Coordination of CcpA and CodY Regulators in *Staphylococcus aureus* USA300 Strains

Saugat Poudel,[a] Ying Hefner,[a] Richard Szubin,[a] Anand Sastry,[a] Ye Gao,[a,b] Victor Nizet,[c,d] Bernhard O. Palsson[a,c]

[a]Department of Bioengineering, University of California San Diego, San Diego, California, USA
[b]Department of Biological Sciences, University of California San Diego, San Diego, California, USA
[c]Collaborative to Halt Antibiotic-Resistant Microbes (CHARM), Department of Pediatrics, University of California San Diego, San Diego, California, USA
[d]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, San Diego, California, USA

ABSTRACT The complex cross talk between metabolism and gene regulatory networks makes it difficult to untangle individual constituents and study their precise roles and interactions. To address this issue, we modularized the transcriptional regulatory network (TRN) of the *Staphylococcus aureus* USA300 strain by applying independent component analysis (ICA) to 385 RNA sequencing samples. We then combined the modular TRN model with a metabolic model to study the regulation of carbon and amino acid metabolism. Our analysis showed that regulation of central carbon metabolism by CcpA and amino acid biosynthesis by CodY are closely coordinated. In general, *S. aureus* increases the expression of CodY-regulated genes in the presence of preferred carbon sources such as glucose. This transcriptional coordination was corroborated by metabolic model simulations that also showed increased amino acid biosynthesis in the presence of glucose. Further, we found that CodY and CcpA cooperatively regulate the expression of ribosome hibernation-promoting factor, thus linking metabolic cues with translation. In line with this hypothesis, expression of CodY-regulated genes is tightly correlated with expression of genes encoding ribosomal proteins. Together, we propose a coarse-grained model where expression of *S. aureus* genes encoding enzymes that control carbon flux and nitrogen flux through the system is coregulated with expression of translation machinery to modularly control protein synthesis. While this work focuses on three key regulators, the full TRN model we present contains 76 total independently modulated sets of genes, each with the potential to uncover other complex regulatory structures and interactions.

IMPORTANCE *Staphylococcus aureus* is a versatile pathogen with an expanding antibiotic resistance profile. The biology underlying its clinical success emerges from an interplay of many systems such as metabolism and gene regulatory networks. This work brings together models for these two systems to establish fundamental principles governing the regulation of *S. aureus* central metabolism and protein synthesis. Studies of these fundamental biological principles are often confined to model organisms such as *Escherichia coli*. However, expanding these models to pathogens can provide a framework from which complex and clinically important phenotypes such as virulence and antibiotic resistance can be better understood. Additionally, the expanded gene regulatory network model presented here can deconvolute the biology underlying other important phenotypes in this pathogen.

KEYWORDS network modeling, *Staphylococcus aureus*, gene regulation, metabolism

Metabolism plays an integral role in infection and antimicrobial resistance (AMR) in the leading human bacterial pathogen *Staphylococcus aureus*. Metabolic requirements specific to infection, intracellular persistence, biofilm formation, and colonization are rapidly being uncovered (1–6). Furthermore, the central role of metabolism in

AMR and persistence is also coming into view, adding to the complexity of known AMR mechanisms (7–9). The complex metabolic circuits and responses underlying these phenomena are nevertheless difficult to unravel. Even relatively well-understood systems such as *S. aureus* central carbon metabolism can be difficult to fully map, as they are layered with multiple levels of gene regulation, posttranslational and biochemical controls, and unexpected molecular interactions (1, 10–12). Some of these complexities can be captured by genome-scale metabolic models (GEMs) that allow rapid query of metabolic complexities through simulations of metabolic flux states, knockout experiments, multistrain metabolic comparisons, and calculation of metabolic characteristics (13, 14). Alternatively, coarse-grained modeling of metabolism attempts to peer beyond the detailed complexity and discover the general principles governing the biological systems of interest (15–17). In the present work, we took guidance from a coarse-grained model proposed in *Escherichia coli* coupled with genome-scale analyses of *S. aureus* transcriptional regulation and metabolism to uncover a similar staphylococcal system that balances resource allocation between carbon and nitrogen metabolism (15, 17).

Biological trade-offs represent an optimization frontier, where the cell must strike a balance between its multiple objectives and their limitations (15, 18). Signatures of these balancing acts can be found in transcriptomes and become apparent when their architecture is viewed at the systems level (19). We previously described one such trade-off and its transcriptional imprint using independent data sets from Gram-negative *E. coli* and Gram-positive *S. aureus*, in which a balance was observed between genes regulated by stress-associated sigma factors and growth-associated translation machinery (20, 21). Here, we expand significantly beyond those observations to describe a trade-off between carbon and nitrogen metabolism in strains of the globally disseminated, hypervirulent *S. aureus* USA300 lineage.

