Appendix A

SUPPLEMENTARY INFORMATION FOR CHAPTER 2

A.1 Supplementary Experimental Procedures

Strain Construction.

See Table A.1 for a full list of strains. An unmarked deletion of *sutA* (DKN1625) was generated by first cloning 1 kb of sequence upstream and downstream of this gene into the pMQ30 suicide vector [1]. This vector carries the URA3 gene from *Saccharomyces cerevisiae*, which facilitated the use of homologous recombination in yeast to stitch together the three DNA pieces. The upstream and downstream 1 kb regions were amplified from *P. aeruginosa* gDNA and cleaned up using the PCR purification kit (Qiagen). Linearized pMQ30 plasmid was transformed along with the 1 kb flanking regions into *S. cerevisiae* using standard methods and successful transformants were selected on media lacking uracil. The pMQ30 plasmid carrying the upstream and downstream sequences for *sutA* was recovered from the yeast colonies by extraction with phenol:chloroform:isoamyl alcohol and transformed into *E. coli* DH5α cells. The construct was verified by sequencing and introduced into *P. aeruginosa* UCBPP-PA14 by triparental conjugation. Successful exoconjugants were selected on VBMM medium containing 100 $\mu g/mL$ gentamicin as described by Choi and Schweizer [2], and were then subjected to counterselection on LB plates lacking NaCl and containing 10% sucrose. Colonies resulting from homologous recombination to remove the wild-type copy of *sutA* and retain the clean deletion were identified by PCR.

The strain overexpressing SutA (DKN1626) was constructed by first cloning the SutA coding sequence into the multiple cloning site of the expression vector pMQ72, placing it under control of the arabinose-inducible P_{ara} promoter, using yeast homologous recombination as described above. The P_{ara} promoter:*sutA* coding sequence cassette was then cloned into the pUC18T-miniTn7T-GmR vector in order to direct its insertion into the attTn7 site of *P. aeruginosa* [2], using the Gibson reaction [3]. This vector was introduced into *P. aeruginosa* UCBPP-PA14 by tetraparental conjugation, and verified by PCR.

To construct the plasmid for overexpression of hemagglutinin-tagged SutA, the *sutA* gene, along with 1 kb upstream and downstream, was cloned from *P. aeruginosa*

gDNA with a 5' overhang encoding the HA epitope (MYPYDVPDYA) and inserted into pMQ30 using the Gibson reaction. The HA-*sutA* gene was then amplified and cloned into the multiple cloning site of pMQ72 between the SacI and KpnI restriction sites (DKN1640). This vector was transformed into *P. aeruginosa* by electroporation.

The GFP- and CFP-marked wild-type and Δ*sutA* strains (DKN1632-1635) carry their respective fluorescent proteins under the control of the strong PA1/04/03 promoter, integrated into the *attTn7* site and marked by a gentamicin resistance cassette. The fluorescent markers were introduced into *P. aeruginosa* by tetraparental conjugation with *E. coli* strains carrying the respective fluorescent protein-encoding plasmids, which were gifts from the laboratory of Gary Schoolnik [4].

The super-folder GFP reporter strains (DKN1627-1628) were generated by first amplifying 1 kb of sequence upstream and the intergenic sequence downstream of the *sutA* and *rpsG* genes from *P. aeruginosa* gDNA. These fragments were cloned upstream and downstream of the sfGFP coding sequence [5] in the pUC18T-miniTn7T-GmR vector using the Gibson reaction, and the resulting construct was introduced into the *attTn7* site in *P. aeruginosa* by tetraparental conjugation.

Media and Growth Conditions.

All cultures were grown at 37 °C with shaking unless otherwise noted. Liquid media were LB (5 g yeast extract, 10 g tryptone, 10 g NaCl per liter), 2xYT (10 g yeast extract, 16 g tryptone, and 5 g NaCl per liter), or phosphate buffered minimal medium (35.9 mM K_2HPO_4, 14.2 mM KH_2PO_4, 9.3 mM NH_4Cl, 42.8 mM NaCl, 1.0 mM $MgSO_4$, 7.5 µM $FeCl_2 \cdot 4\,H_2O$, 0.8 µM $CoCl_2 \cdot 6\,H_2O_4$, 0.5 µM $MnCl_2 \cdot 4\,H_2O$, 0.5 µM $ZnCl_2$, 0.2 µM $Na_2MoO_4 \cdot 2\,H_2O$, 0.1 µM $NiCl_2 \cdot 6\,H_2O$, 0.1 µM H_3BO_3, 0.01 µM $CuCl_2 \cdot 2\,H_2O$) with carbon sources added as noted. All anaerobic cultures were incubated in butyl rubber-stoppered Balch tubes in a Coy anaerobic chamber supplied with an atmosphere of 5% H_2, 15% CO_2 and 80% N_2, with trace amounts of oxygen removed by palladium-catalyzed reaction with the hydrogen gas. Anaerobic cultures were incubated without shaking.

BONCAT Labeling and Enrichment.

L-azidohomoalanine (Aha) [6] and the dialkoxydiphenylsilane (DADPS) biotin-alkyne probe [7] were synthesized as previously described. *P. aeruginosa* PA14 was grown overnight in LB and diluted to OD_{500} 0.02 into minimal medium containing 40 mM arginine, pH 7.2. The culture was grown to OD_{500} 0.4, and split into

aerobic and anaerobic samples. To label aerobic cultures, Aha was added to a final concentration of 1 mM. After 15 min of incorporation, cells were washed once with PBS and cell pellets were frozen at -80 °C. Anaerobic samples were moved to an anaerobic chamber, washed with PBS, resuspended in minimal medium with 40 mM arginine, and sealed in Balch tubes. Anaerobic cultures were allowed to consume residual oxygen and adapt to anoxia for 24 h Aha was then added to a final concentration of 1 mM. After 16 h of incorporation, cells were pelleted, washed with PBS, and lysed immediately. For anaerobic samples, all steps up to and including lysis were performed using degassed solutions in the anaerobic chamber.

All samples were lysed by resuspension in lysis buffer (100 mM Tris-HCl, pH 8, 1% SDS). Lysates were heated to 65 °C for 5 min and clarified by addition of Benzonase Nuclease (Sigma Aldrich) for 1 h at 37 °C followed by centrifugation. For fluorescence detection of Aha-labeled proteins, lysates were reacted with 5 µM TAMRA-alkyne (Click Chemistry Tools), 100 µM chCuSO4, 500 µM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 5 mM aminoguanidine hydrochloride, and 5 mM sodium ascorbate [8] for 15 min at room temperature, precipitated with water, methanol, and chloroform, and washed twice with methanol. Reacted lysates were separated via SDS-PAGE and imaged on a Typhoon gel imager (GE Healthcare). Gels were stained with Colloidal Blue (Life Technologies) to verify equal protein loading.

For protein enrichment, 0.5 mg of each protein lysate was reacted with 100 µM DADPS biotin-alkyne probe as above for 3.5 h at room temperature. Proteins were precipitated with acetone at -20 °C and resuspended in PBS with 0.3% SDS. Streptavidin UltraLink Resin (Pierce Biotechnology) was washed twice with PBS, added to biotinylated lysates, and incubated overnight at 4 °C. Resin was transferred to microfuge spin columns (Pierce Biotechnology) and washed twice with 1% SDS in PBS and once with 0.1% SDS in PBS. Proteins were eluted by cleavage of the DADPS linker via incubation with 5% formic acid and 0.1% SDS in PBS for 2 h at room temperature. Resin was washed with 0.1% SDS in PBS to elute all proteins. Elution fractions were combined and concentrated by centrifugation through Amicon Ultra spin columns (EMD Millipore). The entirety of the concentrated eluents were separated via SDS-PAGE and stained with Colloidal Blue.

