ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

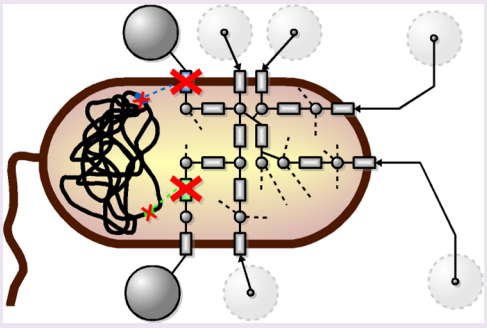
Metabolic Footprinting of Mutant Libraries to Map Metabolite Utilization to Genotype

Richard Baran,[†] Benjamin P. Bowen,[†] Morgan N. Price,[‡] Adam P. Arkin,[‡] Adam M. Deutschbauer,[‡] and Trent R. Northen*,[†]

[†]Life Sciences Division and [‡]Physical Biosciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, California 94720, United States

Supporting Information

ABSTRACT: The discrepancy between the pace of sequencing and functional characterization of genomes is a major challenge in understanding complex microbial metabolic processes and metabolic interactions in the environment. Here, we identified and validated genes related to the utilization of specific metabolites in bacteria by profiling metabolite utilization in libraries of mutant strains. Untargeted mass spectrometry based metabolomics was used to identify metabolites utilized by *Escherichia coli* and *Shewanella oneidensis* MR-1. Targeted high-throughput metabolite profiling of spent media of 8042 individual mutant strains was performed to link utilization to specific genes. Using this approach we identified genes of known function as well as novel transport proteins and enzymes required for the utilization of tested metabolites. Specific examples include two subunits of a predicted ABC transporter encoded by the genes *SO1043* and *SO1044* required for the utilization of citrulline and a predicted histidase encoded by the gene *SO3057* required for the utilization of ergothioneine by *S. oneidensis*. *In vitro* assays with purified proteins showed substrate specificity of *SO3057* toward ergothioneine and histidine betaine in contrast to substrate specificity of a paralogous histidase *SO0098* toward histidine. This generally applicable, high-throughput workflow has the potential both to discover novel metabolic capabilities of microorganisms and to identify the corresponding genes.



The discrepancy between the pace of sequencing and functional characterization of genomes has been recognized as one of the major challenges in microbial genomics.[1] Computational gene annotations have traditionally relied on sequence homology-based assignments of function based on previous biochemical characterization of orthologous gene products from different species.[2,3] Limitations of homology-based functional annotations are widely recognized and are often misleading or incorrect.[4] A range of additional approaches including genome context analysis,[5,6] coexpression,[7] large-scale phenotyping of mutants,[8,9] and others have all contributed to functional annotation of genes. Nonetheless, despite significant efforts in these areas, a large fraction of genes in microbial genomes remains enigmatic, with no known function. Additionally, known biochemical transformations are often not associated with any specific gene (orphan enzymes).[6] Both uncharacterized genes and orphan enzymes limit the engineering of biotechnologically useful microbes and the understanding of the role of microorganisms in health, disease, and global biogeochemical cycles.[6]

Complementing genetic analysis with direct biochemical observations presents an opportunity to establish direct associations between genes and functions. Mass spectrometry (MS) based metabolomics allows the identification of metabolites in complex biological samples[10] with high sensitivity and is well suited for probing the metabolism of microorganisms including discovering novel capabilities.[11,12] Metabolomics experiments can be divided into untargeted or targeted approaches.[13,14] Untargeted experiments are aimed at detecting and identifying the broadest possible set of metabolites. While this often leads to the discovery of novel compounds or metabolic processes, these experiments are very low-throughput. Metabolite profiling of complex samples is usually performed in a hyphenated mode, e.g., by using liquid chromatography (LC) to separate the metabolites prior to analysis by MS. This provides broader metabolite coverage and often sensitivity but comes at the expense of throughput, limiting the number of analyzed samples to a few dozen per day.[15] Targeted approaches focus on the detection of a predefined set of metabolites, often with the aim of quantifying metabolite levels or metabolic fluxes, and are suitable for high-throughput methodologies. MS can also be used in high-throughput either by direct infusion of the sample into the mass spectrometer (without chromatography)[16] or using surface-based methods such as MALDI[17] or NIMS[18] for a limited set of metabolites. In these setups, the coverage of detected compounds in complex samples is decreased due to sample matrix effects.[19,20] Despite the limits in coverage, monitoring

Received: September 6, 2012
Accepted: October 19, 2012
Published: October 19, 2012

multiple metabolites can still be multiplexed in a single analysis.[21] More recently, a high-throughput analysis using direct infusion of over 5000 standard solutions and cell extracts of *Escherichia coli* gene deletion mutants was used to detect hundreds of metabolites.[15]

Untargeted metabolomics approaches to gene annotation have been developed to extend classical functional annotations based on *in vitro* assays of biochemical activity or on phenotypic analysis of mutants. For example, the integration of *in vitro* enzymatic activity assays with complex mixtures of metabolites resulted in identification of two novel phosphatases/phosphotransferases[22] and a hydroxybutyrate dehydrogenase.[23] One of the challenges is the limitation to soluble proteins and correct assay conditions. Integration of untargeted profiling with specific mutants enables identification of biochemical activities without purification. Using this approach Saghatelian *et al.*[24] identified endogenous substrates of mammalian enzyme fatty acid amide hydrolase. Recently, Long *et al.*[25] used HEK293T cell lines transfected with 12 uncharacterized human serine hydrolases coupled with untargeted metabolite profiling to identify ABHD3 as a lipase specific to medium-chain and oxidatively truncated phospholipids. Unfortunately, currently the throughput of untargeted approaches is far from sufficient to address the discrepancy between the pace of sequencing and functional characterization. Allen *et al.*[26] in their breakthrough study showed that raw mass spectra of spent media of yeast (metabolic footprints) can be used to classify the physiological state of WT yeast or 19 specific mutants using multivariate statistical analysis. The ability to identify utilized metabolites for subsequent high-throughput screening would have the favorable attributes of both untargeted (discovery) and targeted approaches (functional assignment).

Here we report a novel integration of untargeted metabolic footprinting to identify uptaken or released metabolites with high-throughput metabolite screening of whole mutant libraries to identify genes related to the utilization of specific metabolites in both *Escherichia coli* K12 (hereafter, *E. coli*) and the metal-reducing bacterium *Shewanella oneidensis* MR-1 (hereafter, *S. oneidensis*) (Figure 1). In the untargeted phase, *E. coli* and *S. oneidensis* were cultured in different complex media. Metabolites utilized by these bacteria were identified using untargeted liquid chromatography coupled to mass spectrometry (LC–MS) by comparing metabolite profiles of spent culture media to control media (metabolic footprinting, Figure 1a). For the screening phase, 10 utilized metabolites were selected for large-scale library screening to link metabolism to specific genes. In total, 8042 individual mutant strains (3901 for *E. coli* and 4141 for *S. oneidensis*) were grown in minimal medium supplemented with target metabolites, and the composition of the spent medium was analyzed by MS in high-throughput in the largest bacterial metabolite screen to date (Figure 1b). The presence of one of the tested metabolites in the spent medium of specific mutants indicated a defect in the utilization of the metabolite pointing to related genes (Figure 1c). Intracellular metabolites of these mutants were then profiled by LC–MS to identify potential accumulation of intermediates related to the utilization of specific metabolites. Whole-genome fitness profiling of *S. oneidensis* mutants was performed to understand the biochemical function of the uptaken metabolites.

RESULTS AND DISCUSSION

Discovery Phase: Metabolic Footprinting of *S. oneidensis* and *E. coli* Identifies Utilized Metabolites.

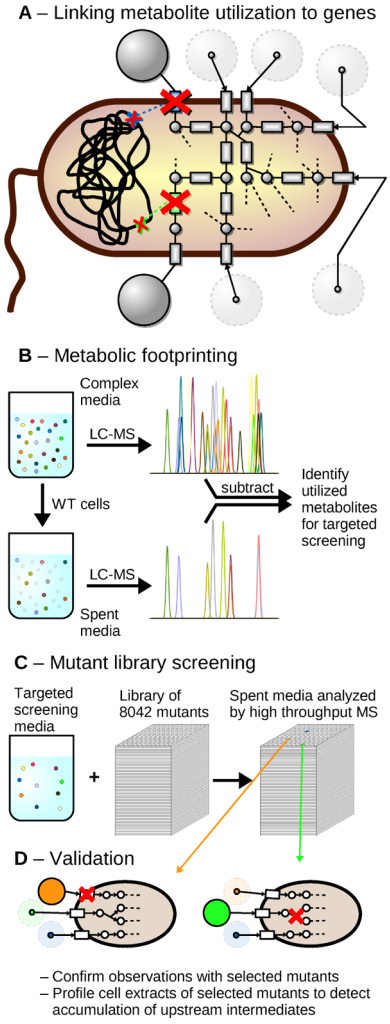


Figure 1. Overview of screening. (A) Mutations in genes encoding transport proteins or enzymes (rectangles) may affect the utilization of specific metabolites (circles). (B) Untargeted metabolic footprinting by LC–MS was used to identify metabolites utilized by wild-type (WT) strains of *E. coli* and *S. oneidensis*. (C) Ten of the utilized metabolites identified in panel B were supplemented to minimal medium for cultivating the mutant strain collections. The spent medium from each individual mutant culture was screened by mass spectrometry (MS) in high-throughput for the 10 metabolites. The presence of a metabolite in the spent medium of a specific mutant is indicative of a potential defect in the utilization of the corresponding metabolite. (D) Illustrations of possible interpretations of metabolite identification in spent medium of single mutant strains. A mutation in a transport protein required for the uptake of the orange metabolite or a mutatation in an intracellular utilization enzyme required to utilize the green metabolite are responsible for the phenotype. Not illustrated is the possibility that certain mutant strains release a metabolite that closely matches the mass spectral features of one of the 10 compounds chosen for high-throughput screening.

