGATAGAGTATAATTAA
AATAAGCTTGAT*GGTACC*AGATCCTAGGAAAGGAGGATGATAAAATGAGAGGATCGCATCAC
CATCACCATCACGGATCGATGACTCAAGTAGCAAAAAAAATTTTAGTAACTTGTGCAAATCCT
TATGCAAATGGTAGTATTCATTTAGGTCACATGTTAGAACATATTCAAGCAGATGTTTGGGTTC
GTTATCAACGTATGAGAGGTCATGAAGTTAATTTTATTTGTGCAGATGATGCTCATGGTACACC
AATTATGTTAAAAGCACAACAATTAGGTATTACTCCTGAACAAATGATTGGTGAAATGAGTCA
AGAACATCAAACAGATTTTGCAGGTTTTAATATTTCTTATGATAATTATCATTCTACTCATTCA
GAAGAAAATCGTCAATTATCTGAATTAATTTATTCAAGATTAAAAGAAAATGGTTTTATTAAA
AATCGTACAATTTCACAATTATATGATCCAGAAAAAGGAATGTTTTTACCTGATAGATTTGTAA
AAGGTACTTGTCCAAAATGTAAATCACCTGATCAATATGGTGATAATTGTGAAGTTTGTGGTG
CAACATATAGTCCAACTGAATTAATTGAACCTAAAAGTGTTGTATCTGGTGCTACACCAGTTAT
GCGTGATTCAGAACATTTCTTTTTCGATTTACCTTCTTTTTCAGAAATGTTACAAGCATGGACT
AGAAGTGGTGCTTTACAAGAACAAGTAGCAAATAAAATGCAAGAATGGTTTGAATCAGGTTT
ACAACAATGGGATATTAGTCGTGATGCACCATATTTTGGTTTTGAAATTCCAAATGCTCCTGGT
AAATATTTTTATGTATGGTTAGATGCACCAATTGGTTTAATGGGTTCATTTAAAAATTTATGTG
ATAAAAGAGGTGATAGTGTTTCTTTTGATGAATATTGGAAAAAAGATAGTACAGCAGAATTAT
ATCATTTTATTGGTAAAGATATTGTATATTTCTTATCTTTATTTTGGCCAGCAATGTTAGAAGGT
AGTAATTTTCGTAAACCTTCTAATTTATTTGTTCATGGTTATGTTACAGTAAATGGTGCAAAAA
TGTCAAAAAGTAGAGGTACATTTATTAAAGCTTCTACTTGGTTAAATCATTTTGATGCAGATTC
ATTACGTTATTATTATACAGCTAAATTATCAAGTAGAATTGATGATATTGATTTAAATTTAGAA
GATTTTGTTCAACGTGTAAATGCAGATATTGTAAATAAAGTTGTAAATTTAGCATCTCGTAATG
CTGGTTTTATTAATAAAAGATTTGATGGTGTTTTAGCATCAGAATTAGCTGATCCACAATTATA
TAAAACATTTACTGATGCAGCTGAAGTTATTGGTGAAGCATGGGAATCTCGTGAATTTGGTAA
AGCTGTAAGAGAAATTATGGCATTAGCTGATTTAGCAAATCGTTATGTTGATGAACAAGCACC
ATGGGTTGTAGCTAAACAAGAAGGTAGAGATGCAGATTTACAAGCTATTTGTAGTATGGGTAT
TAATTTATTTCGTGTTTTAATGACATATTTAAAACCAGTATTACCTAAATTAACTGAAAGAGCA
GAAGCATTTTTAAATACAGAATTAACTTGGGATGGTATTCAACAACCATTATTAGGTCATAAA
GTTAATCCTTTTAAAGCATTATATAATCGTATTGATATGAGACAAGTTGAAGCATTAGTAGAA
GCTTCAAAAGAAGAAGTTAAAGCAGCTGCAGCTCCAGTAACAGGTCCTTTAGCAGATGATCCA
ATTCAAGAAACAATTACTTTTGATGATTTTGCTAAAGTTGATTTACGTGTAGCATTAATTGAAA
ATGCTGAATTTGTAGAAGGTAGTGATAAATTATTAAGATTAACATTAGATTTAGGTGGTGAAA

AACGTAATGTATTTTCAGGTATTAGAAGTGCTTATCCAGATCCTCAAGCATTAATTGGTAG
ACATACAATTATGGTAGCAAATTTAGCTCCACGTAAAATGAGATTTGGTATTTCAGAAGGAAT
GGTTATGGCAGCTGGTCCTGGTGGTAAAGATATTTTTTTATTAAGTCCAGATGCAGGTGCTAAA
CCAGGTCATCAAGTTAAATAAACCTTGTAAAACAAGGTTTCTCTGCAGTTAAGCC*GAATTC*---
pmRFPmars backbone