Protein Digestion, Mass Spectrometry, and Data Analysis.

For GeLCMS, gel pieces were destained by alternating washes with 50 mM ammonium bicarbonate (AB) and 1:1 50 mM AB:acetonitrile. Proteins were reduced by

incubation with 6.7 mM dithiothriitol (DTT) in 50 mM AB at 50 °C for 30 min and alkylated by incubation with 37 mM iodoacetamide in 50 mM AB at room temperature for 20 min. Gel pieces were washed with 100 mM AB and then with acetonitrile. Proteins were digested with 300 ng endoproteinase LysC in 100 mM Tris-HCl at 37 °C for 18 h. Peptides were extracted by sequential washing with: 1% formic acid/2% acetonitrile, 1:1 acetonitrile:water, and 1% formic acid in acetonitrile. Extracted peptides were dried and desalted using C18 StageTips as previously described [9].

For in-solution digestion, proteins were brought to a final concentration of 8 M urea, reduced by incubation with 3 mM tris(2-carboxyethyl) phosphine (TCEP) for 20 min at room temperature, and alkylated by incubation with 10 mM iodoacetamide for 15 min at room temperature in the dark. For immunoprecipitation, proteins were digested with 250 ng endoproteinase LysC for 18 h at room temperature. Samples were further digested by dilution with 100 mM Tris-HCl to a final urea concentration of 2 M and addition of 600 ng trypsin and 1 mM $CaCl_2$ at room temperature for 9 h. Digestion was quenched by addition of 5% formic acid. Digested peptides were desalted by HPLC using a Michrom Bioresources C18 macrotrap, (Buffer A: 0.2% formic acid in chH2O; Buffer B: 0.2% formic acid in acetonitrile) and concentrated *in vacuo*. Peptides were dimethyl labeled following established protocols [10] and mixed in a 1:1 mass ratio.

Liquid chromatography-mass spectrometry were essentially carried out as previously described [11]. Anaerobic vs. aerobic BONCAT and immunoprecipitation experiments were performed on a nanoflow LC system, EASY-nLC 1000 coupled to a hybrid linear ion trap Orbitrap Classic mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nanoelectrospray ion source (Thermo Fisher Scientific) with the following modifications: For the EASY-nLC II system, solvent A consisted of 97.8% chH2O, 2% ACN, and 0.2% formic acid and solvent B consisted of 19.8% chH2O, 80% ACN, and 0.2% formic acid. For the LC-MS/MS experiments, digested peptides were directly loaded at a flow rate of 500 nL/min onto a 16-cm analytical HPLC column (75 µm ID) packed in-house with ReproSil-Pur C_{18}AQ 3 µmresin (120 Å pore size, Dr. Maisch, Ammerbuch, Germany). The column was enclosed in a column heater operating at 30 °C. After 30 min of loading time, the peptides were separated with a 50 min gradient at a flow rate of 350 nL/min. The gradient was as follows: 0–30% B (50 min), and 100% B (10 min). The Orbitrap was operated in data-dependent acquisition mode to automatically alternate between a full scan (m/z=400–1600) in the Orbitrap and subsequent 10 CID MS/MS

scans in the linear ion trap. CID was performed with helium as collision gas at a normalized collision energy of 35% and 30 ms of activation time.

For the BONCAT experiment, raw files were searched using MaxQuant [12] against the *P. aeruginosa* PA14 UniProt entries (5,886 sequences) and a contaminant database (246 sequences). Trypsin was specified as the digestion enzyme with up to two missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification and protein N-terminal acetylation and methionine oxidation were variable modifications. We also included variable modifications of methionine corresponding to Aha, reduced Aha, Aha reacted to the DADPS linker, and Aha reacted to the cleaved DADPS linker. Protein ratios and their standard errors were calculated using bootstrap estimates and pooled variance estimates at the peptide level [13]. Briefly, peptide intensities were normalized to the total intensity for each run and a global estimate of measurement error was calculated using pooled variance from all peptide ratios between each sample. The protein ratio was calculated as the median of peptide ratios. The standard error of the protein ratio was calculated using a bootstrap procedure where resampling of peptide ratios is augmented by adding a random "noise" effect drawn from a normal distribution with mean zero and standard deviation equal to the previously calculated global estimate of measurement error. In total, 1000 bootstrap iterations were performed. The standard error of the protein ratio was then calculated as the standard deviation of the bootstrapped peptide ratios. Z-tests were then used to calculate p-values of overall protein ratios with respect to a 1-to-1 ratio. P-values were adjusted for false discovery by the Benjamini-Hochberg procedure.

For dimethyl labeling experiments, raw files were searched using MaxQuant as above. Trypsin was specified as the digestion enzyme with up to two missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification and protein N-terminal acetylation and methionine oxidation were variable modifications. Dimethyl mass modifications (light and medium) at lysine residues and peptide N-termini were specified for quantification.

Colony Morphology Assay.

Cultures were grown overnight in LB medium, diluted 1:1000, and spotted in a 10 µL volume on solid media (1% tryptone, 1% Bacto Agar, 20 µg/ml Coomassie blue and 40 µg/ml Congo red) [14]. Plates were incubated at room temperature for 6 days and then imaged using a Keyence VHX-1000 digital microscope.

Crystal Violet Assay.

The Crystal Violet assay was performed as previously described [15]. Cultures were grown overnight in LB and diluted 1:1000 into LB. 125 µl of each diluted culture was transferred to 96-well round bottom polystyrene plates coated for tissue culture (Corning). Plates were sealed with parafilm and incubated for 18 h at 37 °C without shaking. Wells were washed with 0.9% NaCl and treated with 150 µl of 0.1% Crystal Violet for 20 min at room temperature. Wells were washed three times with water and Crystal Violet was extracted from adherent cells by addition of ethanol. Ethanol containing Crystal Violet was transferred to a new well-plate and absorbance at 600 nm was measured. The average absorbance for wells containing only LB was subtracted from all measurements. Each strain was measured in two separate experiments, with four wells per experiment.

Phenazine Measurements.

Phenazine concentrations in culture supernatants were measured as described previously [16]. Briefly, culture supernatants were filtered using SpinX columns with a 0.2 µM pore size and were directly loaded onto a Beckman System Gold reverse-phase high-pressure liquid chromatography (HPLC) instrument with a UV-visible light (Vis) detector and a Waters Symmetry C_{18} analytical column (5 µm particle size; 4.6 by 250 mm). A gradient of water-0.1% trifluoroacetic acid (TFA; solvent A) to acetonitrile-0.1% TFA (solvent B) at a flow rate of 1 ml/min was used to elute phenazines, which can be detected based on their characteristic absorption wavelengths and retention times. Peak areas for samples were compared to peak areas from standards of purified phenazine-1-carboxylic acid and pyocyanin.

Competition Assay.