We first greatly expanded on our previous model of transcriptional regulation in USA300 strains to incorporate all publicly available RNA sequencing data from the Sequence Reads Archive (SRA) (21). Models were then generated by applying independent component analysis (ICA), which calculates independently modulated sets of genes (iModulons) and their activities present in the input RNA sequencing samples. iModulons represent sources of signals in the expression data, with transcriptional regulators being the most common source. Our model showed that the activities of two global metabolic regulators, CcpA and CodY, which play critical roles in central carbon and nitrogen metabolism, respectively, are negatively correlated against one another. This negative correlation pointed to a condition-specific reallocation of resources toward different metabolic subsystems. GEMs fitted with metabolomics data confirmed the inferences made from the transcriptomic data. Furthermore, GEMs revealed specific metabolic interfaces where coordination of metabolism by the two regulators is required for optimal biomass production, including glutamate dehydrogenase and the folate cycle. Placing genes from CodY and CcpA-associated iModulons onto the metabolic map demonstrated that they did not share any metabolic reactions but coregulated expression of a gene encoding ribosome hibernation factor. In light of these observations, we propose a model whereby CcpA and CodY coordinate gene expression for carbon metabolism, nitrogen metabolism, and translation, thus modularly controlling protein production at specific stages.

RESULTS

Expanding the USA300 iModulons using RNA sequencing data from the SRA database. Our previous work outlined 29 iModulons for USA300 strains that were generated from 108 in-house RNA sequencing data (21). To expand the previous iModulons coverage of the TRN, we took advantage of the rapidly growing, publicly available *S. aureus* RNA sequencing samples (see Fig. S1 in the supplemental material). We queried the Sequence Reads Archive (SRA) for all available USA300-specific RNA sequencing data and combined it with 64 newly generated samples. Of the 576 sequencing samples available, 385 passed the stringent quality control/quality assurance (QC/QA) pipeline and were
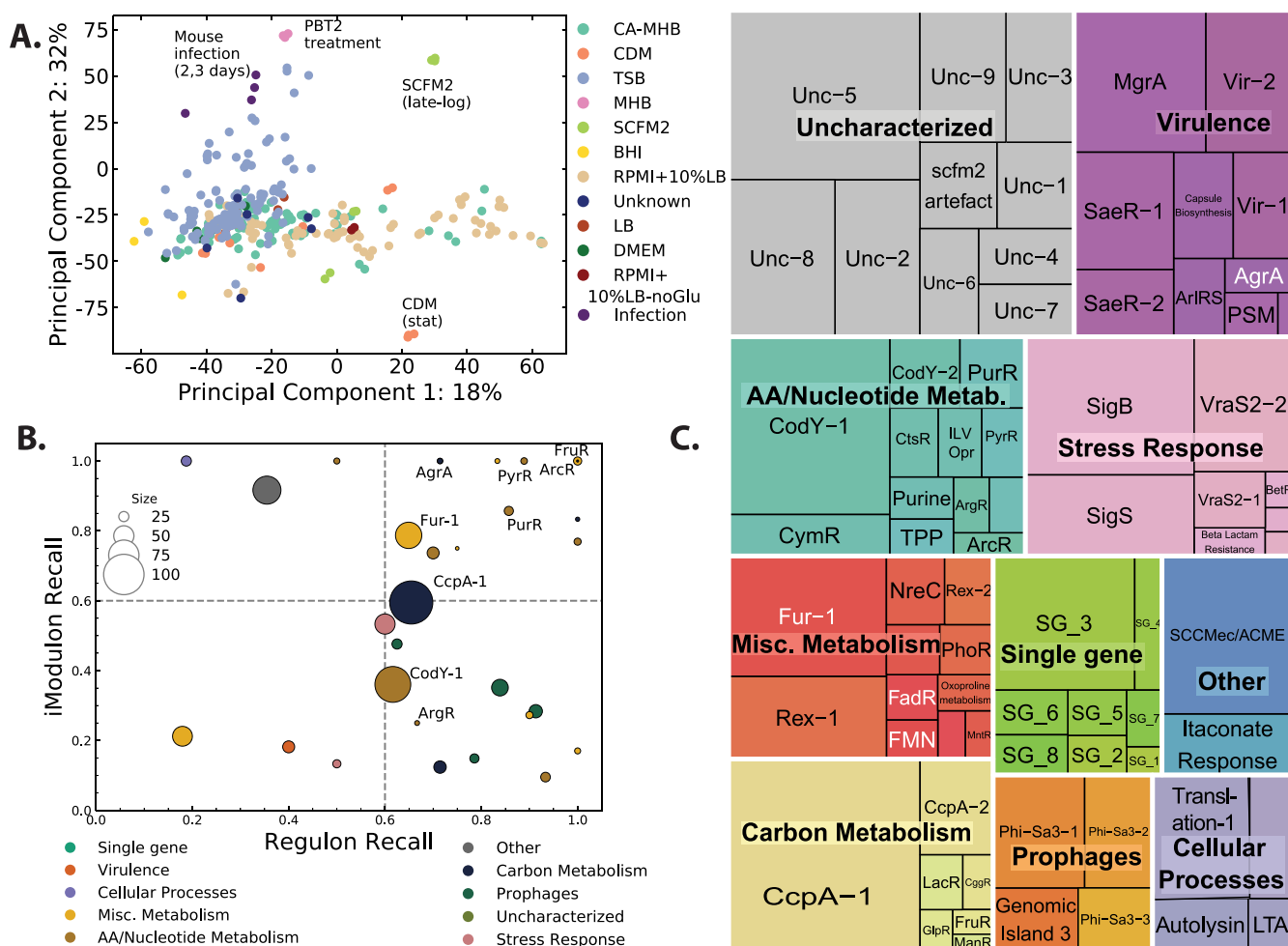
FIG 1 Updated iModulons for USA300 strains. (A) Principal-component analysis (PCA) of 385 RNA sequencing samples from diverse growth conditions that were used to generate the expanded USA300 iModulons. The log TPM of each sample was normalized to project-specific control conditions to reduce signal from batch effect before projecting them onto the principal components. (B) Association between iModulons and previously published regulons. iModulon recall represents the fraction of iModulon genes that were present in the regulon, while regulon recall represents the fraction of regulon genes that were present in the iModulon. iModulons that have high iModulon recall and high Regulon recall (e.g., PyrR, PurR) have very high overlap in gene content. (C) TreeMap of iModulon names, sizes (gene content), and types in the current model after manual curation. The size of the boxes correspond to the number of genes in the iModulon. stat, stationary; exp, exponential; TSB, tryptic soy broth; SCFM2, synthetic cystic fibrosis medium 2; CDM, chemically defined medium; TPM, transcripts per million.