pSS20_hprk insert (pmRFPmars backbone) **"+NLL"**
hprK promoter
his tag
MetRS

pmRFPmars backbone ---
*gtcgac*GTGAGGCTCAACGTGTGAAACTTGCATCTGAACTTCATAAACGTTCAACTGGTAAATCT
ATTTATATCCTAGATGAACCGACAACAGGGTTACATGTTGACGATATTAGTAGATTATTAAAA
GTATTAAACCGATTAGTTGAAAATGGTGATACTGTTGTAATTATTGAACATAACCTAGATGTTA
TCAAAACAGCAGACTATATTATAGACTTAGGTCCTGAAGGTGGTAGTGGCGGTGGTACTATTG
TTGCGACTGGCACACCCGAAGATATTGCTCAGACAAAGTCATCATATACAGGAAAGTATTTAA
AAGAAGTACTTGAACGAGATAAACAAAATACTGAAGATAAATAAGATTAAAAGAAGTGAAGG
ATGTTATAAATTTATCCTTCGCTTCTTTTTATTAATTTAGTAATGAATAGTAGAAAGAAAAGAT
GCGTAAAAAGAATTATGTTAAGATAGGGTCAATCTAGAGTAGTTAAACATAAATCGAACTGG
GAGTGGGACAGAAATGATAAAGAATCACTAATGATTTATTATGTAGTGGTTCTTTGTCATTAG
CCACAGCTATTGTGTACTTAAAAATAGGAATGCATGAGTGCAACTCATGCATAAGAAATACTA
ATTTCTAAAGAAAAAGTATTTCTTTATGTTGGGGCCCCGCCAACTTGCATTGTTTGTAGAATTT
CTTTTCGAAATTCTTTATGTTGGGGCCCCGCCAACTTGCATTGTTTGTAGAATTTCTTTTCGAAA
TTCTTTATGTTGGGGCCCCGCCAACTAATTCCAATATATCATTGTAGAGCTTAGGTCATTGATT
TTTGGCTCGGACTTTTATGGCGATATGAACCATGTAAATTAAGCAAGCAATAAATTAATGATT
GATATTGACTTGTAAAATAATAACAATAATGAACAATTAATATTTATTTTAGCTTTTCAATGTA
GATTGGTGTTATATTTTTGATATGATAAGAAGAGATGTAAGAGTAGGGATAAATACAATTGAG
GTGAACCCATGTTAACGACAGAAAAACTAGTTGAAACATTAAAGTTAGATTTAATCGCTGGTG
AAGAAGGAC*ggtacC*AGATCCTAGGAAAGGAGGATGATAAAATGAGAGGATCGCATCACCATC
ACCATCACGGATCGATGACTCAAGTAGCAAAAAAAATTTTAGTAACTTGTGCAAATCCTTATG
CAAATGGTAGTATTCATTTAGGTCACATGTTAGAACATATTCAAGCAGATGTTTGGGTTCGTTA
TCAACGTATGAGAGGTCATGAAGTTAATTTTATTTGTGCAGATGATGCTCATGGTACACCAATT
ATGTTAAAAGCACAACAATTAGGTATTACTCCTGAACAAATGATTGGTGAAATGAGTCAAGAA
CATCAAACAGATTTTGCAGGTTTTAATATTTCTTATGATAATTATCATTCTACTCATTCAGAAG
AAAATCGTCAATTATCTGAATTAATTTATTCAAGATTAAAAGAAAATGGTTTTATTAAAAATC
GTACAATTTCACAATTATATGATCCAGAAAAAGGAATGTTTTTACCTGATAGATTTGTAAAAG
GTACTTGTCCAAAATGTAAATCACCTGATCAATATGGTGATAATTGTGAAGTTTGTGGTGCAA
CATATAGTCCAACTGAATTAATTGAACCTAAAAGTGTTGTATCTGGTGCTACACCAGTTATGCG
TGATTCAGAACATTTCTTTTTCGATTTACCTTCTTTTTCAGAAATGTTACAAGCATGGACTAGA
AGTGGTGCTTTACAAGAACAAGTAGCAAATAAAATGCAAGAATGGTTTGAATCAGGTTTACAA
CAATGGGATATTAGTCGTGATGCACCATATTTTGGTTTTGAAATTCCAAATGCTCCTGGTAAAT
ATTTTTATGTATGGTTAGATGCACCAATTGGTTTAATGGGTTCATTTAAAAATTTATGTGATAA
AAGAGGTGATAGTGTTTCTTTTGATGAATATTGGAAAAAAGATAGTACAGCAGAATTATATCA
TTTTATTGGTAAAGATATTGTATATTTCTTATCTTTATTTTGGCCAGCAATGTTAGAAGGTAGTA
ATTTTCGTAAACCTTCTAATTTATTTGTTCATGGTTATGTTACAGTAAATGGTGCAAAAATGTC
AAAAAGTAGAGGTACATTTATTAAAGCTTCTACTTGGTTAAATCATTTTGATGCAGATTCATTA
CGTTATTATTATACAGCTAAATTATCAAGTAGAATTGATGATATTGATTTAAATTTAGAAGATT
TTGTTCAACGTGTAAATGCAGATATTGTAAATAAAGTTGTAAATTTAGCATCTCGTAATGCTGG
TTTTATTAATAAAAGATTTGATGGTGTTTTAGCATCAGAATTAGCTGATCCACAATTATATAAA
ACATTTACTGATGCAGCTGAAGTTATTGGTGAAGCATGGGAATCTCGTGAATTTGGTAAAGCT
GTAAGAGAAATTATGGCATTAGCTGATTTAGCAAATCGTTATGTTGATGAACAAGCACCATGG

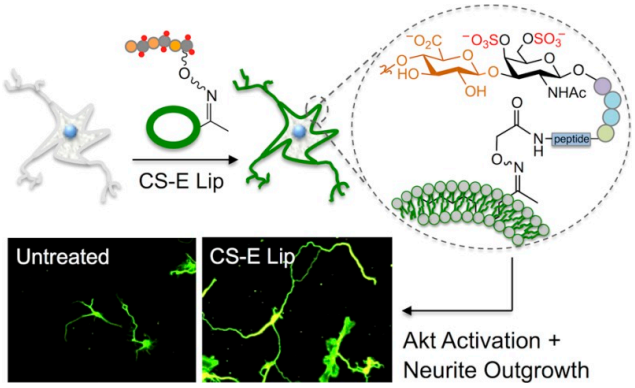
GTTGTAGCTAAACAAGAAGGTAGAGATGCAGATTTACAAGCTATTTGTAGTATGGGTATTA
ATTTATTTCGTGTTTTAATGACATATTTAAAACCAGTATTACCTAAATTAACTGAAAGAGCAGA
AGCATTTTTAAATACAGAATTAACTTGGGATGGTATTCAACAACCATTATTAGGTCATAAAGTT
AATCCTTTTAAAGCATTATATAATCGTATTGATATGAGACAAGTTGAAGCATTAGTAGAAGCT
TCAAAAGAAGAAGTTAAAGCAGCTGCAGCTCCAGTAACAGGTCCTTTAGCAGATGATCCAATT
CAAGAAACAATTACTTTTGATGATTTTGCTAAAGTTGATTTACGTGTAGCATTAATTGAAAATG
CTGAATTTGTAGAAGGTAGTGATAAATTATTAAGATTAACATTAGATTTAGGTGGTGAAAAAC
GTAATGTATTTTCAGGTATTAGAAGTGCTTATCCAGATCCTCAAGCATTAATTGGTAGACATAC
AATTATGGTAGCAAATTTAGCTCCACGTAAAATGAGATTTGGTATTTCAGAAGGAATGGTTAT
GGCAGCTGGTCCTGGTGGTAAAGATATTTTTTTATTAAGTCCAGATGCAGGTGCTAAACCAGG
TCATCAAGTTAAATAAACCTTGTAAAACAAGGTTTCTCTGCAGTTAAGCCGAATTC
TGTAAAACAAGGTTTCTCTGCAGTTAAGCC*GAATTC* ---pmRFPmars backbone

Appendix B

DIRECTING NEURONAL SIGNALING THROUGH CELL-SURFACE GLYCAN ENGINEERING

Pulsipher, A., Griffin, M. E., **Stone, S. E.**, Brown, J. M., Hsieh-Wilson, L. C. "Directing neuronal signaling through cell-surface glycan engineering." J. Am. Chem. Soc. 2014, 136 (19), 6794-6797. DOI: 10.1021/ja5005174

TOC



Abstract

The ability to tailor plasma membranes with specific glycans may enable the control of signaling events that are critical for proper development and function. We report a method to modify cell surfaces with specific sulfated chondroitin sulfate (CS) glycosaminoglycans using chemically modified liposomes. Neurons engineered to display CS-E-enriched polysaccharides exhibited increased activation of neurotrophin-mediated signaling pathways and enhanced axonal growth. This approach provides a facile, general route to tailor cell membranes with biologically active glycans and demonstrates the potential to direct important cellular events through cell-surface glycan engineering.

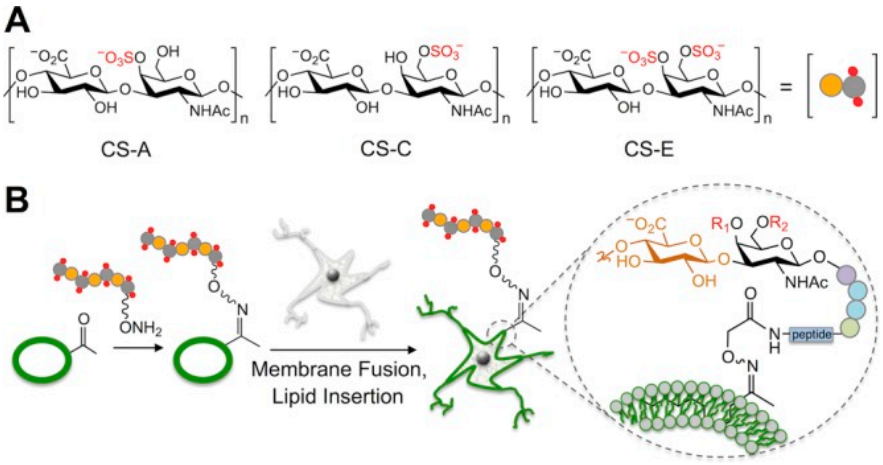


Figure 1: (A) CS polysaccharide structures used in this study. $n \approx 110$. (B) Strategy to remodel cell surfaces with CS GAGs and control signaling pathways.