Individual overnight cultures of wild-type cells carrying a *gfp* or a *cfp* marker and Δ*sutA* cells carrying a *gfp* or a *cfp* marker were grown in 5 ml LB medium. Cultures were diluted 1:1000 in LB medium and mixed in equal proportions based on their ODs in the following combinations: A) wild type, *gfp* marked plus Δ*sutA*, *cfp* marked; B) wild type, *cfp* marked plus Δ*sutA*, *gfp* marked; and C) wild type, *gfp* marked plus wild type, *cfp* marked. The mixtures were allowed to grow to mid-exponential phase (OD approximately 0.4), with shaking at 37 °C. Small aliquots of the mixed cultures were taken for microscopy (time 0 sample) and the remainders were pelleted and transferred to an anaerobic chamber (Coy), where they were resuspended in anaerobic minimal medium with 40 mM arginine and placed in

sealed Balch tubes. Cultures were incubated anaerobically at 37 ºC for 19-20 h, then were removed from the anaerobic chamber, diluted 1:100 or 1:200 into LB medium, and allowed to grow aerobically with shaking at 37 ºC for 4-6 h, back to mid-exponential phase. No significant change in OD occurred during the anaerobic incubation. Once cells reached mid-exponential phase, a small aliquot of the culture was taken for microscopy (transfer 1 sample) and the remainders of the cultures were pelleted and resuspended again in the anaerobic arginine medium in the anaerobic chamber. This process was repeated for 4 transfers. At each transfer, epifluorescence microscopy using a Zeiss Axio Imager microscope was used to observe live cells placed on agarose pads. GFP was detected using the Zeiss 46HE filter cube and CFP was detected using the Zeiss 47HE filter cube. The percentage of cells carrying each marker in each mixed culture was counted. At least 500 cells were counted for each sample. A very small bias in favor of carrying CFP over GFP was detected in the wild type vs. wild type mixed culture (combination C), so at each time point, the proportion of each marker in this culture was taken to reflect the "no advantage" state, and the wild type vs. Δ*sutA* proportions were adjusted by the difference observed due to carrying GFP vs. CFP. The adjusted proportions in the two marker-flipped cultures (combinations A and B) were averaged together. The entire experiment was performed three times.

GFP Reporter Protein Measurement.

For growth, transcript and reporter protein measurements, starter cultures were grown to stationary phase in LB medium, diluted 1:1000 into either LB or pyruvate minimal medium and allowed to grow into early exponential phase (approx. 4 h for LB or 18 h for pyruvate) at which point the "Time 0" measurements were made. Live cells in liquid culture were diluted between 1:250 and 1:1000 into TBS (50 mM Tris, pH 8.0, 150 mM NaCl) containing the SYTO 62 red-fluorescent, cell permeant nucleic acid stain at a concentration of 500 nM. Cells were incubated at room temperature in the dark for 15-20 min to allow for DNA staining. The BD Accuri c6 flow cytometer was used to measure both red fluorescence from the SYTO 62 dye (excitation laser: 640 nM, emission filter: 675/25 nM) and green fluorescence from GFP (excitation laser 488 nM, emission filter: 533/30 nM). Particles were gated on forward scatter vs. red fluorescence, and the mean green fluorescence for particles with red fluorescence and forward scatter values consistent with cells was measured. At each time point, mean green fluorescence in a strain lacking GFP was also measured to determine background autofluorescence and this value was

subtracted from the GFP values for that time point. GFP was measured for biological triplicates for each genotype and condition.

RNA Extraction.

Total RNA was extracted from cells using the RNeasy Mini Kit (Qiagen). Briefly, approximately 10×10^9 cells were pelleted rapidly by centrifugation at top speed in a microfuge, the supernatant was removed, and the pellet was immediately frozen in liquid nitrogen. After all samples were collected, pellets were resuspended in TE buffer (10 mM Tris pH 7.5, 1 mM ethylenediaminetetraacetic acid (EDTA)) plus 15 mg/ml lysozyme (Sigma) and 15 U per sample proteinase K (Qiagen), and incubated for 10 min at 37 ºC to digest the cell wall. Samples were then processed according to the manufacturer's instructions, including on-column DNAse treatment. Purified RNA was quantified by absorbance at 260 nm, and 10 µg per sample was treated with Turbo DNAse Free (Ambion) according to the manufacturer's instructions. Samples were verified to be free of genomic DNA by quantitative PCR.

qRT-PCR.

1 µgof DNAse-treated total RNA was converted to cDNA using the iScript cDNA synthesis kit (BioRad). 1/100th of this reaction mixture (representing 10 ng total RNA) was used per qRT PCR reaction, along with 500 nM each of forward and reverse primers and the iTaq SYBR Green reaction mix (BioRad). Samples were run on the ABI platform (ABI) for 40 cycles with an annealing temperature of 60 ºC. Standard curves for each primer pair were generated using serial dilutions of genomic DNA. The *oprI* gene was used to normalize against potential loading differences. See Table A.2 for primer sequences. Measurements were made on biological triplicates.

Co-Immunoprecipitation

Cultures of Δ*sutA* carrying pMQ72 or pMQ72-HAsutA were grown overnight in minimal medium containing 40 mM sodium pyruvate, 20 mM arabinose, and 50 µg/ml gentamicin to an OD_{500} of approximately 1. Cells were washed once in PBS and frozen at -80 ºC. Cell pellets were resuspended in IP lysis buffer (50 mM HEPES, 70 mM potassium acetate, 5 mM magnesium acetate, 0.2% n-dodecyl-β-D-maltoside, and cOmplete mini protease inhibitor, EDTA free (Roche)). Cells were gently lysed by passage through a 22G needle ten times. Lysates were clarified by incubation with Benzonase Nuclease for 1 h at 37 ºC followed by centrifugation.

For IP of HA-SutA, 50 µl agarose beads conjugated to an anti-HA antibody (Sigma-Aldrich) were washed three times in IP lysis buffer, combined with 1 ml lysate, and incubated with rotation overnight at 4 ºC. For IP of RpoA, 1 ml lysate was incubated with an anti-RpoA antibody (gift of Olaf Schneewind) for 1 h at 4 ºC with rotation. 50 µl Protein A/G PLUS-agarose beads (Santa Cruz Biotechnology) were washed three times with IP lysis buffer, combined with the antibody-lysate mixture, and incubated with rotation overnight at 4 ºC. For both IPs, beads were washed twice with 0.5 ml IP lysis buffer and twice with 0.5 ml 100 mM Tris-HCl, pH 8. Proteins were eluted by incubation with 64 µl 10 M urea in 100 mM Tris-HCl. IP eluents were digested in-solution, reacted with dimethyl labels, and analyzed by LC-MS/MS, as described above.

For Western blotting, 10 µL of each IP fraction (lysate, flow through, four washes, and elution) were separated by SDS-PAGE and transferred to a Hybond ECL membrane (GE Healthcare). Membranes were blocked with 5% milk in TBST (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Tween 20). HA-SutA was detected by anti-HA antibody-Alexa Fluor 594 conjugate (Life Technologies). RpoA was detected by incubation with the primary anti-RpoA antibody described above, followed by incubation with a goat anti-mouse antibody-Alexa-Fluor 633 conjugate (Life Technologies). On a separate gel, the same samples were stained with Coomassie.

RNA Seq Library Preparation.

For RNA-Seq experiments, starter cultures were grown to stationary phase in LB, diluted 1:1000 in pyruvate minimal medium containing 25 mM arabinose, and then allowed to grow 21 h until they reached stationary phase again (OD_{500} of approximately 1), at which point cells were collected for RNA extraction (described above). 3.8 µg of DNAse-treated total RNA was subjected to ribosomal RNA depletion using the Gram Negative Magnetic Ribo-Zero kit (Epicentre), according to the manufacturer's instructions. Following rRNA depletion, samples were cleaned up using the RNeasy MinElute kit (Qiagen) and libraries were generated for sequencing using the NEBNext mRNA Library Prep Kit for Illumina (NEB). Briefly, mRNAs were fragmented by treatment with MgCl-containing fragmentation buffer for 1 min at 94 ºC and cleaned up using the RNeasy MinElute columns. Fragmentation to an average size of approximately 200 bp was verified by running the samples on a Bioanalyzer RNA Pico chip (Agilent). The fragmented RNA was reverse-transcribed to cDNA, which was then end-repaired, dA-tailed, and ligated to adaptors. Each sample was PCR-amplified with a universal primer and a unique bar-coded primer, using 12

amplification cycles. Final libraries were verified using the High-Sensitivity DNA chip on the Bioanalyzer and quantified using the Qubit fluorimeter and dsDNA dye (Invitrogen). Sequencing was performed on biological triplicates for each genotype.