therefore incorporated into the new model (see Materials and Methods). The final set of samples contained data from at least 7 different USA300 isolates, 4 growth phases (exponential, stationary, biofilm, and infection), and 10 base media (Fig. S2).

Before applying ICA, we normalized the log-transformed transcripts per million (log-TPM) data to a project-specific control condition. This reduced batch-specific variation in the data and reduced the presence of iModulons not associated with biological signals. Principal-component analysis (PCA) of the log-TPM data showed that normalized samples tended to cluster with media types and growth phases rather than by data source (Fig. 1a). For example, data from *S. aureus* grown to late-log phase in SCFM2 (synthetic cystic fibrosis sputum medium 2) and to stationary phase in chemically defined medium (CDM) did not cluster together, despite being from the same bioproject.

Application of ICA to this normalized expression data resulted in 76 independent components, and genes with high absolute weightings within each component were assigned to a corresponding iModulon. These enriched iModulon genes were then compared with existing literature of predicted regulons in *S. aureus*. Those iModulons that had significant overlap with other predicted regulons were named after the

associated regulator (Fig. 1b). Last, some iModulons with no known regulators but associated with other biological processes (e.g., prophages, translation) were manually curated. In total, we labeled 60 of the 76 iModulons with either a regulator or a biological process (Fig. 1c). The remaining uncharacterized iModulons represent signals in the *S. aureus* transcriptome with currently unknown origins, thus providing a roadmap to discovery of missing parts of the known TRN. In addition to the structure of each iModulon, the activities of each of the 76 iModulons in the 385 input samples were also calculated. The activity represents the role of each iModulon (and the associated regulator, if known) in shaping the transcriptome in the given sample. Higher iModulon activity represents a higher expression level of genes with positive weightings in the iModulon and lower expression of genes with negatively weighted genes (20).

CcpA and CodY iModulon activities highlight balance of carbon and nitrogen metabolism. Cumulatively, the 70 iModulons captured ~70% of the variance in the input transcriptomic data. The CodY-1, CcpA-1 (henceforth referred to as simply CodY and CcpA iModulons, respectively), and Translation iModulons explained the most variation in the data (Fig. 2a). CcpA is the catabolite repressor protein in *Firmicutes* that represses genes involved in alternate carbon utilization, as well as other central carbon metabolic pathways, such as the tricarboxylic acid (TCA) cycle, in the presence of high concentrations of glucose. CodY, on the other hand, globally represses the genes required for amino acid biosynthesis in response to high branched-chain amino acid (BCAA) or GTP concentrations. Lastly, the Translation iModulon almost entirely consists of ribosomal genes (e.g., *rplK*, *rplA*, etc.) and genes involved in translation, such as *infA* and *fusA*, which encode translation initiation factor IF-1 and elongation factor G, respectively (Fig. 2b). This iModulon has been enriched in almost all bacteria and archaea for which iModulons have been calculated (20, 22–25).

Interestingly, the activities of these three iModulons were highly correlated across all samples (Fig. 2c) and formed a large cluster, along with other metabolic iModulons (Fig. S3). Along with CodY, CcpA, and Translation iModulons, activities of ILVopr (iModulon containing the operon with isoleucine, leucine, and valine biosynthesis genes), MntR, LacR, and PurR iModulons were also highly correlated. The correlation of CcpA with LacR simply reflects the catabolite repression of lactose utilization genes by the regulator CcpA. Similarly, the ILV operon is regulated globally by CodY and locally by leucine attenuator (26). This multilayer regulation likely explains why this operon formed its own iModulon whose activity was closely correlated with CodY. The MntR iModulon contains genes required for manganese uptake, and its coordinated activity with CcpA confirms the association of manganese concentration with glycolytic flux (27).

The correlated activity of CcpA and CodY iModulons suggested that *S. aureus* carefully coordinates its central carbon and nitrogen metabolism (Fig. 2d). Close examination of the activities of these two iModulons showed a biphasic relationship. In conditions with preferred carbon sources, and therefore low CcpA iModulon activity, CodY activity generally increased. This effect was observed when glucose was added to both a complex medium (cation-adjusted Mueller-Hinton broth, or CA-MHB) and to a defined medium (chemically defined medium, or CDM1). Other conditions without explicitly controlled glucose levels that showed low CcpA activity still had concomitant high CodY activity, suggesting that this effect was not glucose specific. In conditions with already low CodY activity, however, removal of glucose (RPMI⁻ glucose; substituted with maltose) did not lead to further change in CodY activity, creating the second phase of the trade-off plane.