To identify exchanged metabolites in *E. coli* and *S. oneidensis*, we performed untargeted metabolic footprinting using methodology as previously described.[27] Wild-type strains of *E. coli* and *S. oneidensis* were grown in minimal medium containing lactate as the carbon source and in the same medium supplemented with yeast extract (YE), a metabolite extract of *Synechococcus* sp. PCC 7002 (SynE), or a metabolite extract of *S. oneidensis* (ShwE). Metabolites from spent media of these cultures were extracted and profiled using LC–MS. Peak areas of metabolites

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189–199

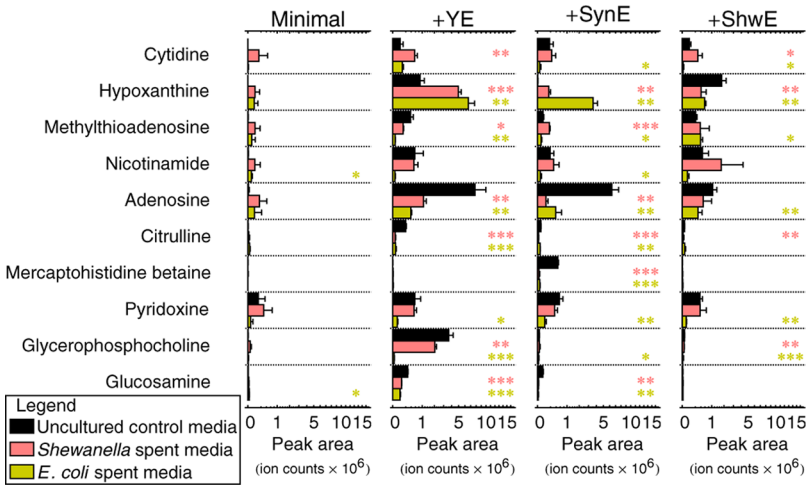


Figure 2. Metabolic footprinting identifies uptake of metabolites in *E. coli* and *S. oneidensis*. Comparison of metabolite levels in different complex spent media and control media (with no bacterial culture) for wild-type *E. coli* and *S. oneidensis* ($n = 4$). Minimal is a defined medium with lactate as the carbon source. The other media are minimal supplemented with one of the complex nutrient sources yeast extract (YE), *Synechococcus* sp. PCC 7002 cell extract (SynE), or *S. oneidensis* MR-1 cell extract (ShwE). Statistically significant differences are highlighted as * ($p < 0.05$), ** ($p < 0.01$), or *** ($p < 0.001$). Supplementary Figure S1 shows this comparison for additional metabolites.

in uncultured control media and spent *S. oneidensis* or *E. coli* media were compared. Figure 2 and Supplementary Figure S1 show this comparison for 71 metabolites that we were able to identify or putatively identify based on our previous studies.[27,28] Both *S. oneidensis* and *E. coli* were able to utilize a broad range of metabolites including amino acids, various dipeptides, and nucleosides (Figure 2; Supplementary Figure S1). Significant release of nucleobases adenine, xanthine, hypoxanthine, and uracil was observed for both bacteria in minimal media supplemented with YE or SynE. There were also cases of metabolites that were significantly utilized by *E. coli* but not *S. oneidensis*. These metabolites include nicotinamide, pyridoxine, and various carbohydrates and their derivatives (Figure 2; Supplementary Figure S1).

Selection of Metabolites for Screening Phase. On the basis of the results of metabolic footprinting we selected 10 metabolites for further investigation: citrulline, cytidine, adenosine, 5′-methylthioadenosine (MTA), hypoxanthine, nicotinamide (NA), pyridoxine, glucosamine, glycerophosphocholine, and ergothioneine (ET). These compounds were chosen from the 71 identified metabolites based on commercial availability and different m/z values of their ions suitable for downstream high-throughput analysis by MS. Metabolites corresponding to well annotated genes were selected for validation of the approach along with metabolites with poorly characterized activities. The intention of selecting common metabolites such as adenosine was to identify known genes and use these results as a positive control. The selection of metabolites with less characterized metabolism should increase the likelihood of annotating novel genes. In some instances, particularly for the well-studied model organism *E. coli*, there is a wealth of knowledge regarding the uptake, consumption, and utilization for most of these 10 compounds. In contrast, *S. oneidensis* is less-studied and our understanding of its metabolism is less complete than that of *E. coli* despite the recent publication of a *S. oneidensis* flux balance analysis (FBA) model.[29] Based on predictions derived from the *S. oneidensis* FBA model and growth curve experiments with different sole sources of carbon, nitrogen, sulfur, and phosphorus, *S. oneidensis* can utilize cytidine and adenosine as the sole source of either carbon or nitrogen.[9] Pyridoxine is a component of the vitamin mix used in the standard growth medium for *S. oneidensis*. For the remaining 7 compounds of interest in this study, there was no evidence that *S. oneidensis* can utilize these compounds as sole substrates for growth or even incorporate these compounds into its metabolism.

High-Throughput Screening of Mutant Libraries for Defects in the Utilization of Metabolites. To link the uptake and utilization of the 10 selected metabolites to specific genes, we developed a high-throughput strategy combining genome-wide mutant libraries and metabolomics (Figure 1). We hypothesized that the mercaptohistidine betaine from SynE could be ET based on the comparison of chromatographic and spectral properties to an authentic ET standard (Supplementary Figure S2), and the latter was used for screening. Preliminary studies in *S. oneidensis* and *E. coli* showed that ET was utilized by the prior and not the latter. Hence, further studies on ET were limited to *S. oneidensis*.

We took advantage of large, archived mutant collections of *E. coli* and *S. oneidensis* for high-throughput metabolomic screening. Specifically, we used the 3,901 member *E. coli* KEIO deletion collection in which each strain contains a precise deletion in a unique nonessential open reading frame.[30] For *S. oneidensis*, we used 4,141 transposon mutant strains from a recently described collection.[9] The screened *S. oneidensis* transposon strains contain insertions in 3,174 different genes (some genes are represented by more than one independent transposon strain). The individual *E. coli* deletion and *S. oneidensis* transposon strains were grown in the respective nine and ten metabolite-supplemented minimal media using a high-throughput microplate format. After the cultures reached saturation, we measured the overall growth of each strain (by $OD_{600\ nm}$) to control for mutants that were unable to grow in the supplemented minimal media.

Following the high-throughput analysis of spent media by a rapid LC−MS method (2.2 min per sample), peak areas of ions with m/z values corresponding to $[M + H]^+$ ions of targeted metabolites were determined (Supplementary Table S1). The peak areas were quartile normalized as described in the Methods section. Elevated levels of at least one target ion (peak