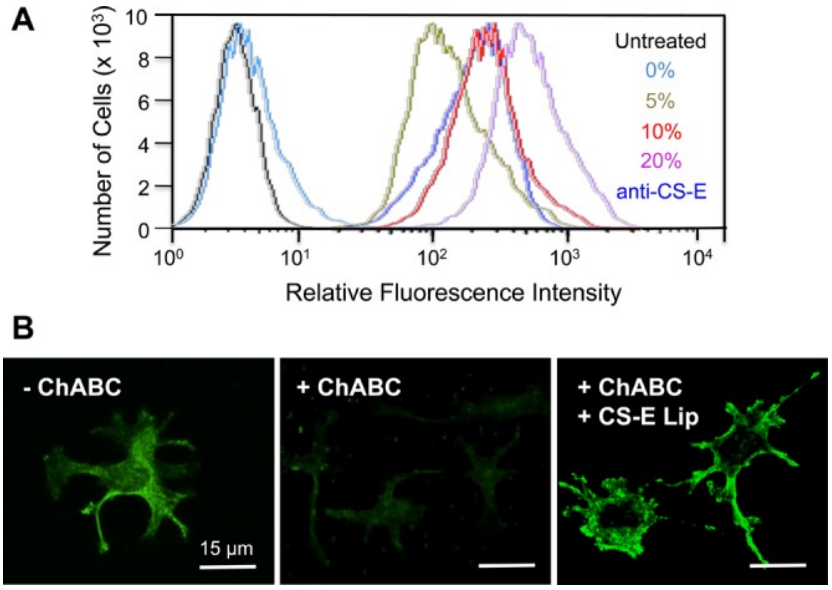


Figure 2: Controlled cell-surface display of CS polysaccharides and fluorophores. (A) FACS analysis of PC12 cells treated with liposomes presenting varying amounts (0–20%) of AF488-hyd. (B) Immunofluorescence detection of CS-E (green) on PC12 cells treated with or without chondroitinase (ChABC) and CS-E-functionalized liposomes as indicated.

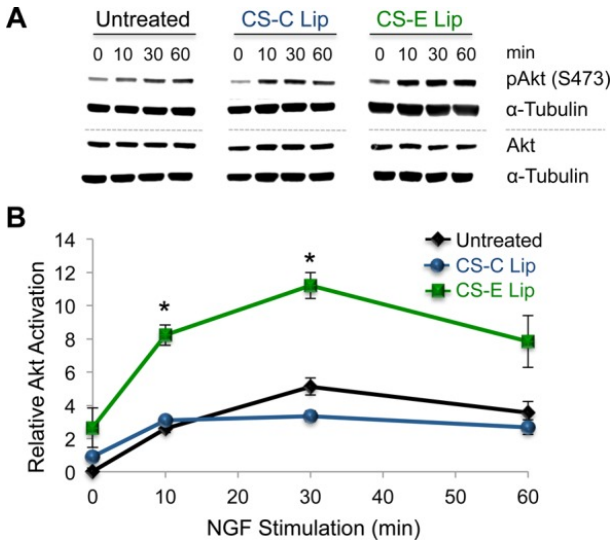


Figure 3: Presentation of CS-E polysaccharides on neuronal cell surfaces enhances NGF-mediated Akt activation. (A) Representative Western blots and (B) quantitation of Akt activation in neurons displaying CS-C- or CS-E-enriched polysaccharides. Akt activation was normalized against total Akt levels at each time point and compared to untreated neurons. Data represent mean ± SEM (*$P < 0.05$) from at least three experiments.

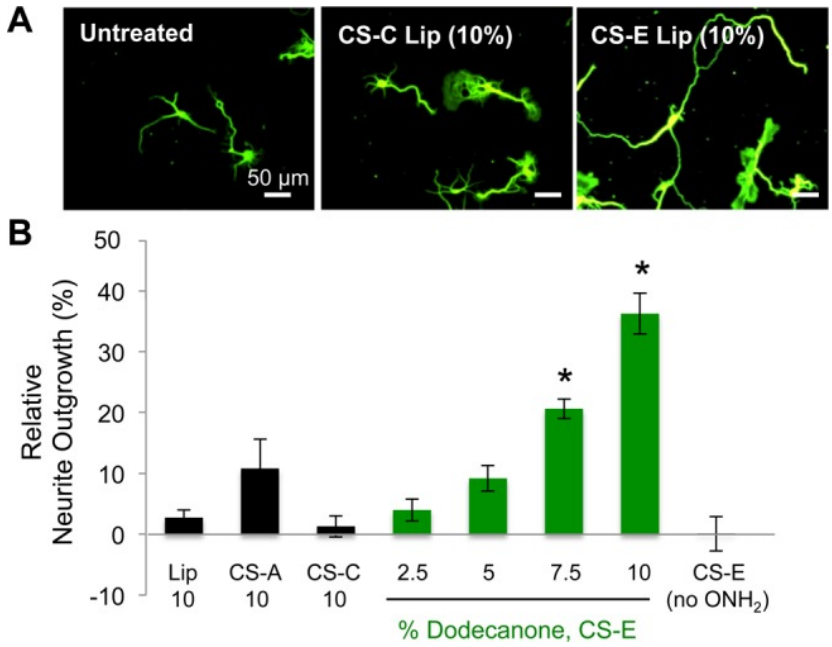


Figure 4: Presentation of CS-E polysaccharides on neuronal cell surfaces stimulates neurite outgrowth. (A) Representative images and (B) quantification of neurite outgrowth for neurons treated with unmodified liposomes or liposomes displaying CS-A-, CS-C-, or CS-E-enriched polysaccharides. Neurite outgrowth was normalized and plotted relative to untreated neurons. Data represent mean ± SEM (*$P < 0.05$) from at least three experiments.

Appendix C

LONG-LIVED ENGINEERING OF GLYCANS TO DIRECT STEM CELL FATE

Pulsipher, A., Griffin, M. E., **Stone, S. E.**, Hsieh-Wilson, L. C. "Long-lived engineering of glycans to direct stem cell fate." Angew. Chem. Int. Ed. **2015**, 54, 1466-1470. DOI: 10.1002/anie.201409258

Abstract

Glycans mediate many critical, long-term biological processes such as stem cell differentiation. However, few methods are available for the sustained remodeling of cells with specific glycan structures. We report a new strategy that enables the long-lived presentation of defined glycosaminoglycans on cell surfaces using HaloTag proteins (HTPs) as anchors. By controlling the sulfation patterns of heparan sulfate (HS) on pluripotent embryonic stem cell (ESC) membranes, we demonstrate that specific glycans cause ESCs to undergo accelerated exit from self-renewal and differentiation into neuronal cell types. Thus, the stable display of glycans on HTP scaffolds provides a powerful, versatile means to direct key signaling events and biological outcomes such as stem cell fate.