On the other hand, an increase in CodY iModulon activity did not necessarily lead to decrease in CcpA activity (Fig. 2d, red markers). Samples from several projects where the *codY* gene was interrupted showed minimal effect on CcpA iModulon activity. These samples fell well outside the CcpA-CodY trade-off line (Fig. 2d, gray dashed lines). Similar effects can also be observed in samples treated with subinhibitory concentrations of mupirocin. Mupirocin activates the stringent response in *S. aureus*, which leads to conversion of GTP to ppGpp and subsequent derepression of the CodY regulon (28). As change in CcpA activity leads to change in CodY activity but not necessarily vice versa, these data suggest that CcpA works "upstream" of CodY.
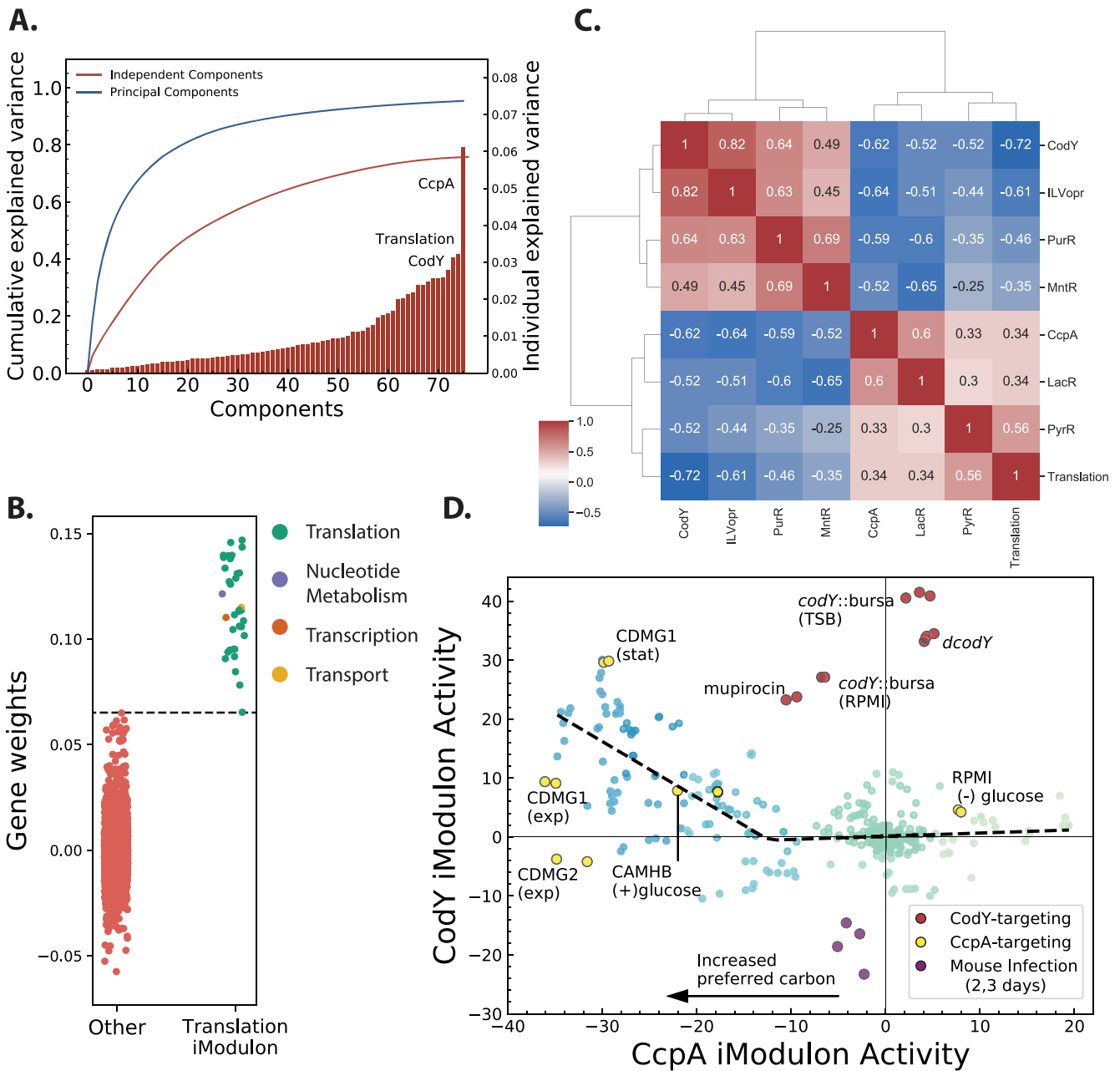
FIG 2 Coordination of metabolic iModulons in USA300 strains. (A) Explained variance of each of the iModulons; CcpA, Translation, and CodY iModulons explain the most variance in the transcriptome data. (B) Weightings of all the genes in the component associated with the Translation iModulon. Genes enriched in the iModulon (right column) are almost entirely part of the translation COG category. (C) Clustermap based on Pearson correlation coefficient of iModulon activities highlights potential coordination between various regulators. (D) Activity of CcpA and CodY iModulons across all USA300 samples shows biphasic association with close correlation in low CcpA activity conditions (blue markers) and little correlation otherwise (green markers). Large changes in CodY activity induced by genetic disruption of the *codY* gene or by mupirocin treatment does not alter CcpA activity (red markers). However, decrease in CcpA activity leads to increase in CodY activity (yellow markers). This asymmetric relationship suggests that CcpA works upstream of CodY. stat, stationary; exp, exponential; TSB, tryptic soy broth.