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189−199

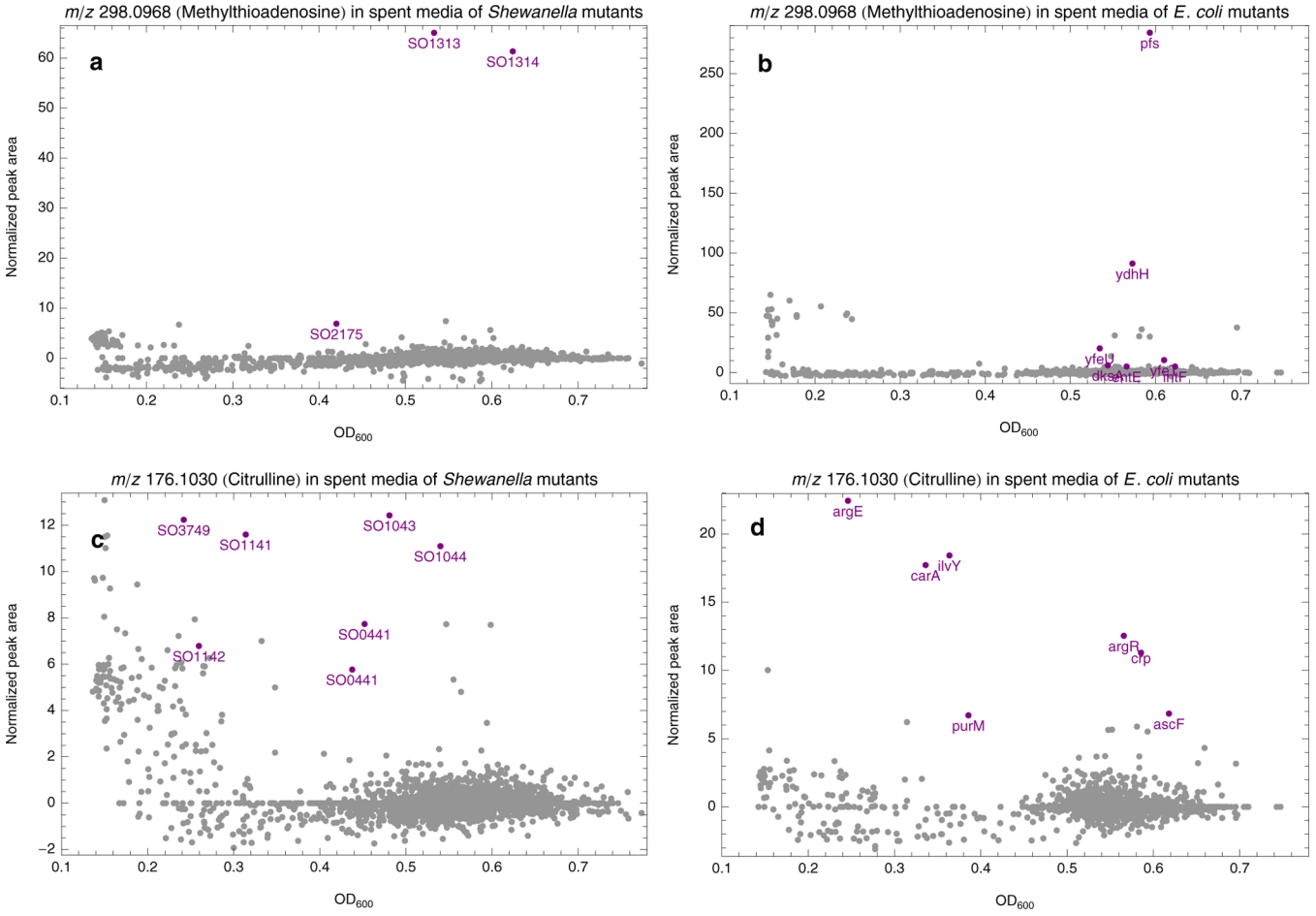


Figure 3. High-throughput mass spectrometry of mutant strain media identifies genes with metabolic defects. Normalized peak areas of selected target ions in spent media extracts of *S. oneidensis* (A, C) and *E. coli* (B, D) mutants were plotted against the $OD_{600\ nm}$ of cultures to discriminate non-growing auxotrophs. Spent media extracts of mutants having elevated levels (5 times the interquartile range above the upper quartile) of the normalized peak areas of at most two target ions are labeled and highlighted in purple. Supplementary Figure S3 shows extended comparison for all targeted ions ranked according to normalized peak areas.

area >1.5 × interquartile range + upper quartile) were found in spent media extracts of 954 *S. oneidensis* mutants and 557 *E. coli* mutants (Figure 3; Supplementary Figure S3). Some mutants (mostly auxotrophs) exhibited limited growth, and significant levels of most of the targeted ions were detected in their spent media and were not investigated further. Other mutants grew comparatively well and showed a significant level of only one or a few target ion(s) in spent media extracts (Figure 3; Supplementary Figure S3). Numbers of tested metabolites of target ions that were significantly elevated in spent media of mutants are listed in Supplementary Figure 3. This number was a critical criterion for judging the specificity of the mutation toward the utilization of a given metabolite. Cases of mutations affecting the utilization of only a single metabolite include a number of *E. coli* strains with mutations in genes of known function and demonstrate the effectiveness of our approach to identify enzymes and transporters required for metabolite uptake and/or utilization. Utilization mutants were detected in enzymes known to hydrolyze the targeted metabolites such as *pfs* (5′-methylthioadenosine/*S*-adenosylhomocysteine nucleosidase) for MTA[31,32] (Figure 3b) and *pncA* (pyrazinamidase/nicotinamidase) for NA[33] (Supplementary Figure S3h). Additionally, individual *E. coli* mutants with deletions in three subunits of mannose PTS permease (*manX*, *manY*, and *manZ*), known to import glucosamine into cells,[34] showed increased levels of the corresponding ion in spent media extracts (Supplementary Figure S3t). On the other hand, no mutants specifically affected in the utilization of only adenosine were detected as the spent media with increased levels of adenosine also showed increased levels of target ions of additional tested metabolites (Supplementary Figure 3ij).

Validation of Mutations and Defects in Metabolite Utilization for Selected Mutants. Based on the genome-wide screening results, 24 *E. coli* and 12 *S. oneidensis* genes were selected for further validation (see Supplementary Table S2 for list). Individual mutant clones and a wild-type control for each bacterium were cultured in minimal medium supplemented with a single metabolite for which the defect in utilization was found during screening. To minimize confounding factors, optical densities of mutant cultures were measured to verify growth using methodology identical to that used for screening the entire library of mutants. Additionally, individual mutant clones were also cultured in minimal medium without supplemented metabolites to verify that they were not producers of the target metabolites. Spent media extracts were analyzed by the rapid LC−MS method, and peak areas of parent ions and characteristic fragments were used for high-specificity quantification (Supplementary Figure S4). For 12 genes (7 in *E. coli* and 5 in *S. oneidensis*), we were able to verify that the mutations affected the metabolite utilization using the

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189−199

strict growth criteria and the more sensitive MS approach described above (see Table 1 for list of validated genes). For

Table 1. Validated *E. coli* and *S. oneidensis* Mutants with Metabolic Defects

organism	gene(s)	affected metabolite	note
E. coli	*pfs*	MTA	5′-methylthioadenosine/*S*-adenosylhomocysteine nucleosidase[31,32]
E. coli	*pncA*	nicotinamide	pyrazinamidase/nicotinamidase[33]
E. coli	*manX, manY, manZ*	glucosamine	subunits of mannose PTS permease[34]
E. coli	*nagB*	glucosamine	glucosamine-6-phosphate deaminase[35]
E. coli	*anmK* (*ydhH*)	ahMurNAc[a]	anhydro-*N*-acetylmuramic acid kinase[39]
E. coli	*argE*	citrulline	acetylornithine deacetylase[b] [50]
S. oneidensis	*SO3749*	citrulline	non-homologous functional analogue of *argE*[b] [9]
S. oneidensis	*SO1043, SO1044*	citrulline	subunits of an ABC transporter
S. oneidensis	*SO3057*	ET	predicted Pal/Histidase
S. oneidensis	*SO1313, SO1314*	ahMurNAc*	

[a]Anhydro-*N*-acetylmuramic acid was not targeted in screening. [b]Reported role in arginine biosynthesis, not citrulline utilization.

three additional genes identified in the high-throughput screening as hits for MTA, *E. coli anmK* (*ydhH*) and *S. oneidensis SO1313* and *SO1314*, we verified that a compound with MS properties similar to those of MTA, anhydro-*N*-acetylmuramic acid, was released into the control medium with no MTA added (see below for further discussion).

Of the 14 genes with validated metabolic phenotypes identified in this study (Table 1), several are expected based on existing genome annotations. As mentioned previously, *pfs*, *pncA*, *manX*, *manY*, *manZ* all have clear impact on the uptake of their respective metabolites (MTA, NA, and glucosamine). Additionally, mutations in *E. coli nagB*, previously known to be involved in the metabolism of glucosamine,[35] resulted in a decreased glucosamine uptake by our high-throughput approach (Supplementary Figure S3t, S4). Importantly, we also found novel phenotypes for poorly annotated and uncharacterized genes. For example, we found that two independent mutants in *S. oneidensis SO3057*, computationally annotated as a pal/histidase, a histidine degrading enzyme, both reduced the utilization of ET (Supplementary Figure S3m). This gene is a paralog of a predicted pal/histidase encoded by *SO0098*, which was previously shown to be required for the utilization of histidine as a nitrogen source.[9] Interestingly, *SO3057* was not essential for histidine utilization,[9] and mutation in *SO0098* did not significantly affect the utilization of ergothioneine in this study.

Metabolic phenotypes are also detected in connected metabolic pathways and related metabolites. Mutants in arginine biosynthesis-related genes in both *S. oneidensis* (*SO3749*, *SO0277*, *SO1141*, *SO1142*) and *E. coli* (*argE*, *carA*, *carB*, *argR*) exhibited limited utilization of citrulline (Figure 3c,d; Supplementary Figure S4). In addition to arginine biosynthetic genes, it was found that mutants in *S. oneidensis* genes *SO1043* and *SO1044* both affected the utilization of citrulline (Figure 3c; Supplementary Figure S4). *SO1043* and *SO1044* are currently annotated as a permease domain and an amino acid binding protein domains of an ABC transporter, respectively. This transporter is annotated as His/Glu/Gln/Arg/opine family transporter and, based on our results, is also responsible for citrulline transport. We tested arginine utilization in these same mutants (*SO1043* and *SO1044*) and did not find changes in the utilization of arginine (Supplementary Figure S4).