Keywords: embryonic stem cells, heparan sulfate, cell-surface engineering, stem cell differentiation, cell signaling

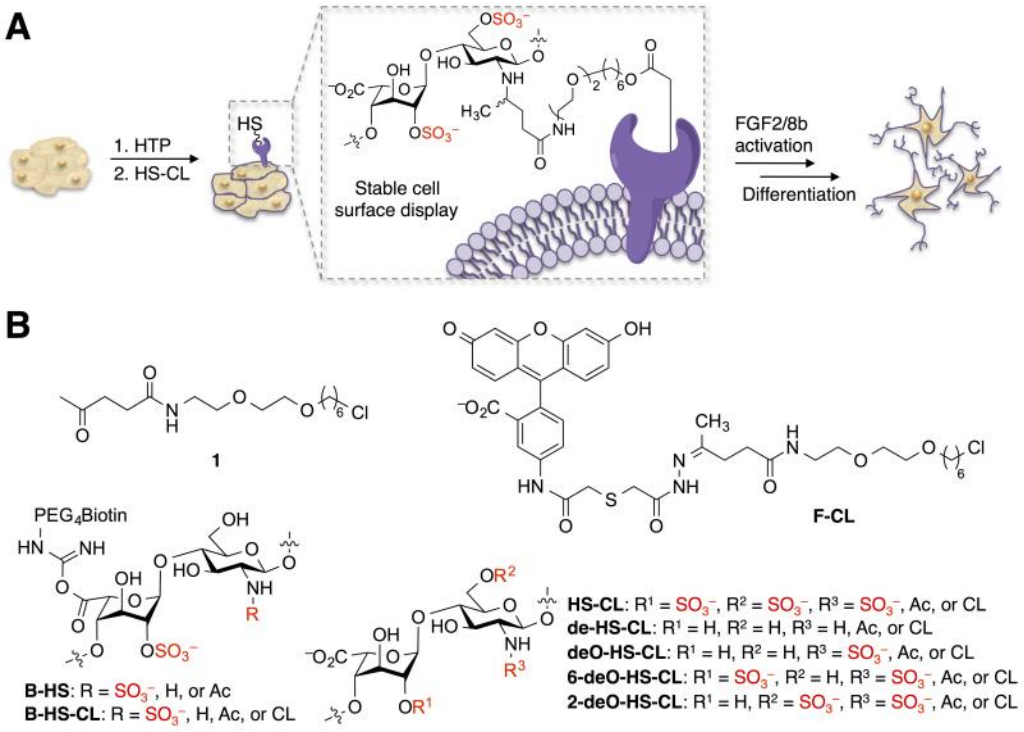


Figure 1: A) Strategy for presenting HS GAGs on cell membranes using HTP anchors to direct stem cell differentiation. B) Molecules used in this study.

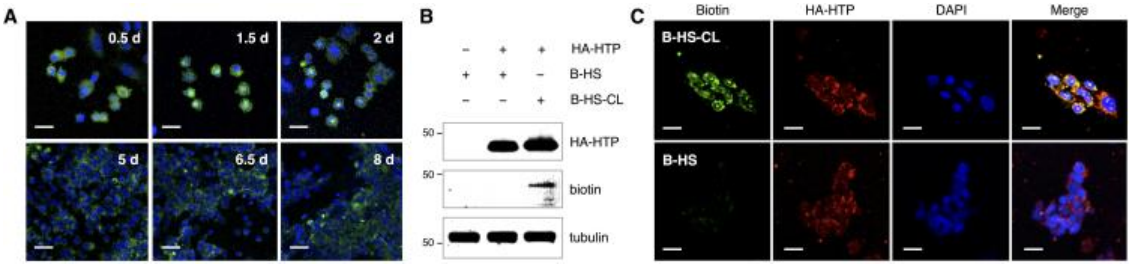


Figure 2: Extended cell-surface display by HTP anchoring. A) CHO cells stably expressing HTP were functionalized with a single treatment of F-CL (shown in green) and imaged over 8 days. Cell nuclei were co-stained at each time point with DAPI (shown in blue). B) Western blot detection and C) fluorescence imaging of hemagglutinin (HA)-tagged HTP and biotinylated HS. Stably transfected CHO cells were labeled with biotinylated HS with or without the chloroalkane linker (B-HS-CL or B-HS, respectively). Tubulin was used as a control for equal protein loading in B. Scale bars represent 20 μm.

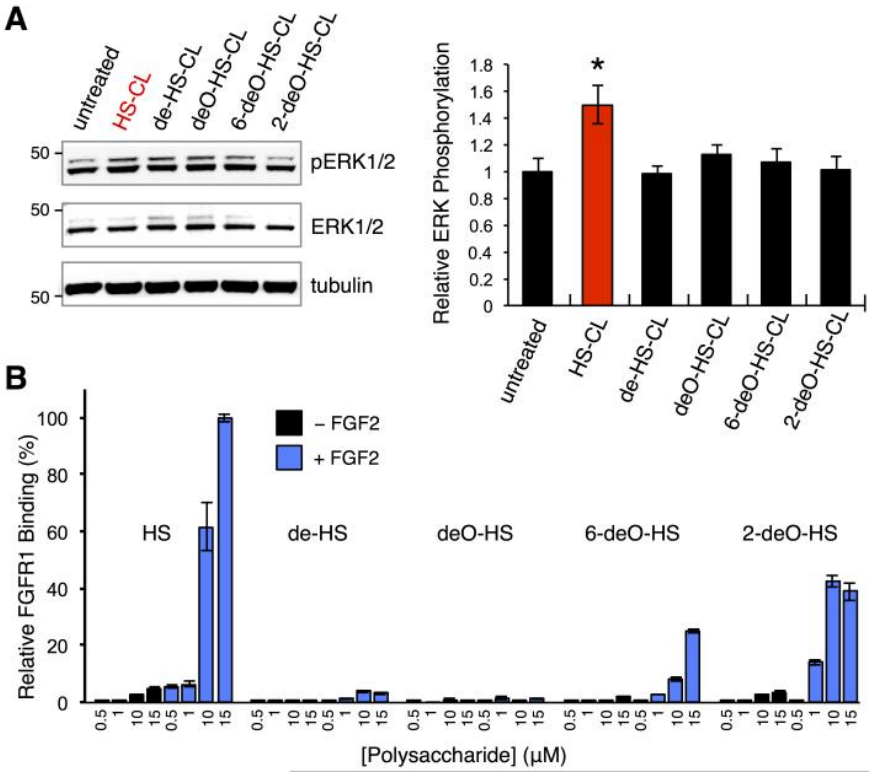


Figure 3: Cell-surface presentation of HS GAGs on ESCs induces FGF2-mediated ERK1/2 activation. A) Representative immunoblots (left) and quantification (right) of ERK1/2 phosphorylation levels in ESCs remodeled with the indicated HS GAGs and stimulated with FGF2. Phospho-ERK levels were normalized with respect to total ERK levels for each condition and compared to untreated ESCs. Tubulin was used as a control for equal protein loading. Data represent the mean ± S.E.M. (*$P < 0.05$) from three experiments. B) FGFR1-Fc binding to glycan microarrays in the presence (blue bars) or absence (black bars) of FGF2. Data represent the mean ± S.E.M. from ten replicate microarray spots.