Chromatin Immunoprecipitation.

Growth conditions were the same as for the RNA-seq experiments except 20 mM arabinose was used and 50 µg/ml gentamicin was added for plasmid maintenance. Stationary phase cultures of the Δ*sutA* strain (DKN1625) carrying either pMQ72 or pMQ72-HA-sutA in pyruvate minimal medium were cross-linked by incubation with 1% formaldehyde at room temperature for 15 min and then crosslinking was quenched by incubation with 125 mM glycine for 10 min. Cells were pelleted and washed twice with TBS (50 mM Tris pH 7.5, 150 mMchNaCl), and then pellets were frozen at -80 ºC. Frozen pellets were resuspended in 1.5 mL IP buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% deoxycholic acid, 1 mg/ml lysozyme) and incubated at 37 ºC for 15 min. Samples were then chilled on ice and sonicated using a microtip sonicator for 4 min at the 4.0 setting, using a cycle of 30 s on, 30 s off. Samples were split in half; one half was subjected to immunoprecipitation by an antibody against RpoA, and the other half was subjected to immunoprecipitation by an antibody against the HA epitope, as described for protein IP above. For the RpoA immunoprecipitation, samples were pre-cleared by incubation with 1/10 volume Protein A/G PLUS-agarose beads (Santa Cruz Biotechnology) for 1 h at 4 ºC, and then were incubated overnight with rotation at 4 ºC with the anti-RpoA antibody. Next 50 µl of the protein A/G agarose beads were added and the mixture was incubated for an additional 1 h at 4 ºC. For the HA-SutA immunoprecipitation, samples were incubated with 50 µl pre-conjugated HA bead slurry overnight with rotation at 4 ºC. The beads from both immunoprecipitations were then washed 5 times for 10 min per wash. Washes 1 and 2 were with IP buffer, wash 3 was with IP buffer with 500 mM NaCl, wash 4 was with stringent buffer (10 mM Tris pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% Nonidet P-40), and wash 5 was with TBS. DNA/protein complexes were eluted from the beads in 100 µL elution buffer (50 mM Tris pH 7.5, 10 mM EDTA, 1%SDS) by incubation for 15 min at 65 ºC. The elution was repeated once and both eluates were combined, then were incubated at 65 ºC overnight to reverse crosslinks. 200 µl TE buffer (10 mM Tris pH 8.0, 1 mM EDTA), 100 µg proteinase K, and 20 µg glycogen were added to each sample and they were incubated for 2 h at 37 ºC to digest proteins. DNA was extracted using 25:24:1 phenol:chloroform:isoamyl alcohol and precipitated with

ethanol. The precipitated DNA was resuspended in 30 µl TE buffer containing 10 µg RNAse A and incubated at 37 ºC for 2 hr to remove RNA contamination, and then was cleaned up using a QIAquick column (Qiagen) with an elution volume of 50 µl [17].

ChIP Seq library preparation.

2-10 ng purified genomic DNA isolated by immunoprecipitation was subjected to further fragmentation by treatment with the NEB ds Fragmentase enzyme cocktail for 10 min at 37 ºC. This reduced the average fragment size from approximately 500-1000 bp to approximately 200-500 bp for optimal high throughput sequencing efficiency. Fragmented DNA was cleaned up using Agencourt AMPure XP magnetic beads (Beckman Coulter). Libraries were prepared from the fragmented gDNA using the NEBNext ChIP Seq Library Prep Reagent Set for Illumina (NEB). DNA fragments were end-repaired, dA-tailed, ligated to adaptors, and PCR amplified with one universal and one bar-coded primer, using 15 amplification cycles. Final libraries sizes were verified using the Bioanalyzer, and library amounts were quantified using the Qubit fluorimeter. All ChIP-Seq was performed on biological triplicates.

Sequencing and Data Analysis.

All sequencing was performed by the Millard and Muriel Jacobs Genetics and Genomics Laboratory at the California Institute of Technology using the Illumina HiSeq 2500 platform. 10-15 million reads of 50 or 75 bp each were collected for each sample. Base-calling and de-multiplexing were performed by the Illumina HiSeq Control Software (HCS, version 2.0). The resulting FASTQ files were concatenated into one file per sample and filtered and trimmed by quality score per base using the Trimmomatic software package with the following parameters: LEADING:27 TRAILING:27 SLIDINGWINDOW:4:20 MINLEN:35 [18]. Surviving reads were mapped to the *P. aeruginosa* UCBPP-PA14 genome sequence (gi|116048575|ref|NC_008463.1) using the Bowtie package with the -n 2 and –best arguments [19]. Specifically for assessing ChIP signal at tRNA genes, Bowtie was run with the -n 2 and -m 1 arguments to require reads to be uniquely mapped in order to be reported. Mapped reads were sorted, indexed, and converted to binary format using the SAMtools package [20]. Reads per 100 bp, gene, or transcriptional unit (TU) were calculated using the easyRNASeq package from the Bioconductor project in R [21]. The .gff file describing the location of genes was generated using the bp_genbank2gff3.pl script from the Bioperl project and the Genbank file for

the RefSeq accession NC_008463.1. The .gff file was modified to additionally include small non-coding RNAs and novel ORFs detected by deep sequencing of the UCBPP-PA14 strain of *P. aeruginosa* [22], and to consistently name genes by their locus tags rather than a mixture of locus tags and gene names. The .gff file describing the locations of transcriptional units was derived from the table of transcriptional units published defined by Wurtzel and colleagues [22], and uses the start of the first coding sequence and the end of the last coding sequence in each operon as the operon boundaries. Average ratios and significance of differential expression or ChIP association between different genotypes or pulldowns were calculated using the Degust web server hosted by the Victorian Bioinformatics Consortium. The Degust project uses the voom and limma packages in R to perform calculations [23].

For viewing ChIP data across genomic loci, the counts per 100 bp for each sample were normalized to the size of the library by converting counts to RPKM, and then further scaled based on the values observed low- and high-signal regions. This method was adapted from one described by Mooney et al. [24]. The baseline value for each sample was defined as the average RPKM value for the 25 transcriptional units at least 1 kb in length that had the lowest signal in the RpoA pulldown from the HA-SutA strain. These transcriptional units were verified to have among the lowest RPKM values from the RNA-Seq data as well, and were assumed to be essentially not transcribed under the conditions of the experiment. The maximum value for each sample was defined as the average RPKM value for the top ten peaks associated with protein-coding genes for that type of pulldown. A peak was defined as two consecutive 100bp regions that fell among the top 100 100bp regions. While some peak regions were the same for both the HA-SutA and the RpoA pulldown, some were distinct. See Dataset A.4 for regions and values used. To scale the RPKM data, the baseline value was subtracted from each 100bp RPKM value and the result was divided by the maximum value, such that nearly all scaled values fall between 0 and 1. The biological triplicates for each pulldown were averaged. The MochiView software package [25] was used to smooth the scaled 100 bp values over a 300 bp rolling window, and then the coordinates of regions with scaled values above 0.20 for the HA-SutA pulldown and scaled values above 0.25 for either RpoA pulldown were extracted. Regions less than 100 bp apart were merged. This set of "high chip signal" regions was then filtered to include only 100 bp regions that also showed a statistically significant enrichment in the HA-SutA pulldown compared to the mock pulldown, which left a total of 2,015 100bp regions that were considered "high

ChIP." 230 transcriptional units starting within a high chip region were identified. There were 405 genes that were contained within these transcriptional units and were considered the list of "high chip" genes that was compared with the list of up- and down-regulated genes. For the aggregate ChIP plot, transcriptional units containing ribosomal protein genes were excluded, since these had already been separately considered, and of the remaining, only transcriptional units with start sites defined by Wurtzel et al. [22] were included. See Dataset A.6 for the transcriptional unit data that was used.