Metabolite Profiling of Cell Extracts of Selected Mutants. Metabolites from cell pellets of selected cultures from the validation stage were extracted and profiled using LC−MS to detect the possible accumulation of intracellular metabolites as a result of defects in metabolic networks. Mutants lacking enzymes participating in the metabolism of tested metabolites would be expected, for example, to accumulate upstream intermediates.[36] These intermediates would be present in higher levels in the cell extracts from test cultures compared to the cell extracts from minimal media cultures or from cultures of other mutants (Supplementary Figure S5).

Metabolite profiling by LC−MS revealed insignificant levels of MTA in cell extracts of mutants *SO1313*, *SO1314*, and *E. coli* mutant *ydhH (anmK)* for all of which an elevated level of the corresponding ion was found in spent media extracts of mutants during screening. Surprisingly, a metabolite putatively identified as anhydro-*N*-acetylmuramic acid was found to be released by the cells (Supplementary Figure S5). The sodium adduct of anhydro-*N*-acetylmuramic acid has a m/z value similar to a protonated form of MTA (298.0897 vs 298.0968) and was not differentiated in the high-throughput screening (Figure 3a,b; Supplementary Figure S4); therefore it was necessary to discriminate this metabolite from MTA based on retention time using LC−MS profiling.

Higher levels of ET, a related metabolite histidine betaine, and additional unidentified metabolites were found in cell extracts of a *S. oneidensis SO3057* mutant strain compared to cell extracts of the wild-type strain grown in a minimal media supplemented with ET (Supplementary Figure S5). Similarly, *E. coli argE* mutant and the functionally related *SO3749 S. oneidensis* mutant showed increases in the levels of citrulline, arginine, upstream metabolites in arginine biosynthesis *N*-acetylglutamate, acetylornithine, and additional unidentified metabolites (Supplementary Figure S5). Mutants in additional genes related to arginine biosynthesis *carA* and *SO0277* (functional homologue of *argF*) also showed increased levels of additional metabolites (Supplementary Figure S5).

Mutant Fitness Profiling in *S. oneidensis* Reveals Significance of Metabolite Uptake. To assess the biological significance of metabolic uptake for citrulline, MTA, hypoxanthine, NA, glucosamine, glycerophosphocholine, and ET, we first tested to see if these compounds could support the growth of *S. oneidensis* as the sole sources of either carbon, nitrogen, or sulfur. We found that none of the 7 compounds were able to support the growth of *S. oneidensis* as a sole substrate in a minimal medium under standard laboratory growth conditions (see Methods). However, given the metabolic footprinting data presented above, it is highly likely that *S. oneidensis* is able to uptake and consume these compounds in some capacity. In the absence of supporting growth of MR-1 as a sole nutrient source, we hypothesized that we could identify the physiological role of the consumed metabolites by looking for rescued phenotypes in mutant strains. To test this, we compared genome-wide fitness data of mutant strains grown in a minimal

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189−199

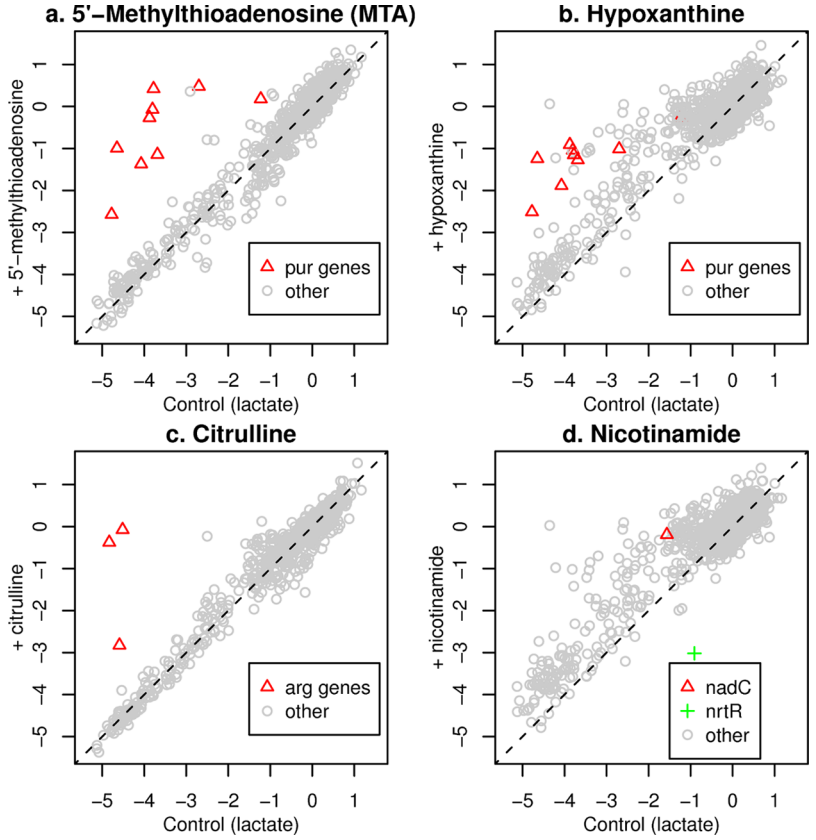


Figure 4. Mutant fitness profiling reveals biological significance of metabolite uptake in *S. oneidensis*. Comparison of mutant fitness values for 3,355 *S. oneidensis* transposon strains grown in a defined lactate minimal media (*x*-axes) or the same media supplemented with a single metabolite identified by metabolic footprinting (*y*-axes): (A) 5′-methylthioadenosine, (B) hypoxanthine, (C) citrulline, (D) nicotinamide. The mutant fitness values are relative values obtained from a pooled fitness assay with DNA barcodes as previously described.[9] Negative values are indicative of growth defects. For each plot, classes of genes whose mutant defects are rescued (for instance, *pur* genes with the addition of 5′-methylthioadenosine or hypoxanthine) or aggravated (*nrtR* with nicotinamide) by the addition of the metabolite are marked.

medium versus a minimal medium supplemented with one of the seven metabolites unexplained by previous growth curve analyses or the FBA model. We performed these mutant fitness experiments using a pooled assay with transposon mutants as recently described (see Methods).[9] For four of the compounds, we found that the fitness defects of some genes in minimal media were rescued by the addition of the metabolite. As illustrated in Figure 4, the addition of either MTA or hypoxanthine rescued purine auxotrophic genes (both metabolites are purines), citrulline rescued arginine biosynthesis genes (citrulline is an intermediate in the pathway), and NA partially rescued mutants in *nadC*, encoding an enzyme in the NAD synthesis pathway (NA is a building block of NAD). Also, NA exacerbated the fitness defect of *nrtR*, which is a repressor of the NAD synthesis pathway.[37] Conversely, the addition of glucosamine, glycerophosphocholine, or ET to minimal media did not rescue any phenotypes (data not shown), implying that these three compounds are not incorporated into biosynthetic pathways that are required for growth in minimal media.

***In Vitro* Validation of Enzymatic Activity of *SO3057*.** To validate the catalytic activity of the predicited pal/histidase from *S. oneidensis* encoded by *SO3057*, which we found by high-throughput screening to be required for the utilization of ET, we cloned the gene in an expression vector in *E. coli* and purified the protein using affinity purification. The gene *SO0098*, encoding a pal/histidase required by *S. oneidensis* for the utilization of histidine as a nitrogen source[9] was also cloned and purified for comparison (Supplementary Figure S6). The purified proteins were incubated with histidine, ET, and metabolite extract of *Synechococcus* (containing histidine betaine), and the compositions of the stopped reaction mixtures were analyzed by LC−MS. Accumulation of urocanate, a product of nonoxidative deamination of histidine, was observed after incubation with the purified *SO0098* protein (Supplementary Figure S7a). Conversion of ET and histidine betaine was observed after incubation with the purified *SO3057* protein (Supplementary Figure S7b,c). Conversion of histidine by *SO3057* protein and conversion of ET or histidine betaine by *SO0098* protein was not observed (Supplementary Figure S7). We repeated the assays with histidine and ET with higher concentrations of substrates and modified concentrations of enzymes and analyzed the stopped reaction mixtures using the rapid LC−MS method. The same catalytic activities and substrate specificities were observed (Figure 5). These results demonstrate the full potential of our system from metabolic footprinting of large mutant libraries to the *in vitro* validation of specific enzymatic activities that could not be predicted on the basis of genome-sequence or automated metabolic models alone.