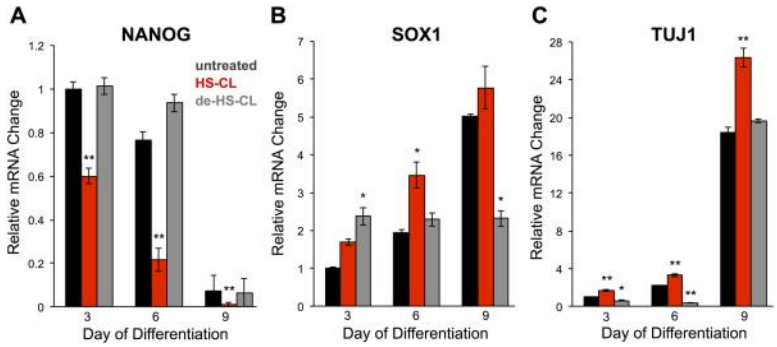


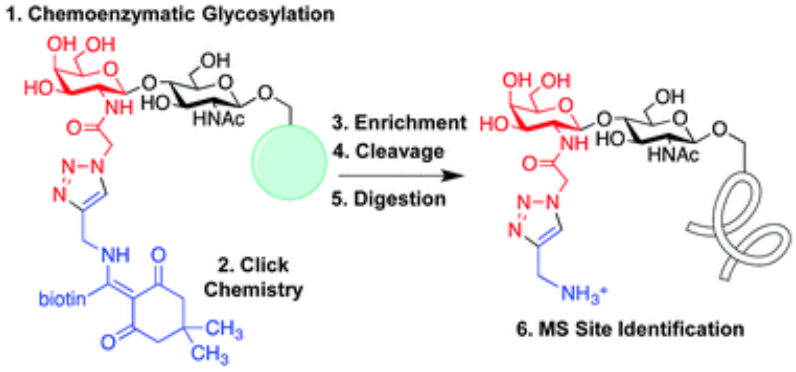
Figure 4: Remodeling the glycocalyx of ESCs with highly sulfated HS induces accelerated self-renewal exit, neural lineage commitment, and differentiation into mature, neuronal cells. qRT-PCR quantification of mRNA levels for A) pluripotent marker NANOG, B) neuroectoderm marker SOX1, and C) neuronal marker TUJ1. Data points were normalized to the housekeeping genes GAPDH and SDHA for cross comparison and to the untreated condition at day 3 for data presentation. Values represent the mean ± S.E.M. (*$P < 0.05$, **$P < 0.01$ when compared to the untreated control at each time point) from two independent experiments.

Appendix D

COMPREHENSIVE MAPPING OF *O*-GLCNAC MODIFICATION SITES USING CHEMICALLY CLEAVABLE TAG

Griffin, M. E., Jensen, E. H., Mason, D. E., Jenkins, C. L., **Stone, S. E.**, Peters, E. C., Hsieh-Wilson, L. C. "Comprehensive mapping of *O*-GlcNAc modification sites using a chemically cleavable tag." Mol. Biosys., **2016**, 12, 1756-1759. DOI: 10.1039/c6mb00138f

TOC



Abstract

The post-translational modification of serine or threonine residues of proteins with a single *N*-acetylglucosamine monosaccharide (*O*-GlcNAcylation) is essential for cell survival and function. However, relatively few *O*-GlcNAc modification sites have been mapped due to the difficulty of enriching and detecting *O*-GlcNAcylated peptides from complex samples. Here we describe an improved approach to quantitatively label and enrich *O*-GlcNAcylated proteins for site identification. Chemoenzymatic labelling followed by copper(I)-catalysed azide–alkyne cycloaddition (CuAAC) installs a new mass spectrometry (MS)-compatible linker designed for facile purification of *O*-GlcNAcylated proteins from cell lysates. The linker also allows subsequent quantitative release of *O*-GlcNAcylated proteins for downstream MS analysis. We validate the approach by unambiguously identifying several established *O*-GlcNAc sites on the proteins α-crystallin and *O*-GlcNAc transferase (OGT), as well as discovering new, previously unreported sites on OGT. Notably, these novel sites on OGT lie in key functional domains of the protein, underscoring how this site identification method may reveal important biological insights into protein activity and regulation.

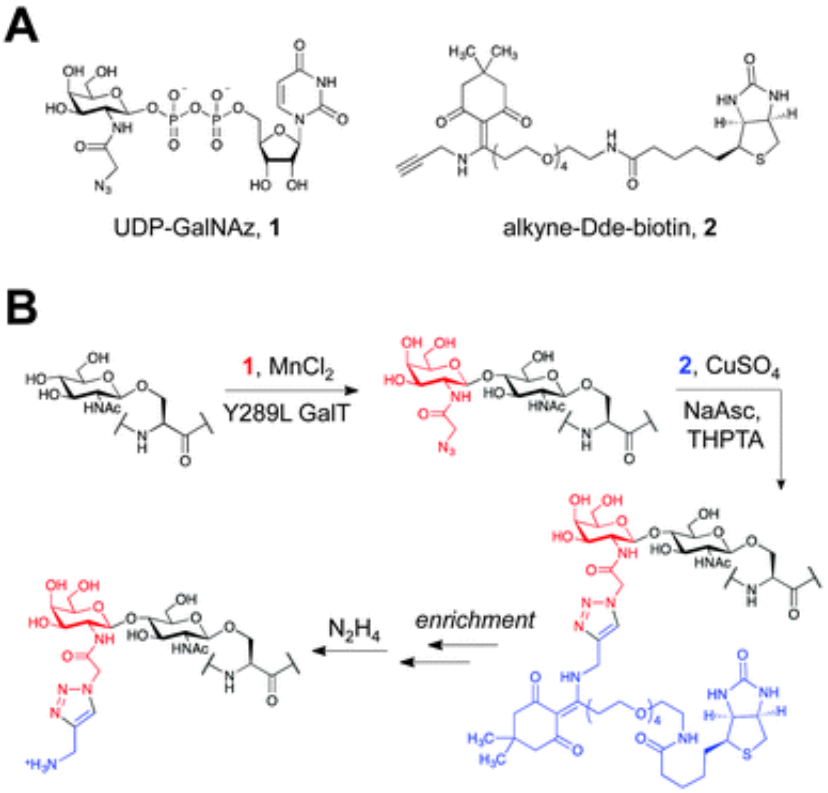


Figure 1: (A) Chemicals used in the labelling protocol. (B) Schematic of *O*-GlcNAc protein enrichment and elution using the two-step chemoenzymatic/CuAAC labelling protocol.

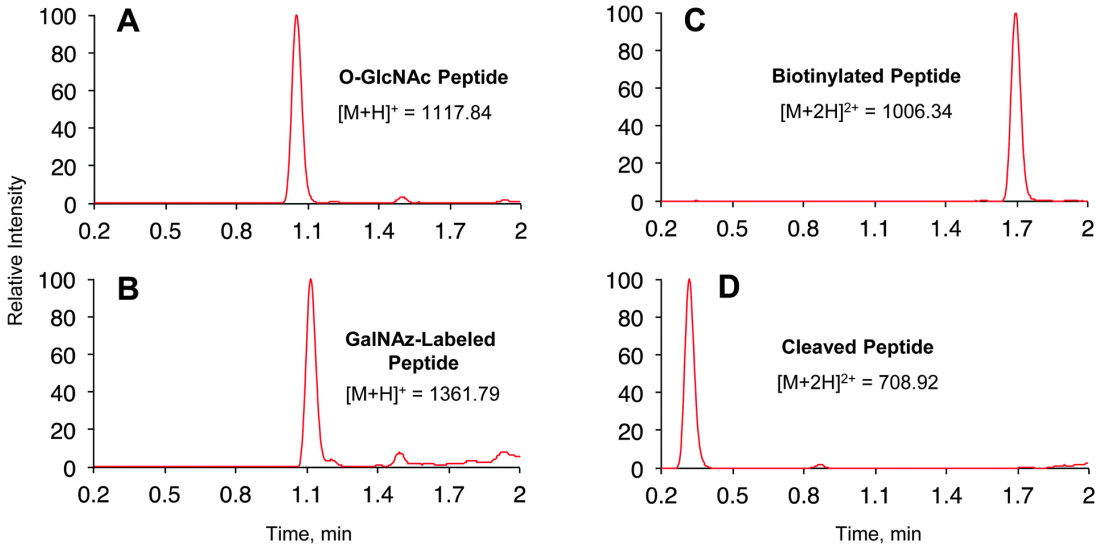


Figure 2: Labelling and cleavage reactions proceed quantitatively. Reverse phase LC-MS analysis of *O*-GlcNAc peptide labelling reactions at (A) time 0, (B) 16 h after addition of **1** and Y289L GalT, (C) 1 h after CuAAC with **2**, and (D) 1 h after cleavage with 2% aqueous hydrazine. See ESI† for

experimental details. (A) and (B) show base peak chromatograms. (C) and (D) show extracted ion chromatograms of the starting material and product within ±1 *m*/*z* of calculated values.