Functional analysis of genes transcriptionally affected more than 2-fold was carried out using the COG category designations recorded in the Pseudomonas Genome Database (`http://www.pseudomonas.com`) [26]. For simplicity, several COG categories were grouped together for each bar in the bar plot. The category designated "unknown" contains COG categories R and S ("General functional prediction only" and "No functional prediction") in addition to genes that did not have an associated COG. The category designated "maintenance and secondary metabolism" contains COG categories C, I, P, O, and Q ("Energy production and conversion", "Lipid transport and metabolism", "Inorganic ion transport and metabolism", "Post-translational modification, protein turnover, and chaperones", and "Secondary metabolites biosynthesis, transport, and catabolism"). The category designated "growth and primary metabolism" contains COG categories D, E, F, G, H, J, L, and M ("Cell cycle control, cell division, chromosome partitioning", "Amino acid transport and metabolism", "Nucleotide transport and metabolism", "Carbohydrate transport and metabolism", "Coenzyme transport and metabolism", "Translation, ribosomal structure and biogenesis", "Replication, recombination and repair", and "Cell wall/membrane/envelope biogenesis"). The category designated "motility, defense, and signaling" contains COG categories N, T, U, and V ("Cell motility", "Signal transduction mechanisms", "Intracellular trafficking, secretion, and vesicular transport", and "Defense mechanisms"). The category designated "transcription and nucleic acid processing" contains COG categories A, B, and K ("RNA processing and modification", "Chromatin structure and dynamics", and "Transcription") (`ftp://ftp.ncbi.nih.gov/pub/COG/COG2014/static/lists/homeCOGs.html`) [27].

Software analysis and data presentation.

This section describes software packages that were not mentioned above. Data processing and statistical analysis were performed with Python version 2.7.9 with

NumPy version 1.9.2, SciPy version 0.15.1, and Pandas version 0.16.1. Data were plotted with Matplotlib version 1.4.3 [28] and Seaborn version 0.5.1. Gel images were analyzed with ImageJ 64-bit version 1.45 [29]. Figures were assembled in Adobe Illustrator CS5.

A.2 Supplementary Tables

Table A.1: **Chapter 2: Strains and plasmids.** Strains and plasmids used in this study. Plasmids are stored as *E. coli* strains carrying the plasmid, and requests should be for the *E. coli* strain.

***Pseudomonas aeruginosa* Strains**		
Name	Genotype	Source
DKN263	*P. aeruginosa* UCBPP-PA14	
DKN1625	UCBPP-PA14 Δ*sutA*	This Study
DKN1626	UCBPP-PA14 attTn7::mini-Tn7T-Gm^R P_{ara}*:sutA*	This Study
DKN1627	UCBPP-PA14 attTn7::mini-Tn7T-Gm^R P_{sutA}*:gfp*	This Study
DKN1628	UCBPP-PA14 attTn7::mini-Tn7T-Gm^R P_{rpsG}*:gfp*	This Study
DKN1632	UCBPP-PA14 attTn7::mini-Tn7T-Gm^R P_{A10403}*:gfp*	This Study
DKN1633	UCBPP-PA14 attTn7::mini-Tn7T-Gm^R P_{A10403}*:cfp*	This Study
DKN1634	UCBPP-PA14 Δ*sutA* attTn7::mini-Tn7T-Gm^R P_{A10403}*:gfp*	This Study
DKN1635	UCBPP-PA14 Δ*sutA* attTn7::mini-Tn7T-Gm^R P_{A10403}*:cfp*	This Study
Transposon insertion mutants	UCBPP-PA14 Gene::MAR2xT7	[30]
***Escherichia coli* Strains**		
Name	Genotype	Source
DKN1299	SM10, pTNS1	[2]
DKN1299	HB101 pRK2013	[2]
DKN1323	S17-1 λpir pMCM11(containing attTn7::mini-Tn7T-Gm^R P_{A10403}*:gfp*)	Gary Schoolnik
DKN1325	S17-1 λpir pMCM11 derivative (containing attTn7::mini-Tn7T-Gm^R P_{A10403}*:cfp*)	Gary Schoolnik
DKN1637	DH5α pMQ30_sutA	This Study
DKN1639	Mach1 pUC18T-mini-Tn7T-Gm^R P_{ara}*:sutA*	This Study
DKN1640	Mach1 pMQ72_HA-sutA	This Study
DKN548	DH5α pMQ72	George O'Toole
DKN1641	DH10β pUC18T-mini-Tn7T-Gm^R P_{sutA}*:gfp*	This Study
DKN1642	DH5α pUC18T-mini-Tn7T-Gm^R P_{rpsG}*:gfp*	This Study
***Saccharomyces cerevisiae* Strains**		
Name	Genotype	Source
DKN569	InvSc1:MATα/MATα his3D1/his3D1 leu2/leu2 trp1-289/trp1-289 ura3-52/ura3-52	Invitrogen

Table A.2: **Chapter 2: Primers.** Primers used in strain construction and qRT-PCR experiments.

Name	Purpose	Sequence
6977del1	Generating SutA deletion construct	tgggtaacgccagggttttcccagtcacgacgttgtaaaaCTGCTCACCGGGATCTTCGC
6977del2	Generating SutA deletion construct	TGGCGGGCCTTGGGATGACGCGAAAGGTCAACCTCTCGGTGCTGCAAAAG
6977del3	Generating SutA deletion construct	CTTTTGCAGCACCGAGAGGTTGACCTTTCGCGTCATCCCAAGGCCCGCCA
6977del4	Generating SutA deletion construct	tgtgagcggataacaatttcacacaggaaacagctatgacGTTCAGCCGGGCGGCAGCGA
Para:sutA1	Cloning SutA into pMQ72	ccatacccgttttttgggctagcgaattcgagctcAGGAGGGGTTGACCATGAGCGAAG
Para:sutA2	Cloning SutA into pMQ72	gcaaattctgttttatcagaccgcttctgcgttctgatttaaAAATCAGATGGGGCGGCT
sutA_gfp1	Generating SutA:gfp reporter construct	agtataggaacttcagagcgcttttgaagctaattcgatcCTGCTCACCGGGATCTTCGC
sutA_gfp2	Generating SutA:gfp reporter construct	TGAACAGCTCTTCGCCTTTACGCATGGTCAACCTCTCGGTGCTGCAAAAGC
sutA_gfp3	Generating SutA:gfp reporter construct	GCTTTTGCAGCACCGAGAGGTTGACCATGCGTAAAGGCGAAGAGCTGTTCA

Continued from previous page.