Metabolic modeling supports coupling of CcpA and CodY activities. To independently confirm the metabolic interaction between CodY and CcpA, we used a previously published USA300 strain-specific genome-scale metabolic model (GEM) (29). GEMs are curated mathematical models of an organism's metabolic network that can be used to simulate, study, and design the metabolic pathways using a wide range of constraints-based reconstruction and analysis (COBRA) tools (14, 30).

One such method, parsimonious flux balance analysis (pFBA), can be used to

calculate the metabolic flux state that optimizes a phenotype while minimizing total metabolic flux in a given condition (13, 31). pFBA thus represents a parsimonious use of the metabolic proteome. Here, we used pFBA to determine the metabolic flux states that maximize *S. aureus* biomass production given the measured uptake and secretion rates of various amino acids and sugars in chemically defined medium (CDM) and CDM with glucose (CDMG) (10). In agreement with increased CodY iModulon activity in CDMG, total flux through reactions catalyzed by enzymes that are encoded in CodY iModulon genes ("CodY reactions" for short) increased from 3.9 mmol/gram dry weight (gDW)/h to 5.5 mmol/gDW/h in the presence of glucose (Fig. 3a). The total sum of fluxes through CcpA reactions in our pFBA simulations did not change drastically between CDM and CDMG (Fig. S4). Many CcpA fluxes simply reversed directions (e.g., reactions involved in the tricarboxylic acid cycle), some of which are described in detail in the next section.

pFBA, however, gives an exact optimal solution and therefore does not account for variations or errors in input uptake data. We addressed this issue by sampling the CDM- and CDMG-specific models, which give a distribution of feasible fluxes in each of the respective conditions. We then mapped the condition-specific flux distributions to various amino acid biosynthetic pathways. For simple interpretation, we excluded amino acids that serve as intermediates for biosynthesis of other amino acids (e.g., glutamine, glutamate, and serine) and included only those amino acids for which unique biosynthetic pathways could be defined (see Materials and Methods). Confirming pFBA analysis, 5 out of the 6 amino acid biosynthetic pathways had increased flux in CDMG compared to CDM (Fig. 3b). The results of these two TRN-agnostic metabolic modeling methods are in agreement with our observation that CodY iModulon activity increases in the presence of glucose.

Transcriptional coordination of CcpA and CodY is likely due to flux coupling at metabolic interfaces. CcpA and CodY iModulons contained 110 and 86 genes, respectively. Most of these genes are involved in central carbon and amino acid metabolism. Despite the large iModulon sizes and close metabolic proximity of the regulated genes, the two iModulons did not share any genes encoding metabolic enzymes. The correlation in iModulon activities, however, suggested that CcpA reactions and CodY reactions must be coordinated at a metabolic level. Using the USA300 GEM, we looked for this coordination at the metabolite interface between CcpA and CodY reactions, i.e., metabolites that are involved in both CcpA and CodY reactions. We found these metabolic interfaces by systematically identifying all metabolites in USA300 GEM that can be found in both CodY and CcpA reactions. After taking out "nonspecific" metabolites and cofactors (e.g., ATP, H_2O, NADH, etc.), we were left with 22 metabolites at the interface (Table S1). While some of these metabolites, such as pyruvate, glutamate, and oxaloacetate, are expected, as they play a crucial role in both carbon and nitrogen metabolism, other metabolites, such as *N*-succinyl-2-L-amino-6-oxoheptanedioate and tetrahydrofolate (THF), are less understood in the context of this trade-off. To further understand how a change in simulated flux through CcpA and CodY reactions in CDM and CDMG altered these key metabolic interfaces, we mapped the pFBA solution fluxes from each media to the reactions around two of these interfaces, glutamate and methyl-THF.