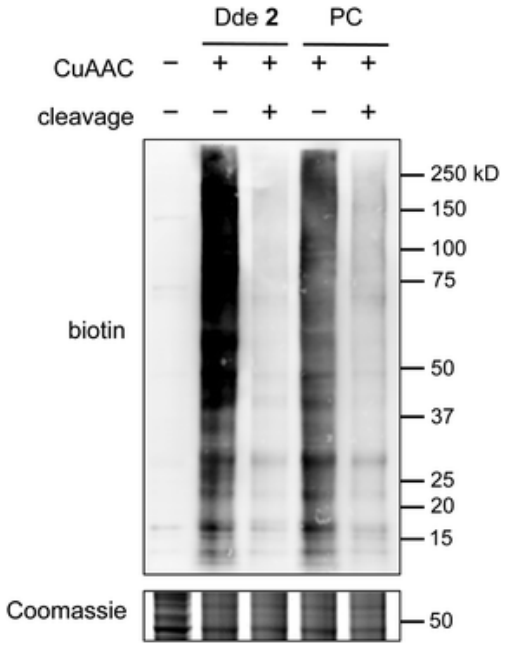


Figure 3: Alkyne-Dde-biotin linker **2** outperforms a widely used alkyne-photocleavable-biotin linker (PC). Protein lysates labelled with **2** show higher biotin signal after labelling (lane 2 *vs*. lane 4) and lower residual signal after cleavage (lane 3 *vs*. lane 5) compared to the PC linker. See Fig. S2 (ESI†) for full Coomassie gel.

Protein	Peptide sequence	Site(s)	Mascot ion score	Mascot delta ion score	Method
α-Crystallin A	AIPVSREEKPSSAPSS	Ser-162	24.9	23.5	ETD
sOGT	ISPTFADAYSNMGNTLK	Ser-10*/Thr-12*	46.5	—	ETD
sOGT	ISPTFADAYSNMGNTLK	Ser-20*	21.6	13.6	ETD
sOGT	ISPTFADAYSNMGNTLK	Ser-10*/Thr-12*, Ser-20*	38.4	—	ETD
sOGT	EMQDVQGALQCYTR	Thr-38	41.8	35.0	CID
sOGT	AIQINPAFADAHSNLASIHKDSGNIPEAIASYR	Ser-52*	53.5	7.9	ETD
sOGT	AIQINPAFADAHSNLASIHKDSGNIPEAIASYR	Ser-56*	56.8	15.7	ETD
sOGT	LYLQMWEHYAAGNKPDHMIKPVEVTESA	Thr-662	33.1	8.0	ETD

Table 1: *O*-GlcNAc sites identified following labelling with **2**, Neutravidin affinity purification, and hydrazine-mediated elution. Sites and regions of modification are denoted in red or red underline, respectively. Novel site identifications are marked by an asterisk.

Appendix E

BIOORTHOGONAL NONCANONICAL AMINO ACID TAGGING (BONCAT) ENABLES TIME-RESOLVED ANALYSIS OF PROTEIN SYNTHESIS IN NATIVE PLANT TISSUE

Glenn, W. S., **Stone, S. E.**, Ho, S. H., Sweredoski, M. J., Moradian, A., Hess, S., Bailey-Serres, J., Tirrell, D. A. "Bioorthogonal noncanonical amino acid tagging (BONCAT) enables time-resolved analysis of protein synthesis in native plant tissue." Plant Phys. **2017**, 173 (3), 1543-1553. DOI: 10.1104/pp.16.01762

Summary

Pulsing the noncanonical amino acid azidohomoalanine into *Arabidopsis thaliana* seedlings enables in-gel visualization, physical enrichment, and identification of newly synthesized proteins.

Abstract

Proteomic plasticity undergirds stress responses in plants, and understanding such responses requires accurate measurement of the extent to which proteins levels are adjusted to counter external stimuli. Here, we adapt bioorthogonal noncanonical amino acid tagging (BONCAT) to interrogate protein synthesis in vegetative *Arabidopsis thaliana* seedlings. BONCAT relies on the translational incorporation of a noncanonical amino acid (ncAA) probe into cellular proteins. In this study, the probe is the methionine surrogate azidohomoalanine (Aha), which carries a reactive azide moiety in its amino acid side chain. The azide handle in Aha can be selectively conjugated to dyes and functionalized beads to enable visualization and enrichment of newly synthesized proteins. We show that BONCAT is sensitive enough to detect *Arabidopsis* proteins synthesized within a 30-min interval defined by an Aha pulse, and that the method can be used to detect proteins made under conditions of light stress, osmotic shock, salt stress, heat stress and recovery from heat stress. We further establish that BONCAT can be coupled to tandem liquid chromatography-mass spectrometry (LC-MS) to identify and quantify proteins synthesized during heat stress and recovery from heat stress. Our results are consistent with a model in which, upon the onset of heat stress, translation is rapidly reprogrammed to enhance the synthesis of stress mitigators and is again altered during recovery. All experiments were carried out with commercially available reagents, highlighting the accessibility of the BONCAT method to researchers interested in stress responses as well as translational and post-translational regulation in plants.

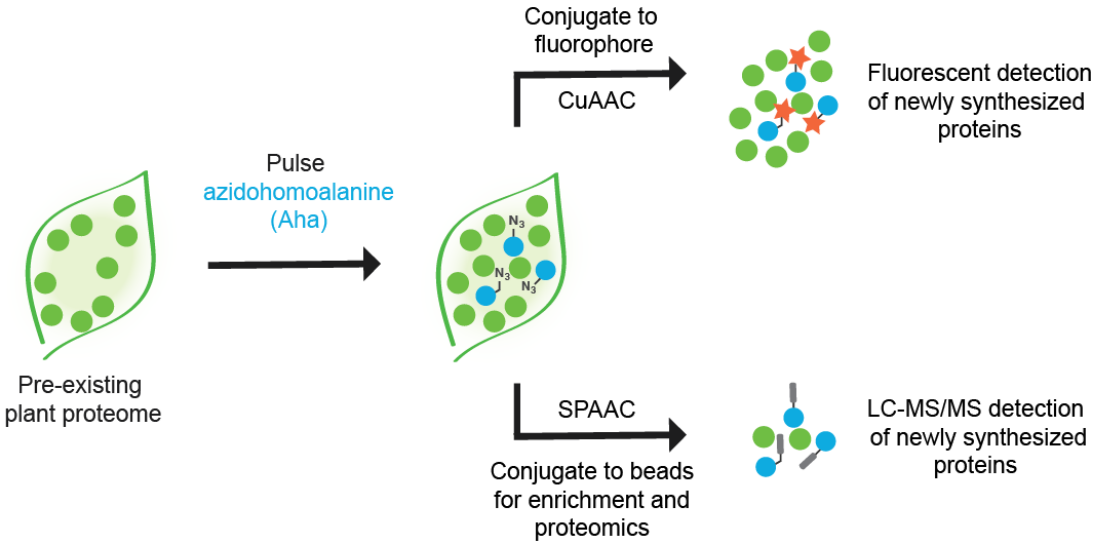


Figure 1: Scheme of BONCAT in native plant tissues. Azidohomoalanine (Aha) is pulsed into aerial tissues where it can be incorporated into nascent proteins. The azide enables conjugation to fluorophores or beads for visualization or enrichment, respectively. Copper-catalyzed azide-alkyne cycloaddition (CuAAC) was used to conjugate TAMRA alkyne to nascent proteins. Strain-promoted azide-alkyne cycloaddition (SPAAC; a biocompatible 'click reaction') was employed to conjugate nascent proteins to beads for enrichment.