Discussion. In this study, we aimed to improve functional annotations of *E. coli* and *S. oneidensis* genes by screening metabolite uptake in large mutant libraries using mass spectrometry. Screening spent media extracts of libraries of bacterial mutant strains using minimal medium supplemented

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189−199

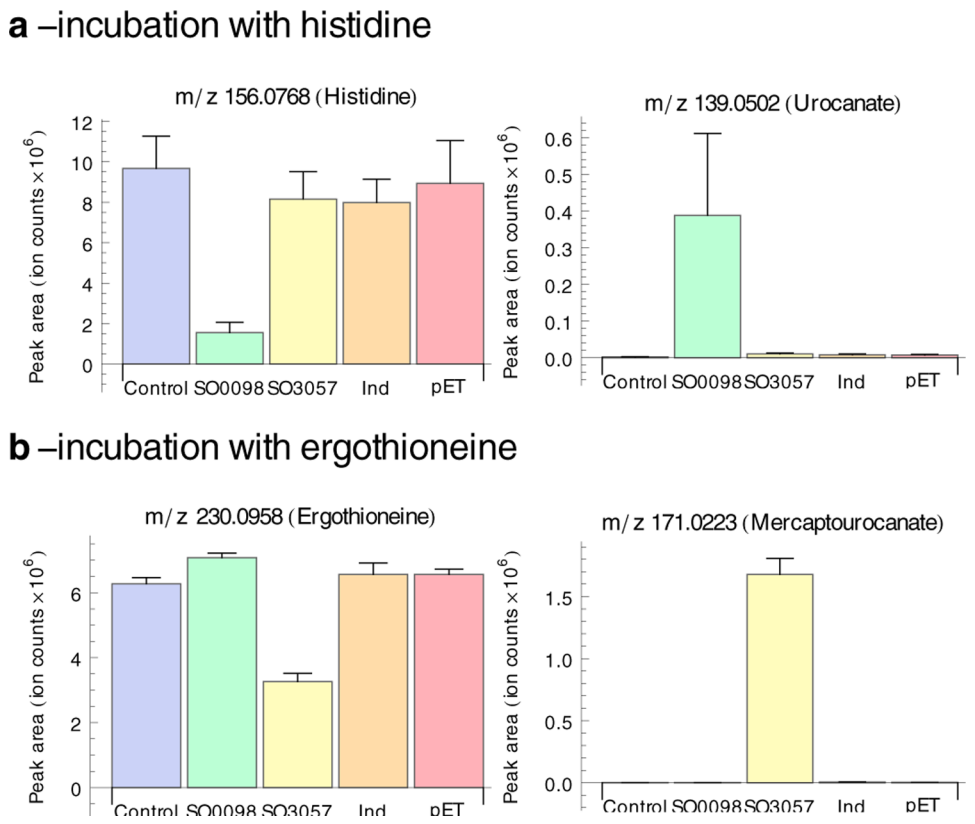


Figure 5. *In vitro* enzymatic activities of purified proteins from *S. oneidensis* encoded by predicted histidases *SO0098* and *SO3057* along with three controls (Control, buffer only; Ind, purification from Induction Control J (Millipore); pET, purification from plasmid pET-32a(+) with no insert). Levels of selected metabolites in stopped reactions mixtures are shown (n = 3). Gene product of *SO0098*, which was previously reported to be required for the utilization of histidine as a nitrogen source,[9] catalyzes the deamination of histidine but does not catalyze the conversion of ergothioneine (A). Gene product of *SO3057* does not deaminate histidine but catalyzes the conversion of ergothioneine to mercaptourocanate as suggested by our high-throughput metabolic footprinting results (B).

with 10 selected metabolites allowed analysis in high-throughput. All 10 metabolites were detectable in spent media extracts using a rapid LC–MS method (2.2 min per sample). However, identification of the metabolites in these screens on the basis of accurate mass only was not sufficient, and further analysis of characteristic fragment ions of these metabolites (Supplementary Figure S4) and downstream metabolite profiling of cell extracts of selected mutants (Supplementary Figure S5) were necessary for full validation.

Genes affecting the uptake or utilization of tested metabolites were detected for half of the metabolites in the study (Table 1). We attribute incomplete coverage to the fact that mutant libraries do not contain mutations in all genes and potential functional redundancy. Genetic redundancy in transport and metabolism is likely to limit the number of defects in metabolite utilization that can be detected using a single gene deletion library. In such cases, a mutation in a single gene would not affect the ability to utilize a given metabolite and would not be detected in this screening. Absence of mutants affecting the specific utilization of only adenosine or cytidine may therefore be caused by the presence of multiple transport systems importing these metabolites into cells or multiple metabolic fates. This approach is thus best suited for screening for metabolites with low-network connectivity. However, the described approach led to the identification of genes of known function as well as putative transport proteins and enzymes (Table 1). Identification of independent mutants in the same gene (*SO3057*), multiple subunits of the same transport complex (*manXYZ* or *SO1043* and *SO1044*), and multiple genes of the same operon or pathway (*SO1313* and *SO1314* or arginine biosynthesis genes) provide additional confidence for the biological origin of the observation. Additionally, using two different bacteria proved useful for the identification of functionally analogous genes. These include *SO3749*, a recently discovered non-homologous functional analogue of *E. coli argE*[9] and *ydhH* (*anmK*) and *SO1313*. Additionally, we used our high-throughput metabolic footprinting data as a baseline for testing specific hypotheses of enzymatic activity by demonstrating the different substrate specificity of *SO3057* (ET, histidine betaine) compared to paralogous pal/histidase *SO0098* (histidine). *S. oneidensis* does not appear to have ergothioneine biosynthesis genes; however, *SO3057* may enable the utilization of ergothioneine produced by other microorganisms.[38]

Elevated levels of anhydro-*N*-acetylmuramic acid were detected in the cell extracts of an *anmK* mutant of *E. coli* (also known as *ydhH*) as well as from *SO1313* and *SO1314* mutants from *S. oneidensis*. AnmK catalyzes the hydrolysis and phosphorylation of anhydro-N-acetylmuramic acid during murein recycling.[39] The following step, hydrolysis of the *N*-acetylmuramic acid 6-phosphate to *N*-acetyl-D-glucosamine-6-phosphate and lactate is catalyzed by MurQ.[40] *SO1313* is homologous to *anmK*, so our data suggest that *SO1313* is also anhydro-*N*-acetylmuramic acid kinase. *S. oneidensis* does not contain any likely ortholog of *murQ*, but *SO1314* could play this role, as it contains a peptidase M23 or lytM domain and is similar to *E. coli yebA*, which has a role in peptidoglycan

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189–199

remodeling.[41] Furthermore, *SO1313* and *SO1314* have similar fitness patterns across diverse conditions (correlation = 0.65).[9] *SO1314* is upstream of *SO1313*, but the similarity in fitness patterns does not seem to be due to polar effects, as knockout strains for *SO1314* but not *SO1313* have reduced abundance when grown with propionate as the carbon source.[9] In addition to mutations in *SO1043* and *SO1044*, which encode two different subunits of an ABC transport system, citrulline utilization was also affected by mutations in genes of arginine biosynthesis pathway. Citrulline is an intermediate in arginine biosynthesis; however, the mutations affecting citrulline utilization (most notably *argE* and *SO3749*) are upstream of citrulline in the pathway This suggests that citrulline may be utilized by the reverse of the arginine biosynthesis pathway via ornithine, at least in certain mutant strains.

Genes of both cytoplasmic enzymes as well as membrane proteins were linked to the utilization of tested metabolites in this study. Annotation of transport proteins is especially valuable as *in vitro* studies are challenging and computational homology-based annotations can be misleading since even closely related transporters have different specificities.[42] Using multiple metabolites in the supplemented screening media allows parallel identification of corresponding genes, providing a significant speed improvement compared to targeted assays. The throughput of the rapid LC−MS assay used in this study to screen for the utilization of 10 metabolites was approximately 660 mutants per day (2.2 min per sample), resulting in approximately two weeks of net running time of mutant spent media analysis by LC−MS during the screening stage. Increasing the scale and throughput of the mutant spent media analysis is of critical importance to enable large scale applicability of our approach. These increases can be achieved with a higher degree of multiplexing (using more test metabolites), application of faster surface-based mass spectrometry methods,[18] or smart-pooling[43,44] spent media extracts prior to analysis. Coupling of acoustic printing to NIMS analysis has been shown to decrease the analysis throughput below 1 min per sample.[45] In addition to improving the throughput and coverage of the experimental screening method, increases in the discovery rate can be achieved by a suitable selection of target metabolites. Use of complex media that approximate the composition of environmental niches of tested microbes for untargeted metabolic footprinting can point to the utilization of unexpected metabolites in natural contexts. Alternatively, target metabolites can also be selected on the basis of their previously reported utilization and lack of detailed biochemical characterization of underlying metabolic pathways to target specific orphan enzymes.[6]

We have demonstrated that metabolite utilization screening in bacterial mutant libraries can provide useful biochemical information allowing gene identification and annotation. This approach is not intended to substitute for *in vitro* characterization. Rather, the intention is to bridge the gap between widely used high-throughput phenotyping assays and downstream metabolic processes. For instance, we used our high-throughput results to predict the catalytic activity of *SO3057*, which we validated using an *in vitro* assay. As shown here, the screening can be preceded by the discovery of novel metabolic capabilities identified using untargeted metabolite profiling approaches. The detection of multiple mutations in the same genes and multiple genes within operons using high-throughput screening can be used to increase confidence in annotations. In this initial work using *E. coli* and *S. oneidensis* mutant libraries we identified 14 genes related to the utilization of specific metabolites. Extension of this approach to larger metabolite libraries and intracellular metabolites could be used to provide additional annotations. We believe that this workflow can contribute significantly to improve the functional characterization of microbial genomes, improve the mapping of microbial metabolism and map metabolic interactions among microorganisms and their environment.