The glutamate-alpha-ketoglutarate (αkg) link is a closely studied interface in *S. aureus* that connects amino acid and central carbon metabolism (5, 10). The main enzyme at the interface, glutamate dehydrogenase (GLUDy), reversibly interconverts αkg and glutamate and is encoded by the *gudB* gene, a constituent of the CcpA iModulon. However, this interconversion also acts as an amine group donor or acceptor to 3 CcpA reactions and 8 CodY reactions (Table S1). In glucose-free CDM, pFBA solution was consistent with a previous observation showing proline is converted to αkg via glutamate and eventually fuels gluconeogenesis (10). However, in CDMG, the flux through GLUDy changes direction and catalyzes conversion of αkg to glutamate instead (Fig. 3C). This makes up ~98% of total flux that consumes αkg. The glutamate, in turn, acts as an amine group donor for biosynthesis of various amino acids and accounts for ~80% of total flux generating αkg in CDMG (Fig. 3D). pFBA solution of
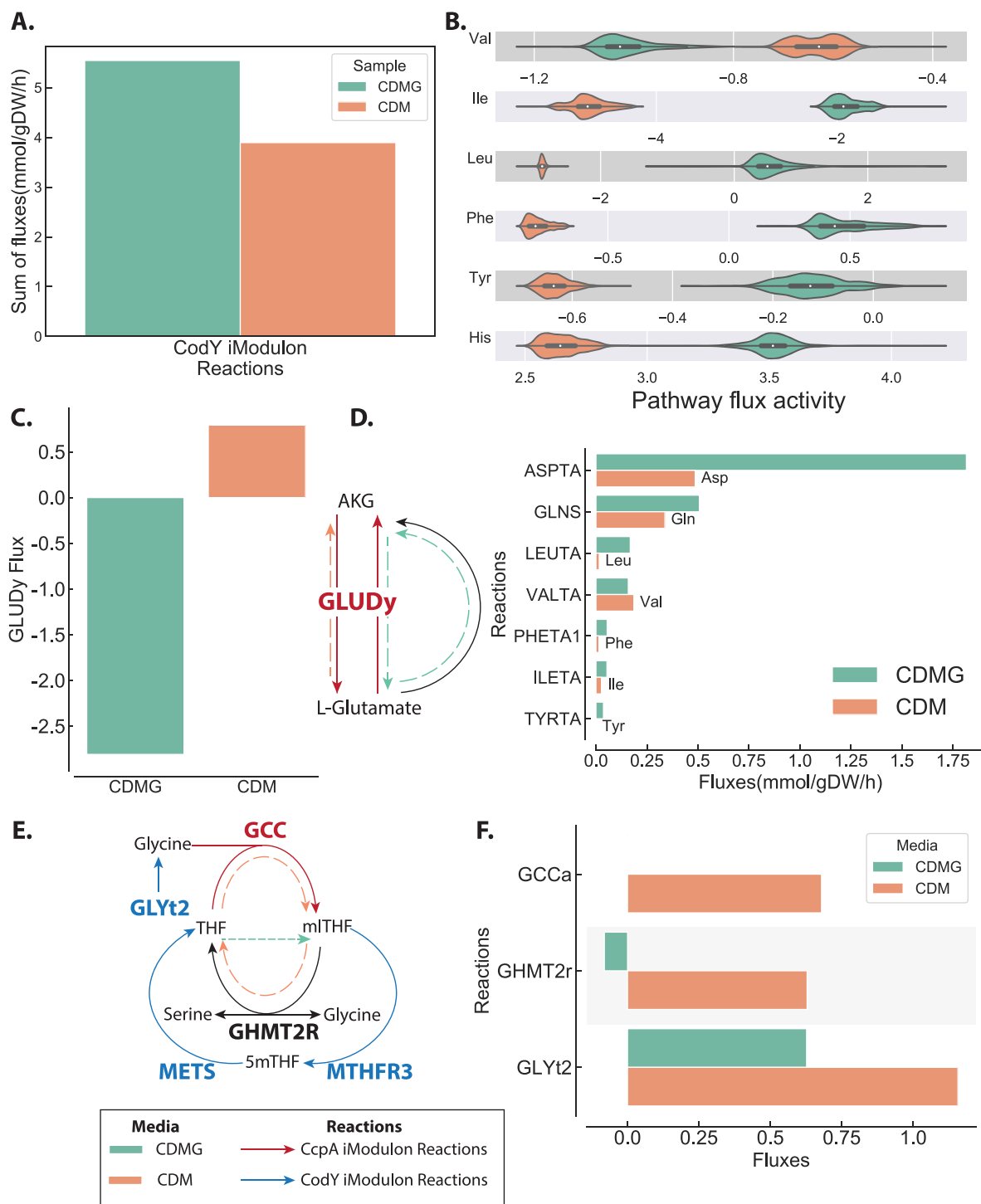
FIG 3 Metabolic modeling shows coordination of CcpA and CodY fluxes. (A) The solution of space of the models with CDM- and CDMG-specific metabolic constraints was sampled. The sum of sampled fluxes through CodY and CcpA reaction shows increased flux through CodY in CDMG. (B) Sampled fluxes through several amino acid biosynthetic pathways also show increased flux in CDMG. Definition of the biosynthetic pathways is found in Materials and Methods. (C) Flux through GLUDy reaction changes direction when glucose is added; alpha-ketoglutarate (AKG) is converted to L-glutamate in CDMG and vice versa in CDM. L-Glutamate is then utilized in biosynthesis of other amino acids in CDMG regenerating AKG. (D) Amino acids generated by accepting amine groups from L-glutamate, coupling the CcpA reaction GLUDy with amino acid-generating CodY reactions. (E) Metabolic map of folate cycle where CcpA and CodY regulated metabolism intersect. (F) Simulated flux through reactions in folate cycle in CDM and CDMG shows alternate metabolic reactions used in each medium. In CDM, reactions that consume glycine (GCC and GHMT2r) to interconvert between THF and methyl-THF (mlTHF). In CDMG, glycine is instead generated by GHMT2R running in reverse.
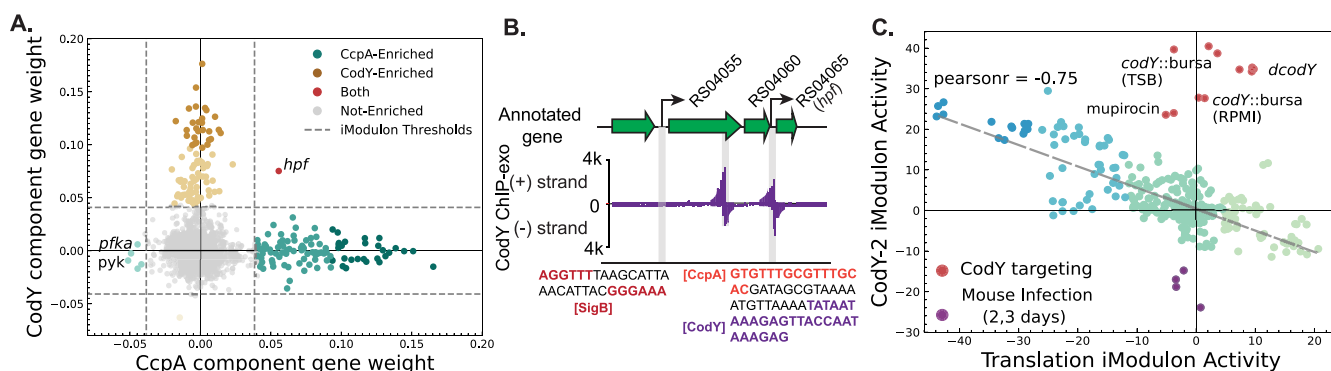
FIG 4 Coordination of translation with metabolism. (A) Gene weights in CcpA and CodY iModulon-associated components shows that only the *hpf* gene is a part of both iModulons. (B) Upstream region of hpf gene with its two alternative transcription start sites. Two CodY binding sites were detected by ChIP-exo (purple bars). The previously recognized SigB(red) and CodY(purple) binding sites and newly proposed CcpA (orange) binding site are highlighted. (C) The negative correlation between CodY and Translation iModulon activity suggests coordination of metabolism and translation in *S. aureus*. As with CcpA and CodY association, samples where CodY activity is altered by genetic disruption of *codY* or by treatment with mupirocin (in red) break this coordination. Other samples are colored in green-blue palette based on Translation iModulon activity.