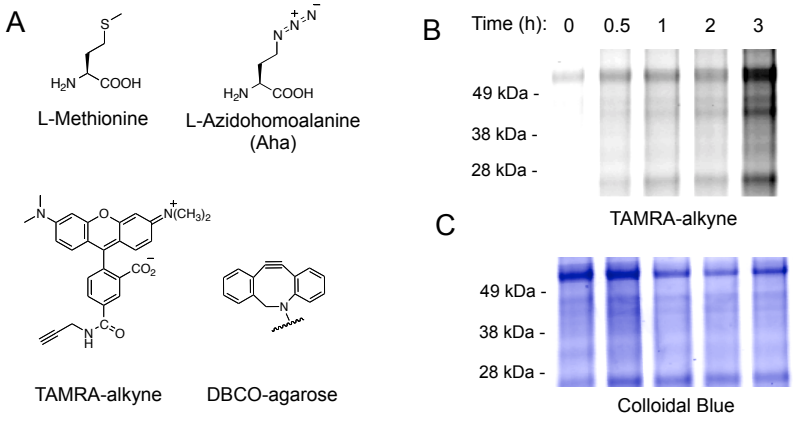


Figure 2: **A**. Probes used in this study. L-Azidohomoalanine (Aha) is a methionine surrogate replete with an azide moiety, which renders labeled proteins amenable to cycloaddition with fluorescent alkynyl probes (5,6-TAMRA alkyne) and strained cyclooctyne reagents (DBCO-agarose). **B**. Time course of Aha incorporation into nascent proteins of aerial tissues in *A. thaliana* seedlings. TAMRA-alkyne was conjugated to newly synthesized Aha-tagged proteins to render them fluorescent. The gel was visualized with an excitation laser at 532 nm and an emission band pass filter at 580 nm. **C**. Loading control. After measuring fluorescence, the gel was stained with

colloidal blue to confirm equal loading. Abbreviations: TAMRA – tetramethylrhodamine; DBCO – dibenzoazacyclooctyne.

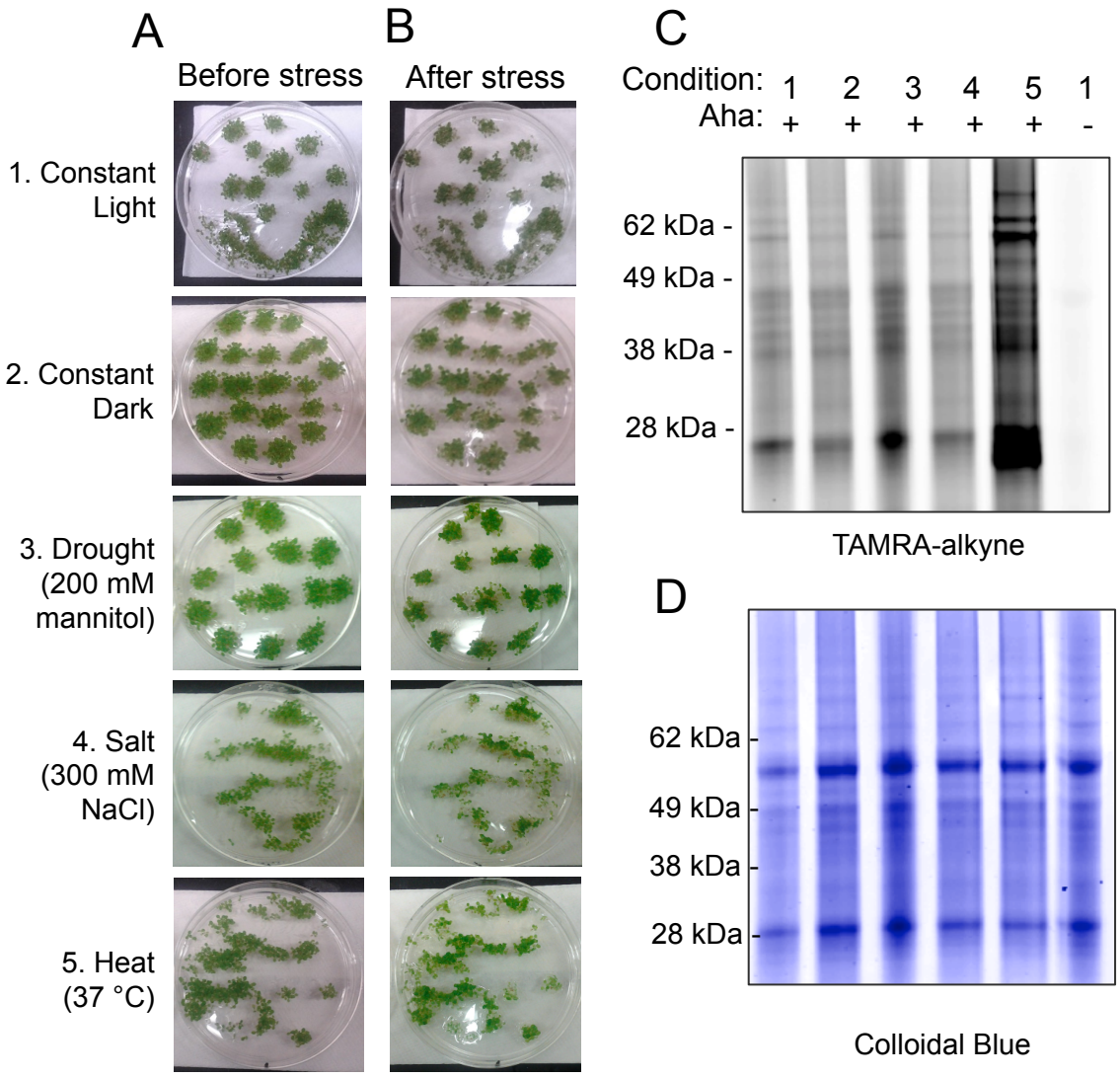


Figure 3: Labeling five separate stress conditions in *A. thaliana.* **A**. *A. thaliana* seedlings prior to Aha pulse and stress exposure. **B.** Seedlings 3 h after Aha pulse (1 mM) and constant exposure to stress conditions 1-5 (defined at left of figure). **C.** In-gel fluorescence assay to demonstrate labeling under stress conditions. Newly synthesized proteins incorporate Aha; TAMRA-alkyne is conjugated to proteins containing Aha. Labeling is observed under each condition. Little background signal is observed in a negative control, where plants were not exposed to Aha. **D.** Colloidal Blue loading control to demonstrate equal loading across lanes. Gels are representative examples of at least 3 biological replicates.