Name	Purpose	Sequence
sutA_gfp4	Generating SutA:gfp reporter construct	TGGCGGGCCTTGGGATGACGCGAAATCATCATTTGTACAGTTCATCCATA
sutA_gfp5	Generating SutA:gfp reporter construct	TATGGATGAACTGTACAAATGATGATTTCGCGTCATCCCAAGGCCCGCCA
sutA_gfp6	Generating SutA:gfp reporter construct	atagtttggaactagatttcacttatctggttggcctgcaGGGATGACAACCGATGTGTC
rpsG_gfp1	Generating RpsG:gfp reporter construct	agtataggaacttcagagcgcttttgaagctaattcgatcATCAAAGGCGACCAGGTGGA
rpsG_gfp2	Generating RpsG:gfp reporter construct	TGAACAGCTCTTCGCCTTTACGCATTGATAAGCCCTCAAACGGTCTTCAG
rpsG_gfp3	Generating RpsG:gfp reporter construct	CTGAAGACCGTTTGAGGGCTTATCAATGCGTAAAGGCGAAGAGCTGTTCA
rpsG_gfp4	Generating RpsG:gfp reporter construct	CCTTTTCTGATGGCAGGATCAGCGATCATCATTTGTACAGTTCATCCATA
rpsG_gfp5	Generating RpsG:gfp reporter construct	TATGGATGAACTGTACAAATGATGATCGCTGATCCTGCCATCAGAAAAGG
rpsG_gfp6	Generating RpsG:gfp reporter construct	atagtttggaactagatttcacttatctggttggcctgcaGACCTCAGACTCCAATTTAC
HAsutA1	Generating HA-SutA	GACCGCATGTACGCCGAAGcggggatcctctagagtcgacctgcaggca
HAsutA2	Generating HA-SutA	cagctatgaccatgattacgaattc

Continued from previous page.

Name	Purpose	Sequence
HAsutA3	Generating HA-SutA	tgcctgcaggtcgactctagaggatccccgCTTCGGCGTACATGCGGTC
HAsutA4	Generating HA-SutA	cagcaccgagaggttgaccATGTACCCATACGATGTTCCAGATTACGCT
HAsutA5	Generating HA-SutA	ATGTACCCATACGATGTTCCAGATTACGCTatgagcgaagaagaactggaac
HAsutA6	Generating HA-SutA	cagctatgaccatgattacgaattcACGAGATTGAACGGGGTAAC
HAsutA7	Moving HA-SutA to pMQ72	atatggtaccCTTCGGCGTACATGCGGTC
HAsutA8	Moving HA-SutA to pMQ72	atatgagctcACGAGATTGAACGGGGTAAC
Sfgfp_f	QPCR	TGGTGTTCAGTGCTTTGCTC
Sfgfp_r	QPCR	TGTACGTGCCGTCATCCTTA
oprI_f	QPCR	AGCAGCCACTCCAAAGAAAC
oprI_r	QPCR	CAGAGCTTCGTCAGCCTTG
Intergenic_f	QPCR	GGGGTGGGGGTAGTTAAAGA
Intergenic_r	QPCR	GCAAAACAAGCCCCTACAAA
16Sleader_f	QPCR	ACGAAAGCCTTGACCAACTG
16Sleader_r	QPCR	TTGCGCTGCTGATAATCTTG

A.3 Supplementary Datasets

Dataset A.1: Proteins more abundant or uniquely identified in the anaerobic sample. All proteins identified by LC-MS/MS from the BONCAT-enriched samples are listed. Columns 1-3 give the locus ID in both the PA14 and PAO1 strains as well as the gene name if available. Columns 4-8 give LC-MS/MS measurements for each protein or protein group: the number of unique peptides identified, number of evidences in the anaerobic and aerobic samples respectively, and the total peak intensities in the anaerobic and aerobic samples respectively. Columns 9 and 10 give the log_2-transformed median of all intensity ratios for peptides shared between the two samples, and the probability that the ratio is not different from zero, with an adjustment for the pooled variance of the experiment. Sheet1 lists proteins identified in both samples, Sheet2 lists proteins identified only in the anaerobic sample, and Sheet3 lists proteins only identified in the aerobic sample. NA: not available. NQ indicates that there was insufficient information for that protein to quantify a ratio between the anaerobic and aerobic samples.

Dataset A.2: Proteomic results from co-immunoprecipitations. All co-precipitated proteins identified by LC-MS/MS following pull-down of either SutA (Sheet1) or RpoA (Sheet 2) are listed. For the SutA immunoprecipitation, two independent experiments were performed, and the results are listed separately. Columns 1-3 list the locus IDs for both the PA14 and PAO1 strains, and the gene name if available. Column 4 lists the log2-transformed ratios between protein abundance in the HA-SutA sample and the untagged control sample, as quantified by dimethyl labeling, and columns 5-7 give the number of evidences and total peak intensities for the differentially labeled peaks. Columns 8-11 give this information for the second IP experiment. For the RpoA pulldown, the gene identification information is the same as for the HA-SutA pulldown, and number of evidences and total peak intensities are given, ordered by peak intensity.

Dataset A.3: ChIP-Seq and RNA-Seq data per gene. The first six columns give the locus ID for both the PA14 and PAO1 strains, the gene name if available, and genomic locus of the gene. Columns seven and eight give the log_2-transformed ratio between normalized ChIP-Seq counts for the HA-SutA immunoprecipitation versus the mock immunoprecipitation from the strain lacking HA-SutA, and the empirical Bayes F-test corrected p-value (FDR) indicating the probability that this ratio is not different from zero. Columns eight and nine give the log_2-transformed

ratio between normalized ChIP-Seq counts for the RpoA immunoprecipitation from the strain containing HA-SutA versus the RpoA immunoprecipitation from the strain lacking HA-SutA, and the empirical Bayes F-test corrected p-value (FDR) indicating the probability that this ratio is not different from zero. Columns ten through thirteen give the average normalized (RPKM) ChIP-Seq counts per gene for each of the four immunoprecipitation samples. Columns fourteen and fifteen give the log_2-transformed ratio between the normalized RNA-Seq counts in the P_{ara}*:sutA* strain versus the Δ*sutA* strain, and the empirical Bayes F-test corrected p-value (FDR) indicating the probability that this ratio is not different from zero. Columns sixteen through nineteen give the RNA-Seq count ratios between the Δ*sutA* strain and the wild-type strain and between the P_{ara}*:sutA* strain and the wild-type strain, and the empirical Bayes F-test corrected p-value (FDR) indicating the probability that there is no differential expression among the three strains. Columns twenty through twenty-two give the average RNA-Seq RPKM values for each of the three strains. Full raw data, and processed data for individual replicate samples, are available through the NCBI GEO repository (`http://www.ncbi.nlm.nih.gov/geo/`) accession GSE66181.

Dataset A.4: Values for baseline and maximum regions used to scale ChIP data. The RPKM values for the 25 transcriptional units that had the lowest RPKM values in the RpoA immunoprecipitation in the HA-SutA containing strain were used to estimate a baseline level for each ChIP sample. Additionally, the values from the top ten peak regions associated with protein coding genes (where a peak is defined as two consecutive 100bp tiles that fall in the top 100 100bp tiles in the genome) for each type of pulldown were used to estimate the maximum level for each ChIP sample. The baseline and peak regions that were chosen for this analysis are shown in this table. The dynamic range for the HA pulldown in the Δ*sutA*/pMQ72 empty vector strain was significantly lower than those for the other pulldowns, as expected for a control pulldown in which no specific association occurs. For the purpose of comparing association patterns in this strain to association patterns in the Δ*sutA*/pMQ72-HA-SutA strain, reads per 100 bp in the empty vector strain were scaled to the baseline and maximum values observed in the HA-SutA ChIP samples.