this interface therefore shows that in the absence of glucose, the GLUDy reaction converts glutamate to αkg to fuel gluconeogenesis, but in the presence of glucose, it converts αkg to glutamate to fuel amino acid biosynthesis.

The folate cycle represents another metabolic interface of CcpA and CodY reactions. The folate cycle is required for one-carbon metabolism, nucleotide biosynthesis, and amino acid metabolism, and the pathway leading up to the cycle is the target of sulfonamide-class antibiotics (32). The folate cycle consists of 2 CodY reactions, MTHFR3 and methionine synthase (METS), and one CcpA reaction, GCCabc (glycine cleavage complex) (Fig. 3E). In CDM, THF is converted into 5,10-methylenetetrahydrofolate (mlTHF) by the GCCabc reaction, which cleaves glycine in the process (Fig. 3F). THF is then regenerated from mlTHF by the GHMT2r reaction, which also consumes glycine and generates serine. This consumption of glycine in the folate cycle by the CcpA reaction is coupled with increased transport of glycine by CodY-regulated GLYt2. However, in CDMG, where CcpA iModulon activity is low, there is no flux through the CcpA reaction, GCCabc. Instead, GHMT2r runs in "reverse" to convert THF from mlTHF, consuming serine and generating glycine instead. Together, combining iModulon structure with metabolic simulation demonstrates how *S. aureus* coordinates flux through CcpA and CodY iModulon reactions at these key metabolic interfaces, despite not sharing any genes at the regulatory level.

Expression levels of translation-associated genes are responsive to CcpA and CodY activities. While CcpA and CodY iModulons do not share any metabolic genes, *hpf*, which encodes ribosomal hibernation-promoting factors (HPFs), is a gene found in both iModulons (Fig. 4a). HPF is a small peptide that dimerizes 70S ribosomal subunits to form inactive 100S subunits (33, 34). It plays an important role in stress response and nutrition limitation and protects ribosomal pools from degradation (35–37). Previous studies in *S. aureus* have shown that SigB and CodY regulate *hpf* expression in response to heat and nutritional stress (35). iModulon structure confirms the role of the CodY and suggests an additional layer of control by CcpA.

Chromatin immunoprecipitation coupled with exonuclease treatment (ChIP-exo) data from our previous work found two CodY binding sites in the regulatory region of the *hpf* gene (Fig. 4b) (38). To confirm the role of CcpA in *hpf* expression, we searched for the catabolite repressor element motif (WTGNNARCGNWWWCAW) from *Bacillus subtilis* in the same region (39). A matching motif was found in the region between the two CodY binding peaks (false-discovery rate [*q*] = 0.00905). This architecture, with two CodY binding sites flanking the CcpA binding site, is also found in the regulatory region of the *B. subtilis* BCAA operon, where both regulators contribute to the expression of the operon genes (40). The signal from expression data and the presence of
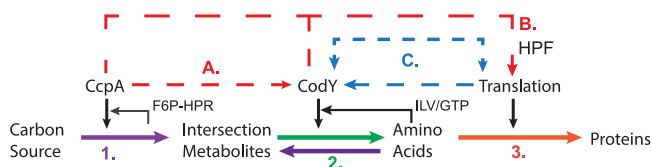
FIG 5 Proposed coarse-grained model of protein biosynthesis regulation in *S. aureus*. The solid lines represent the parts of the protein synthesis pathway controlled by CcpA (purple) and CodY (green). The dashed lines represent new proposed roles of these regulators in coordinating carbon and nitrogen metabolism (A) and linking metabolic gene expression with expression of translation-associated proteins (B, C). At the regulatory level, CcpA and CodY activities are coordinated at high glucose conditions. Similarly, CodY, but not CcpA, activity is correlated with the activity of Translation iModulon. CcpA and CodY also feed forward to regulate the expression of the *hpf* gene, which encodes a hibernation factor that can alter the fraction of active ribosomes in the cell. Through this mechanism, *S. aureus* may be changing the concentration of the active ribosome in response to metabolic cues.

binding motifs suggests that CcpA regulates *hpf* along with the previously identified regulators, CodY and SigB.