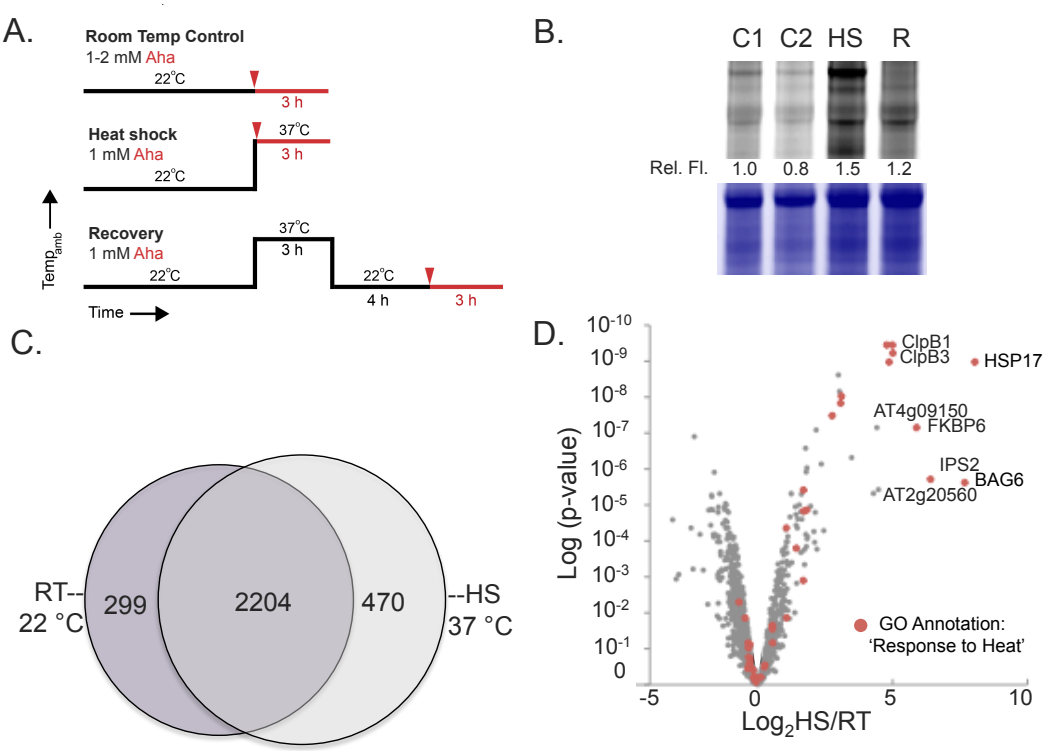


Figure 4: Enrichment of newly synthesized proteins for proteomics. **A.** Treatments used for this study. Arrowheads indicate time of introduction of Aha. **B.** Labeling under various conditions. **C.** Venn diagram of proteins identified in control conditions versus heat shock conditions. **D.** Volcano plot of ratios of expression levels of proteins shared between heat shock and control conditions. Proteins with higher average expression in RT samples fall on the left side of the plot, whereas proteins with higher average expression in HS samples are displayed on the right. To construct the plot, LFQ values were averaged for each condition. Then, the HS average was divided by the RT average and the Log_2 value was taken. Each point represents a protein. Proteins shown in red have the GO annotation "Response to Heat". Abbreviations: HS – heat shock, RT – room temperature, LFQ – label free quantitation, avg – average.

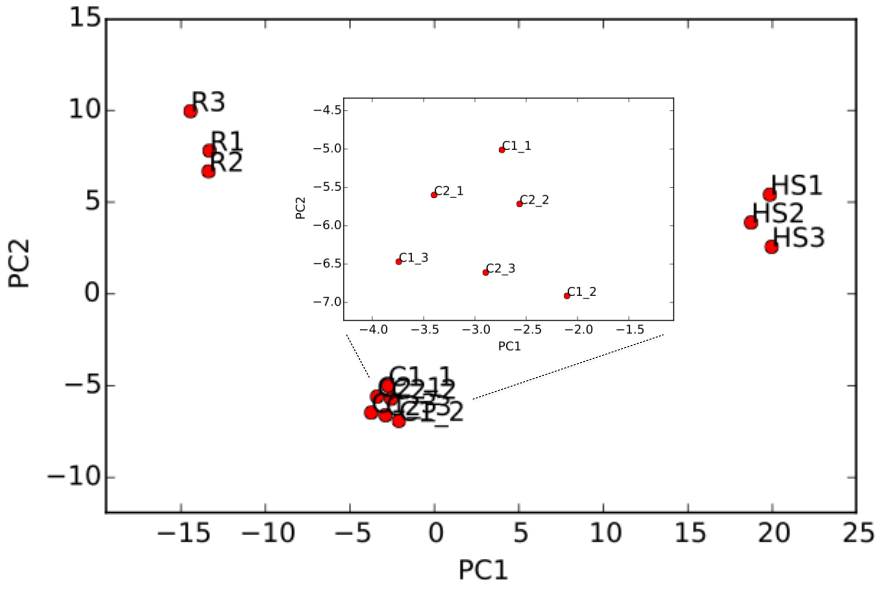


Figure 5: Principal component analysis of mass spectrometry results based on LFQ values. This analysis shows clear separation of control samples, heat shock samples and recovery samples. Inset shows zoom-in of controls cluster.

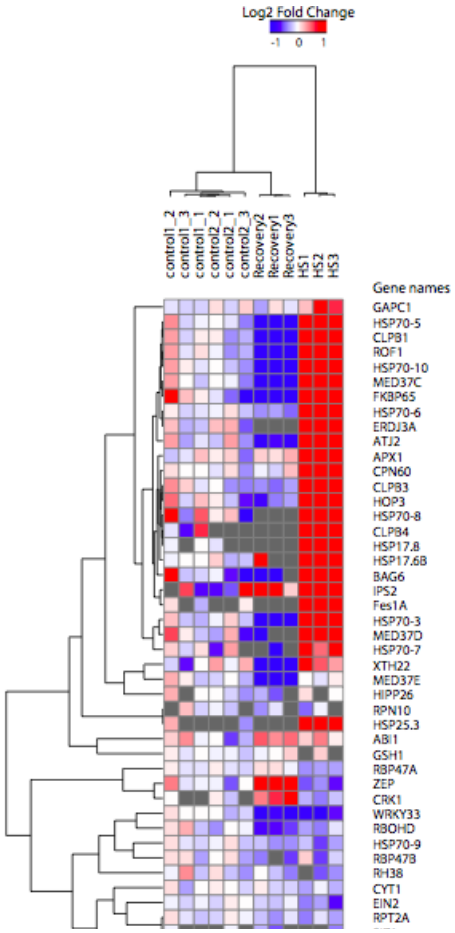


Figure 6: Partial heat map of proteins with GO annotation 'response to heat' found in this study. Significance of each fold change was calculated using the R package limma. Heat maps were created using GENE-E where the sample clustering was performed using the average linkage and Euclidean distance and the gene clustering was performed using the average linkage and 1-Pearson correlation coefficient. For heat map visualization, proteins had to be quantified in at least two control samples and two "treated" samples (either heat shock or recovery). Relative protein expression was normalized individually for each protein so that the average control expression was zero.

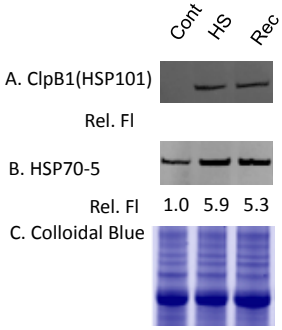


Figure 7: Immunoblotting analysis of select proteins shown in BONCAT screen to be up-regulated in response to heat stress. **A.** ClpB1(HSP101) and **B**. Heat Shock Protein 70-5 (HSP70-5) were found to be highly up-regulated in response to heat stress. These proteins are not synthesized at high levels during the recovery period. Neither are they rapidly degraded during the recovery period. Steady state ClpB1 levels during recovery are 0.95 ± 0.08 when the fluorescent signal of heat shock samples is normalized to 1.00. Relative fluorescence values are provided for the control (room temperature), heat shock and recovery for HSP70-5. **C.** Loading control. All fluorescence signals were normalized to Colloidal Blue staining. Abbreviations: Rel. Fl – relative fluorescence, Cont – control (room temperature), HS – Heat Shock, Rec – Recovery.