METHODS

Chemicals and Strains. Chemicals for culture media preparation and solvents for LC−MS analysis were purchased from Sigma. Ergothioneine was purchased from American Advanced Scientific Inc. and Sigma. Wild-type *Shewanella oneidensis* MR-1 was purchased from ATCC (700550). For wild-type *E. coli* K12, we used the parental strain of the KEIO deletion collection, BW25113. *Synechococcus* sp. PCC 7002 was purchased from ATCC and cultured in A+ medium[46] without TRIZMA base as described previously.[27] The construction of the *S. oneidensis* transposon mutant collection is previously described.[9] For the high-throughput metabolomic screening used in this study, we used the 4141 individual *S. oneidensis* strains of the upPool collection.[9] For *E. coli*, we assayed the indidivual deletion mutants of the KEIO collection.[30]

Media and Microbial Culturing. We used the same base minimal medium with lactate as a carbon source for all experiments in this study, both *S. oneidensis* and *E. coli*. This base medium contained salts (per liter: 1.5 g NH_4Cl, 0.1 g KCl, 1.75 g NaCl, 0.61 g $MgCl_2{\cdot}6H_20$, 0.6 g NaH_2PO_4), 30 mM DL-lactate, Wolfe's vitamins, and Wolfe's minerals. We adjusted the pH to 7.0 with NaOH. For metabolic footprinting experiments, this standard medium was supplemented with either yeast extract to a final concentration of 0.1% (w/v), a cellular extract of *S. oneidensis*, or a cellular extract of *Synechococcus* sp. PCC 7002. The *S. oneidensis* extract was prepared from a 500 mL saturated culture of wild-type *S. oneidensis* grown in our baseline minimal medium. The metabolite extract of *Synechococcus* sp. PCC 7002 was prepared from a 500 mL culture using methanol extraction as described previously.[27] For high-throughput screening of the microbial mutant libraries, we supplemented the baseline medium with a mixture of metabolites: citrulline (50 μM), cytidine (50 μM), adenosine (50 μM), 5′-methylthioadenosine (10 μM), hypoxanthine (10 μM), nicotinamide (10 μM), pyridoxine (10 μM), glucosamine (100 μM), glycerophosphocholine (50 μM), and ergothioneine (50 μM). For screening of the *E. coli* KEIO collection, we omitted ET. For validation studies of single mutants, we supplemented the minimal medium with a single metabolite at the following final concentrations. For pooled *S. oneidensis* MR-1 mutant fitness experiments, we supplemented the base minimal medium with either citrulline (50 μM final concentration), 5′-methylthioadenosine (50 μM), hypoxanthine (10 μM), nicotinamide (10 μM), glucosamine (100 μM), glycerophosphocholine (50 μM), or ergothioneine (50 μM).

For all high-throughput culturing of the bacterial mutant strains, we used miniaturized growth assays in 96-deep-well microplates (Corning 3960). All liquid handling described below was performed using the AP96 tool of a Biomek FxP robot with AP250 filter tips. All microplates were grown in a Multitron II incubator shaker set to 30 °C for *S. oneidensis* and 37 °C for *E. coli*. The 4141 strain *S. oneidensis* transposon upPool collection is stored in 47 96-well microplates; the 3901 strain *E. coli* KEIO collection is stored in 48 96-well microplates. For *S. oneidensis*, we first inoculated 5 μL of each glycerol stock into 295 μL of LB and grew the cells overnight to saturation. We inoculated 2 μL of these LB overnight cultures into 400 μL of our standard lactate minimal medium and grew the cells in minimal medium to saturation (around 36 h total). For *E. coli*, we directly inoculated 1 μL of the glycerol stock into 400 μL of our standard lactate minimal medium and grew the cells for ∼18 h until they reached saturation. After growth in minimal medium, we checked the optical density ($OD_{600\ nm}$) of each culture (diluted 1:4 in water) in a Safire II microplate reader (Tecan). The remainder of the cultures were pelleted at 4000 rpm for 10 min in a microplate centrifuge (Eppendorf). Using a Biomek FxP,

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189−199

we carefully removed 200 μL of the supernatant (spent medium) into a fresh microplate for downstream metabolite analyses. Both spent media supernatants and cell pellets were frozen and kept at −20 °C prior to metabolite extraction.

Metabolite Extraction. Samples (1.8 mL) of culture media supernatants from untargeted metabolic footprinting experiments were dried down using a Savant SpeedVac Plus SC110A. The samples were redissolved in 500 μL of methanol and left overnight at 4 °C. The samples were then centrifuged for 10 min at 2348 × *g*, dried down, redissolved in 100 μL of methanol, left overnight at 4 °C, filtered using 0.20 μm PVDF membrane microcentrifugal filters (Millipore), and analyzed using LC−MS. Spent culture media supernatants from high-throughput screening experiments (120 μL) and validation experiments (120 μL) were freeze-dried in 96-well format using Labconco FreeZone 2.5 freeze-dry system. The samples were redissolved in 120 μL of methanol, left overnight at 4 °C and centrifuged for 10 min 3320 × *g*. The 96-well plates were combined into 384-well plates by taking 50 μL of the supernatant. The samples in 384-well plates were analyzed by LC−MS. These extractions and analyses were performed in batches using eight or fewer 96-well plates per batch.

Extraction of metabolites from cell pellets of selected mutants was performed by resuspending frozen cell pellets in 1 mL of cold (−20 °C) methanol, keeping the suspension at −20 °C for 1 h, centrifuging for 10 min at 2348 × *g*, drying down the supernatant, and resuspending in 100 μL of methanol. The samples were then kept overnight at 4 °C and filtered using 0.20 μm PVDF membrane microcentrifugal filters (Millipore) prior to analysis by LC−MS.

Analytical Conditions. LC−MS analysis of samples from untargeted metabolic footprinting experiments and cell extracts of selected mutants was performed using normal phase liquid chromatography (ZIC-HILIC capillary column, 150 mm × 1 mm, 3.5 μm 100 Å, Merck Sequant; Agilent 1200 series capillary LC system) coupled to a time-of-flight mass spectrometer (Agilent 6520 dual-ESI-Q-TOF) as described previously.[28] The acquisition was performed in fast polarity switching mode.

High-throughput analysis of spent media extracts of deletion mutant cultures was performed using the same LC−MS system with a ZORBAX SB-C18 30 mm × 1 mm, 3.5 μm (Agilent) under isocratic conditions using 60% acetonitrile with 0.1% formic acid (v/v) as the mobile phase. The flow rate was 80 μL/min, injection volume was 1 μL, and the stop time of the pump was set at 1.8 min. An overlapped injection starting at 1.1 min was used resulting in approximately 2.2 min of cycle time per sample. The high-throughput LC−MS analysis of the samples was performed in the same batches as the metabolite extraction.

Data Analysis. Raw data sets from untargeted metabolic footprinting experiments were exported into mzData format using Agilent MassHunter Workstation Software Qualitative Analysis (Version B.03.01). These data sets were processed using the MathDAMP package[47] to highlight differences between metabolite profiles of control media, spent *S. oneidensis* media, and spent *E. coli* media as described previously.[27] These comparisons were performed for the four types of tested media (minimal, minimal + YE, minimal + SynE, minimal + ShwE). Identified or putatively identified metabolites from our previous studies[27,28] were highlighted in our comparison results based on the correspondence of retention times and mass spectral features. Bounds of characteristic peaks ($[M + H]^+$ and $[M - H]^-$ ions) of identified or putatively identified metabolites in their mass spectra were identified manually in all data sets, and their peak areas were integrated with a ±20 ppm integration window for relative comparisons.

Raw data sets from high-throughput analyses of spent media from deletion mutants were exported into mzData format. Areas of $[M + H]^+$ peaks of targeted metabolites were integrated with a ±100 ppm integration window over a fixed retention time range (0.5−0.9 min). A broader integration window was chosen for these data sets to avoid false negatives resulting from temporal decreases in mass accuracy during the measurement. For each targeted metabolite, quartiles of peak areas were calculated within each extraction and LC−MS analysis batch. Outliers among peak areas of specific metabolites and specific deletion mutants within each batch were determined by subtracting the peak area of the metabolite from the upper quartile (for peak areas > upper quartile) or lower quartile (for peak areas < lower quartile) and divided by the interquartile range. This measure was used to rank deletion mutants according to the amounts of metabolites not utilized from the media for each metabolite.