Dataset A.5: ChIP-Seq data per 100bp region. This table summarizes ChIP-Seq results by 100bp region. The first column gives the region number for the 100bp region. 100bp regions were numbered in order throughout the genome. Columns

two and three give the log_2-transformed ratio between normalized ChIP-Seq counts for the HA-SutA immunoprecipitation versus the mock immunoprecipitation from the strain lacking HA-SutA, and the empirical Bayes F-test corrected p-value (FDR) indicating the probability that this ratio is not different from zero. Columns four and five give the log_2-transformed ratio between normalized ChIP-Seq counts for the RpoA immunoprecipitation from the strain containing HA-SutA versus the RpoA immunoprecipitation from the strain lacking HA-SutA, and the empirical Bayes F-test corrected p-value (FDR) indicating the probability that this ratio is not different from zero. Columns six through nine give the average normalized (RPKM) ChIP-Seq counts per gene for each of the four immunoprecipitation samples. Columns ten through seventeen give the average scaled values for each immunoprecipitation following linear scaling to the baseline and maximum values described in Dataset A.4, and the standard deviations for the three biological replicates for each immunoprecipitation. Columns eighteen and nineteen give the mean difference between the scaled value for the RpoA pulldown in the strain lacking HA-SutA and the strain containing HA-SutA, and the uncorrected p-value indicating the probability that this difference is not zero. Column twenty indicates whether the 100bp region was included in our "high chip" subset, which satisfied criteria of having scaled ChIP values above a threshold of 0.20 for the HA-SutA ChIP and above 0.25 for the RpoA ChIP in either strain, plus having a statistically significant enrichment in the HA-SutA ChIP compared to the mock control. Full raw data, and processed data for individual replicate samples, are available through the NCBI GEO repository (`http://www.ncbi.nlm.nih.gov/geo/`) accession GSE66181.

Dataset A.6: Transcription unit data. Transcription unit (operon) predictions made by Wurtzel and colleagues were used in this study and are presented here for convenience. Additional information on transcription unit sizes, distances between transcription units, and transcription unit orientation compared to neighboring transcription units is also collected here.

References

(1) Shanks, R. M. Q.; Caiazza, N. C.; Hinsa, S. M.; Toutain, C. M.; O'Toole, G. A. *Appl. Environ. Microb.* **2006**, *72*, 5027–5036.

(2) Choi, K.-H.; Schweizer, H. P. *Nat. Protoc.* **2006**, *1*, 153–161.

(3) Gibson, D. G. In *Methods in Enzymology*; Elsevier: 2011; Vol. 498, pp 349–361.

(4) Fong, J. C. N.; Karplus, K.; Schoolnik, G. K.; Yildiz, F. H. *J. Bacteriol.* **2006**, *188*, 1049–1059.

(5) Pédelacq, J.-D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S. *Nat. Biotechnol.* **2006**, *24*, 79–88.

(6) Link, A. J.; Vink, M. K. S.; Tirrell, D. A. *Nat. Protoc.* **2007**, *2*, 1879–1883.

(7) Szychowski, J.; Mahdavi, A.; Hodas, J. J. L.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A. *J. Am. Chem. Soc.* **2010**, *132*, 18351–18360.

(8) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. *Angew. Chem. Int. Ed.* **2009**, *121*, 10063–10067.

(9) Rappsilber, J.; Mann, M.; Ishihama, Y. *Nat. Protoc.* **2007**, *2*, 1896–1906.

(10) Boersema, P. J.; Raijmakers, R.; Lemeer, S.; Mohammed, S.; Heck, A. J. R. *Nat. Protoc.* **2009**, *4*, 484–494.

(11) Kalli, A.; Hess, S. *Proteomics* **2012**, *12*, 21–31.

(12) Cox, J.; Mann, M. *Nat. Biotechnol.* **2008**, *26*, 1367–1372.

(13) Bagert, J. D.; Xie, Y. J.; Sweredoski, M. J.; Qi, Y.; Hess, S.; Schuman, E. M.; Tirrell, D. A. *Mol. Cell Proteomics* **2014**, *13*, 1352–1358.

(14) Dietrich, L. E. P.; Okegbe, C.; Price-Whelan, A.; Sakhtah, H.; Hunter, R. C.; Newman, D. K. *J. Bacteriol.* **2013**, *195*, 1371–1380.

(15) Müsken, M.; Di Fiore, S.; Dötsch, A.; Fischer, R.; Häussler, S. *Microbiology* **2010**, *156*, 431–441.

(16) Dietrich, L. E. P.; Price-Whelan, A.; Petersen, A.; Whiteley, M.; Newman, D. K. *Mol. Microbiol.* **2006**, *61*, 1308–1321.

(17) Gilbert, K. B.; Kim, T. H.; Gupta, R.; Greenberg, E. P.; Schuster, M. *Mol. Microbiol.* **2009**, *73*, 1072–1085.

(18) Bolger, A. M.; Lohse, M.; Usadel, B. *Bioinformatics* **2014**, *30*, 2114–2120.

(19) Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L. *Genome Biol.* **2009**, *10*, R25.

(20) Fröhlich, K. S.; Vogel, J. *Curr. Opin. Microbiol.* **2009**, *12*, 674–682.

(21) Delhomme, N.; Padioleau, I.; Furlong, E. E.; Steinmetz, L. M. *Bioinformatics* **2012**, *28*, 2532–2533.

(22) Wurtzel, O.; Yoder-Himes, D. R.; Han, K.; Dandekar, A. A.; Edelheit, S.; Greenberg, E. P.; Sorek, R.; Lory, S. *PLoS Pathog.* **2012**, *8*, e1002945.

(23) Law, C. W.; Chen, Y.; Shi, W.; Smyth, G. K. *Genome Biol.* **2014**, *15*, R29.

(24) Mooney, R. A.; Davis, S. E.; Peters, J. M.; Rowland, J. L.; Ansari, A. Z.; Landick, R. *Mol. Cell* **2009**, *33*, 97–108.

(25) Homann, O. R.; Johnson, A. D. *BMC Biology* **2010**, *8*, 49.

(26) Winsor, G. L.; Lam, D. K. W.; Fleming, L.; Lo, R.; Whiteside, M. D.; Yu, N. Y.; Hancock, R. E. W.; Brinkman, F. S. L. *Nucleic Acids Res.* **2011**, *39*, D596–D600.

(27) Tatusov, R. L. et al. *BMC Bioinformatics* **2003**, *4*, 41.

(28) Hunter, J. D. *Comput. Sci. Eng.* **2007**, *9*, 90–95.

(29) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. *Nat. Methods* **2012**, *9*, 671–675.

(30) Liberati, N. T.; Urbach, J. M.; Miyata, S.; Lee, D. G.; Drenkard, E.; Wu, G.; Villanueva, J.; Wei, T.; Ausubel, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2833–2838.

Appendix B

OTHER CONTRIBUTIONS

B.1 Contributions to Ngo, et al.

Published as:

(1) Ngo, J. T.; Babin, B. M.; Champion, J. A.; Schuman, E. M.; Tirrell, D. A. *ACS Chem. Biol.* **2012**, *7*, 1326–1330.

Abstract

Transcriptional activity from a specified promoter can provide a useful marker for the physiological state of a cell. Here we introduce a method for selective tagging of proteins made in cells in which specified promoters are active. Tagged proteins can be modified with affinity reagents for enrichment or with fluorescent dyes for visualization. The method allows state-selective analysis of the proteome, whereby proteins synthesized in predetermined physiological states can be identified. The approach is demonstrated by proteome-wide labeling of bacterial proteins upon activation of the P_{BAD} promoter and the SoxRS regulon and provides a basis for analysis of more complex systems including spatially heterogeneous microbial cultures and biofilms.

Contributions

To evaluate our ability to label cells in response to oxidative stress, I treated *E. coli* engineered to express NLL-MetRS under control of the *soxS* promoter with paraquat and Anl. I reacted labeled cells with akyne-TAMRA and quantified the response by measuring Anl incorporation via fluorescent imaging. The figures below correspond to Figure 3D, Supplementary Figure 2, and Supplementary Figure 3 in the publication. I also contributed to writing the manuscript.

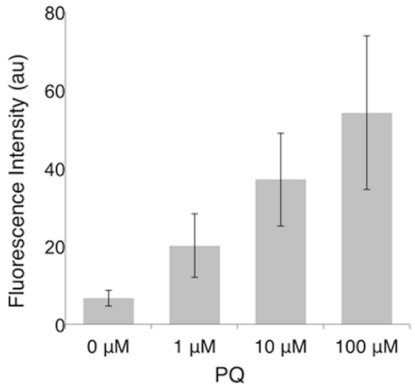

Figure B.1: **Proteomic labeling with Anl under conditions of oxidative stress.** The degree of tagging is sensitive to the level of oxidative stress induction by paraquat. Cells were treated with various concentrations of PQ and pulsed with a fixed concentration of Anl (125 µM). The extent of tagging was assessed by conjugation to alkyne-TAMRA and measurement of individual cell intensities by fluorescence microscopy. Error bars represent the standard deviation of each population examined.

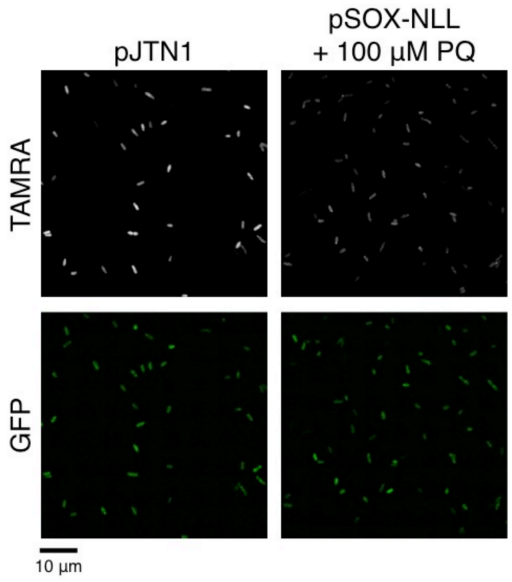

Figure B.2: **Tagging rate in the SoxRS system is less than 10%** *E. coli* cells harboring the pJTN1 plasmid pulsed with Anl under conditions that yield a 10% substitution rate are compared with *E. coli* harboring pSOX-NLL induced with paraquat and pulsed with Anl. Incorporation of Anl is assessed by conjugation with alkyne-TAMRA and subsequent detection by fluorescence microscopy. SoxRS-directed labeling with Anl yields a substitution rate of less than 10%, as fluorescence emission from pJTN1 cells is more intense than that observed with SoxRS-directed labeling. Cells constitutively expressed GFP, which was separately detected to confirm the presence of cells.

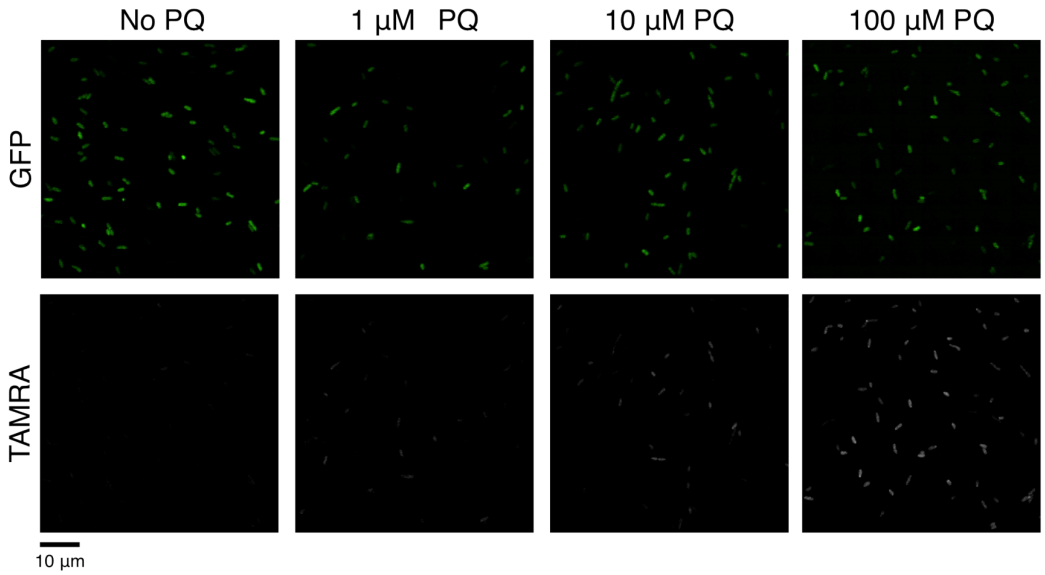

Figure B.3: **Tagging rate in the SoxRS system is dependent on the degree of induction of NLL-MetRS expression.** The NLL-MetRS is under control of the *soxS* promoter and is activated by addition of paraquat (PQ) to the culture medium. As the degree of transcription from the *soxS* promoter is dependent on the concentration of PQ used, so is the level of NLL-MetRS induction. Cells induced with PQ and pulsed with 125 μM Anl exhibit increasing levels of Anl incorporation as more PQ is added. Incorporation of Anl is assessed by conjugation to alkyne-TAMRA and subsequent detection by fluorescence microscopy. Cells constitutively expressed GFP, which was separately detected to confirm the presence of cells.

B.2 Contributions to Hatzenpichler, et al.

Published as:

(1) Hatzenpichler, R.; Scheller, S.; Tavormina, P. L.; Babin, B. M.; Tirrell, D. A.; Orphan, V. J. *Environ. Microbiol.* **2014**, *16*, 2568–2590.

Abstract

Here we describe the application of a new click chemistry method for fluorescent tracking of protein synthesis in individual microorganisms within environmental samples. This technique, termed bioorthogonal non-canonical amino acid tagging (BONCAT), is based on the *in vivo* incorporation of the non-canonical amino acid L-azidohomoalanine (AHA), a surrogate for L-methionine, followed by fluorescent labelling of AHA-containing cellular proteins by azide-alkyne click chemistry. BONCAT was evaluated with a range of phylogenetically and physiologically diverse archaeal and bacterial pure cultures and enrichments, and used to visualize translationally active cells within complex environmental samples including an oral biofilm, freshwater and anoxic sediment. We also developed combined assays that couple BONCAT with ribosomal RNA (rRNA)-targeted fluorescence in situ hybridization (FISH), enabling a direct link between taxonomic identity and translational activity. Using a methanotrophic enrichment culture incubated under different conditions, we demonstrate the potential of BONCAT-FISH to study microbial physiology *in situ*. A direct comparison of anabolic activity using BONCAT and stable isotope labelling by nano-scale secondary ion mass spectrometry ($^{15}NH_3$ assimilation) for individual cells within a sediment-sourced enrichment culture showed concordance between AHA-positive cells and ^{15}N enrichment. BONCAT-FISH offers a fast, inexpensive and straightforward fluorescence microscopy method for studying the *in situ* activity of environmental microbes on a single-cell level.

Contributions

I contributed to the chemical synthesis of Aha; provided advice on Aha labeling experiments, fluorescence gel measurements, and cell imaging; and contributed to writing the manuscript.