Fitness Assays. Pooled fitness assays of the S. oneidensis MR-1 mutant upPool and dnPool were performed as previously described.[9] Briefly, the *S. oneidensis* transposon mutant strains contain unique DNA TagModules that enable the pooling and parallel fitness analysis of thousands of strains in parallel using a microarray readout.[48] We grew aliquots of the upPool and dnPool in minimal medium supplemented with one of seven different metabolites. After ∼7 population doublings, we assayed the DNA tag abundance for all strains via microarray hybridization (expt) and compared these tag signals to those from the preinoculation culture (start). The relative fitness of all mutants in the pools is calculated as $\log_2$ (tag $\text{signal}_{\text{expt}}$/tag $\text{signal}_{\text{start}}$). Negative $\log_2$ ratios are indicative of a strain with a relative growth defect compared to the average strain in the mutant pool.

Cloning, Protein Purification, and *in Vitro* Enzyme Assays. We amplified *SO0098* from *S. oneidensis* genomic DNA with primers SO0098_for (ggccatatgcaccatcatcatcatcatAAATCAGTCAATCATTTAGTATT) and SO0098_rev (gaaccgcgtggcaccagaccagaagaTCAGAGGCTAGGTAAAACACCTAAG). *SO3057* was amplified with primers SO3057_for (ggccatatgcaccatcatcatcatcatGAACAAGTGGTGCTTGACGGTAAA) and SO3057_rev (gaaccgcgtggcaccagaccagaagaTTAATAATGTGTAATAAAGTCAGT). The expression vector pET-32a(+) (Millipore) was amplified and linearized using primers pET32a_for (tcttctggtctggtgccacgcggttc) and pET32a_rev (atgatgatgatgatggtgcatatggcc). We used Phusion high-fidelity DNA polymerase (New England Biolabs) for all PCR reactions. The lowercase and underlined sequences represent homology regions used to generate the final pET-32a(+) vectors with *SO0098* (APA720) or *SO3057* (APA721) by Gibson assembly[49] using the manufacturer's recommendations (New England Biolabs). The final vectors encode N-terminal 6X-His tagged versions of the two *S. oneidensis* histidases, which we transformed into *E. coli* BL21(DE3) cells (New England Biolabs). To induce expression of our recombinant proteins, we grew the two transformed strains along with two control strains (Ind, Induction Control J (Millipore); pET, strain with plasmid pET-32a(+) without an insert) in LB medium to $OD_{600\ nm}$ of 0.8−1 and induced with IPTG with final concentration of 1 mM for 2.5 h. A total of 10 mL of each cell culture was used for protein purification. Protein purification was performed using TALON kit (Clontech) in batch/gravity flow mode according to the manufacturer's protocol and as described previously.[22] 2 mL of TALON metal affinity resin (Clontech) were placed into TALON disposable gravity flow columns (Clontech) and equilibrated with 12 × 1 mL of TALON equilibration buffer (Clontech). Two ml of HisTALON xTractor buffer containing ProteoGuard protease inhibitor cocktail (Clontech) were added to each cell pellet, cells were resuspended and sonicated on ice 3 × 20 s using a Misonix S-4000 sonicator with a fine tip probe and a power output of 50. Protein purification was performed at 4 °C. Supernatants were loaded onto equilibrated columns and placed on a vertical shaker for 20 min. Each column was then washed with 5 × 1 mL of wash buffer (66:934 v/v HisTALON elution buffer: HisTALON equilibration buffer, Clontech) and eluted with HisTALON elution buffer containing 100 mM imidazole (Sigma). Brandford assay with BSA as calibration standard was used to estimate protein concentrations in the crude extract and all fractions. SDS PAGE (Bio-Rad) was performed to verify the size of purified proteins and estimate their purity (Supplementary Figure S6). A 400 μL sample from the third elution fraction (containing the highest amount of total protein) for each purification was taken, and the buffer was exchanged to assay buffer (50 mM MOPS, 10 mM KCl, 10 mM $MgSO_4$, 100 μM $MnCl_2$, pH 7.2) using 10 kDa microcentrifugal ultrafiltration units (Millipore). The final volume was adjusted to 400 μL after buffer exchange. Five microliters of purified protein solutions in the assay buffer were added to 15 μL of 1 mM histidine, 1 mM ergothioneine, or metabolite extract of *Synechococcus* (all in assay buffer) and incubated at RT. After 2 h, 80

dx.doi.org/10.1021/cb300477w | *ACS Chem. Biol.* 2013, 8, 189−199

μL of methanol were added to the assay mixtures, and the samples were filtered using a 0.20 μm PVDF membrane microcentrifugal filters (Millipore) and analyzed by LC−MS. Then 200 μL of glycerol was added to the remaining protein extracts in the assay buffer, and the solutions were stored at −20 °C. For the follow-up enzyme assays using the rapid LC−MS method, the buffer of the stored protein extracts was changed to a modified assay buffer (5 mM MOPS, 10 mM KCl, 10 mM $MgCl_2$, 100 μM $MnCl_2$, pH 7.2). The extract corresponding to *SO0098* was concentrated 1.7 times and the extract corresponding to *SO3057* was diluted 5 times compared to the previous assay. The final concentrations of histidine and ET in the incubation mixtures were 7.5 mM.

■ ASSOCIATED CONTENT

Supporting Information

This material is free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: trnorthen@lbl.gov.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank K. Wetmore for assistance in high-throughput bacterial culturing and J. Ray for assistance with protein purification. This work conducted by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

■ REFERENCES

(1) Galperin, M. Y., and Koonin, E. V. (2010) From complete genome sequence to 'complete' understanding? *Trends Biotechnol. 28*, 398−406.

(2) Koonin, E. V., and Wolf, Y. I. (2008) Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Res. 36*, 6688−6719.

(3) Gerdes, S., El Yacoubi, B., Bailly, M., Blaby, I. K., Blaby-Haas, C. E., Jeanguenin, L., Lara-Núñez, A., Pribat, A., Waller, J. C., Wilke, A., Overbeek, R., Hanson, A. D., and de Crécy-Lagard null, V. (2011) Synergistic use of plant-prokaryote comparative genomics for functional annotations. *BMC Genomics 12* (Suppl1), S2.

(4) Gerlt, J. A., Allen, K. N., Almo, S. C., Armstrong, R. N., Babbitt, P. C., Cronan, J. E., Dunaway-Mariano, D., Imker, H. J., Jacobson, M. P., Minor, W., Poulter, C. D., Raushel, F. M., Sali, A., Shoichet, B. K., and Sweedler, J. V. (2011) The enzyme function initiative. *Biochemistry 50*, 9950−9962.

(5) Fonknechten, N., Perret, A., Perchat, N., Tricot, S., Lechaplais, C., Vallenet, D., Vergne, C., Zaparucha, A., Paslier, D. L., Weissenbach, J., and Salanoubat, M. (2009) A conserved gene cluster rules anaerobic oxidative degradation of L-ornithine. *J. Bacteriol. 191*, 3162−3167.

(6) Hanson, A. D., Pribat, A., Waller, J. C., and de Crécy-Lagard, V. (2010) 'Unknown' proteins and 'orphan' enzymes: the missing half of the engineering parts list−and how to find it. *Biochem. J. 425*, 1−11.

(7) Horan, K., Jang, C., Bailey-Serres, J., Mittler, R., Shelton, C., Harper, J. F., Zhu, J.-K., Cushman, J. C., Gollery, M., and Girke, T. (2008) Annotating genes of known and unknown function by large-scale coexpression analysis. *Plant Physiol. 147*, 41−57.

(8) Nichols, R. J., Sen, S., Choo, Y. J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K. M., Lee, K. J., Wong, A., Shales, M., Lovett, S., Winkler, M. E., Krogan, N. J., Typas, A., and Gross, C. A. (2011) Phenotypic landscape of a bacterial cell. *Cell 144*, 143−156.

(9) Deutschbauer, A., Price, M. N., Wetmore, K. M., Shao, W., Baumohl, J. K., Xu, Z., Nguyen, M., Tamse, R., Davis, R. W., and Arkin, A. P. (2011) Evidence-based annotation of gene function in Shewanella oneidensis MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genet. 7*, e1002385.

(10) Garcia, D. E., Baidoo, E. E., Benke, P. I., Pingitore, F., Tang, Y. J., Villa, S., and Keasling, J. D. (2008) Separation and mass spectrometry in microbial metabolomics. *Curr. Opin. Microbiol. 11*, 233−239.

(11) Baran, R., Reindl, W., and Northen, T. R. (2009) Mass spectrometry based metabolomics and enzymatic assays for functional genomics. *Curr. Opin. Microbiol. 12*, 547−552.

(12) Saito, N., Ohashi, Y., Soga, T., and Tomita, M. (2010) Unveiling cellular biochemical reactions via metabolomics-driven approaches. *Curr. Opin. Microbiol. 13*, 358−362.

(13) Lu, W., Bennett, B. D., and Rabinowitz, J. D. (2008) Analytical strategies for LC-MS-based targeted metabolomics. *J. Chromatogr., B 871*, 236−242.

(14) Patti, G. J. (2011) Separation strategies for untargeted metabolomics. *J. Sep. Sci. 34*, 3460−3469.

(15) Fuhrer, T., Heer, D., Begemann, B., and Zamboni, N. (2011) High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. *Anal. Chem. 83*, 7074−7080.

(16) Liesener, A., and Karst, U. (2005) Monitoring enzymatic conversions by mass spectrometry: a critical review. *Anal. Bioanal. Chem. 382*, 1451−1464.

(17) Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 Da. *Anal. Chem. 60*, 2299−2301.

(18) Northen, T. R., Yanes, O., Northen, M. T., Marrinucci, D., Uritboonthai, W., Apon, J., Golledge, S. L., Nordström, A., and Siuzdak, G. (2007) Clathrate nanostructures for mass spectrometry. *Nature 449*, 1033−1036.

(19) Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T. (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom. 2*, 151−153.

(20) Annesley, T. M. (2003) Ion suppression in mass spectrometry. *Clin. Chem. 49*, 1041−1044.

(21) Reindl, W., Deng, K., Gladden, J. M., Cheng, G., Wong, A., Singer, S. W., Singh, S., Lee, J. C., Yao, C. H., Hazen, T. C., Singh, A. K., Simmons, B. A., Adams, P. D., and Northen, T. R. (2011) Colloid-based multiplexed screening for plant biomass-degrading glycoside hydrolase activities in microbial communities RID C-1373−2010. *Energy Environ. Sci. 4*, 2884−2893.

(22) Saito, N., Robert, M., Kitamura, S., Baran, R., Soga, T., Mori, H., Nishioka, T., and Tomita, M. (2006) Metabolomics approach for enzyme discovery. *J Proteome Res. 5*, 1979−1987.

(23) Saito, N., Robert, M., Kochi, H., Matsuo, G., Kakazu, Y., Soga, T., and Tomita, M. (2009) Metabolite profiling reveals YihU as a novel hydroxybutyrate dehydrogenase for alternative succinic semialdehyde metabolism in Escherichia coli. *J. Biol. Chem. 284*, 16442−16451.

(24) Saghatelian, A., Trauger, S. A., Want, E. J., Hawkins, E. G., Siuzdak, G., and Cravatt, B. F. (2004) Assignment of endogenous substrates to enzymes by global metabolite profiling. *Biochemistry 43*, 14332−14339.

(25) Long, J. Z., Cisar, J. S., Milliken, D., Niessen, S., Wang, C., Trauger, S. A., Siuzdak, G., and Cravatt, B. F. (2011) Metabolomics annotates ABHD3 as a physiologic regulator of medium-chain phospholipids. *Nat. Chem. Biol. 7*, 763−765.

(26) Allen, J., Davey, H. M., Broadhurst, D., Heald, J. K., Rowland, J. J., Oliver, S. G., and Kell, D. B. (2003) High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol. 21*, 692−696.

(27) Baran, R., Bowen, B. P., and Northen, T. R. (2011) Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by Synechococcus sp. PCC 7002. *Mol. Biosyst. 7*, 3200−3206.

(28) Baran, R., Bowen, B. P., Bouskill, N. J., Brodie, E. L., Yannone, S. M., and Northen, T. R. (2010) Metabolite identification in Synechococcus sp. PCC 7002 using untargeted stable isotope assisted metabolite profiling. *Anal. Chem. 82*, 9034−9042.
(29) Pinchuk, G. E., Hill, E. A., Geydebrekht, O. V., De Ingeniis, J., Zhang, X., Osterman, A., Scott, J. H., Reed, S. B., Romine, M. F., Konopka, A. E., Beliaev, A. S., Fredrickson, J. K., and Reed, J. L. (2010) Constraint-based model of Shewanella oneidensis MR-1 metabolism: a tool for data analysis and hypothesis generation. *PLoS Comput. Biol. 6*, e1000822.
(30) Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol. 2*, 2006.0008.
(31) Della Ragione, F., Porcelli, M., Cartenì-Farina, M., Zappia, V., and Pegg, A. E. (1985) Escherichia coli S-adenosylhomocysteine/5′-methylthioadenosine nucleosidase. Purification, substrate specificity and mechanism of action. *Biochem. J. 232*, 335−341.
(32) Cornell, K. A., and Riscoe, M. K. (1998) Cloning and expression of Escherichia coli 5′-methylthioadenosine/S-adenosylhomocysteine nucleosidase: identification of the pfs gene product. *Biochim. Biophys. Acta 1396*, 8−14.
(33) Pardee, A. B., Benz, E., Jr., St Peter, D. A., Krieger, J. N., Meuth, M., and Trieshmann, H., Jr. (1971) Hyperproduction and purification of nicotinamide deamidase, a microconstitutive enzyme of Escherichia coli. *J. Biol. Chem. 246*, 6792−6796.
(34) Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1993) Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev. 57*, 543−594.
(35) Calcagno, M., Campos, P. J., Mulliert, G., and Suástegui, J. (1984) Purification, molecular and kinetic properties of glucosamine-6-phosphate isomerase (deaminase) from Escherichia coli. *Biochim. Biophys. Acta 787*, 165−173.
(36) Ohashi, Y., Hirayama, A., Ishikawa, T., Nakamura, S., Shimizu, K., Ueno, Y., Tomita, M., and Soga, T. (2008) Depiction of metabolome changes in histidine-starved Escherichia coli by CE-TOFMS. *Mol. Biosyst. 4*, 135−147.
(37) Rodionov, D. A., De Ingeniis, J., Mancini, C., Cimadamore, F., Zhang, H., Osterman, A. L., and Raffaelli, N. (2008) Transcriptional regulation of NAD metabolism in bacteria: NrtR family of Nudix-related regulators. *Nucleic Acids Res. 36*, 2047−2059.
(38) Seebeck, F. P. (2010) In vitro reconstitution of mycobacterial ergothioneine biosynthesis. *J. Am. Chem. Soc. 132*, 6632−6633.
(39) Uehara, T., Suefuji, K., Valbuena, N., Meehan, B., Donegan, M., and Park, J. T. (2005) Recycling of the anhydro-N-acetylmuramic acid derived from cell wall murein involves a two-step conversion to N-acetylglucosamine-phosphate. *J. Bacteriol. 187*, 3643−3649.
(40) Uehara, T., Suefuji, K., Jaeger, T., Mayer, C., and Park, J. T. (2006) MurQ Etherase is required by Escherichia coli in order to metabolize anhydro-N-acetylmuramic acid obtained either from the environment or from its own cell wall. *J. Bacteriol. 188*, 1660−1662.
(41) Uehara, T., Dinh, T., and Bernhardt, T. G. (2009) LytM-domain factors are required for daughter cell separation and rapid ampicillin-induced lysis in Escherichia coli. *J. Bacteriol. 191*, 5094−5107.
(42) Gelfand, M. S., and Rodionov, D. A. (2008) Comparative genomics and functional annotation of bacterial transporters. *Phys. Life Rev. 5*, 22−49.
(43) Thierry-Mieg, N. (2006) A new pooling strategy for high-throughput screening: the Shifted Transversal Design. *BMC Bioinf. 7*, 28.
(44) Jin, F., Hazbun, T., Michaud, G. A., Salcius, M., Predki, P. F., Fields, S., and Huang, J. (2006) A pooling-deconvolution strategy for biological network elucidation. *Nat. Methods 3*, 183−189.
(45) Greving, M., Cheng, X., Reindl, W., Bowen, B., Deng, K., Louie, K., Nyman, M., Cohen, J., Singh, A., Simmons, B., Adams, P., Siuzdak, G., and Northen, T. (2012) Acoustic deposition with NIMS as a high-throughput enzyme activity assay. *Anal. Bioanal. Chem. 403*, 707−711.
(46) Sakamoto, T., and Bryant, D. A. (1998) Growth at low temperature causes nitrogen limitation in the cyanobacterium Synechococcus sp. PCC 7002. *Arch. Microbiol. 169*, 10−19.
(47) Baran, R., Kochi, H., Saito, N., Suematsu, M., Soga, T., Nishioka, T., Robert, M., and Tomita, M. (2006) MathDAMP: a package for differential analysis of metabolite profiles. *BMC Bioinf. 7*, 530.
(48) Oh, J., Fung, E., Price, M. N., Dehal, P. S., Davis, R. W., Giaever, G., Nislow, C., Arkin, A. P., and Deutschbauer, A. (2010) A universal TagModule collection for parallel genetic analysis of microorganisms. *Nucleic Acids Res. 38*, e146.
(49) Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., Denisova, E. A., Baden-Tillson, H., Zaveri, J., Stockwell, T. B., Brownley, A., Thomas, D. W., Algire, M. A., Merryman, C., Young, L., Noskov, V. N., Glass, J. I., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2008) Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science 319*, 1215−1220.
(50) Meinnel, T., Schmitt, E., Mechulam, Y., and Blanquet, S. (1992) Structural and biochemical characterization of the Escherichia coli argE gene product. *J. Bacteriol. 174*, 2323−2331.