In addition to coordinated regulation of the translation-associated *hpf* gene, CodY activity was also strongly correlated with Translation iModulon activity. In contrast, CcpA and Translation iModulon activities showed little correlation between them (Fig. S5). Similar to CcpA and CodY activity correlation, *codY* knockout and stringent response activation by mupirocin also disrupted correlation with Translation iModulon (Fig. 4c). This also suggested that the signal controlling Translation iModulon gene expression also works upstream of CodY, as interruption of CodY had little effect on Translation iModulon activity. While the coordination of the two iModulon activities is apparent, we were unable to further interrogate the nature of this relationship since the signal behind the Translation iModulon is yet to be identified.

DISCUSSION

Based on the data presented here, we propose a coarse-grained model of transcriptional regulation of metabolism involved in protein synthesis in *S. aureus* USA300 strains (Fig. 5). It is motivated by the model of proteome coordination in *E. coli* and extends its principles to nonmodel pathogenic organisms (15). The coarse-grained model simplifies metabolism underlying protein synthesis into three steps, (i) the generation of precursors from the carbon source, (ii) biosynthesis of amino acids from precursors or direct transport from the medium, and (iii) synthesis of peptides from amino acids via translation. The generation of precursors from carbon sources is largely regulated by CcpA (Fig. 5, purple arrow). CcpA represses alternate carbon sources (including amino acids such as proline, glutamine, and aspartate) in the presence of preferred carbon (such as glucose) and regulates other key aspects of central metabolism, such as gluconeogenesis and the TCA cycle, that are necessary to generate various precursors (1, 10, 41–43). The precursors in our model are represented by the metabolites at the CcpA-CodY interface derived from the USA300 GEM (see Table S1 in the supplemental material). These precursors are then converted to amino acids via CodY-regulated gene products (green arrow) and polymerized by ribosomes into proteins (light blue arrow) (38, 43, 44).

Our analysis suggests that *S. aureus* USA300 strains coordinate their CcpA and CodY activity to regulate carbon and nitrogen flow through the system (dashed orange arrow). Metabolic modeling in CDMG shows increased flux through amino acid biosynthetic reactions compared to CDM. The results of this TRN-agnostic metabolic model agree with the increased CodY activity in CDMG and other glucose-containing media. Additionally, we also found a feed-forward regulation whereby CcpA and CodY control the expression of the gene encoding the HPF protein, which sequesters ribosomes into inactive 100S forms, suggesting a mechanism by which translation is coordinated with the metabolic state of the cell (dashed red arrows) (33, 35).

Last, the activity of the Translation iModulon is also closely correlated with CodY activity,

which may act as an additional layer of coordination between metabolism and translation (dashed blue arrows). However, we have yet to identify the signal or regulator controlling the Translation iModulon activity, and therefore, the source of this concomitant change in expression along with CodY is unclear. rRNA expression is regulated by ppGpp during the stringent response, which can be activated by mupirocin treatment (28, 45). We therefore expected mupirocin to also have an effect on Translation iModulon activity, but we found that while CodY activity increased in response to mupirocin as expected, there was minimal change in Translation activity (Fig. 4c). This suggests that stringent response, at least when induced by mupirocin treatment, does not play a major role in the expression of Translation iModulon genes.

Despite close coordination of metabolic flux at different interfaces between CcpA and CodY reactions, it is still not clear how CcpA and CodY activities are coordinated. In *E. coli*, Kochanowski et al. have observed similar coordination between anabolic and catabolic fractions of metabolism (46). The authors attributed active regulation by Crp and passive changes in metabolic fluxes in response to change in metabolite concentrations as the source of the coordination. Other nonmetabolic constraints could also play a role in this coordination. The agreement between our metabolic model and the iModulon activities only arises after the metabolic model is constrained using experimentally measured amino acid uptake rates. This suggests that other nonmetabolic factors may be leading to differential amino acid uptake rates in glucose media and subsequently leading to the change in CodY activity. In this sense, the problem may be similar to the inability of unconstrained metabolic models to predict overflow metabolism, which has been postulated to result from other nonmetabolic constraints, such as proteome and membrane real estate allocation (47, 48). The root of this relationship, whether metabolic or nonmetabolic in nature, is yet to be determined.

The analysis of the coarse-grained model of metabolic gene regulation presented here was enabled by a computable model of TRN. iModulons enable us to query the TRN at multiple scales, giving insights into TRN from single-gene membership level to global coordination of regulators. By modularizing the TRN, our analysis enabled us to unravel complex regulatory and metabolic interactions to understand regulation of the central metabolism one regulator at a time. This modularization can also be used to continually expand on the presented model. For example, our previous works have shown that Translation iModulon activity in *E. coli* and *S. aureus* is closely correlated with stress-associated alternate sigma factors (20, 21). This points to a possible entry point for incorporation of general stress response with metabolism and protein synthesis. Similarly, we have also found that both PyrR and PurR activity is correlated with CodY and CcpA, which may provide insights into the regulation of nucleotide biosynthesis in response to carbon or nitrogen availability. While we mainly focused on 3 iModulons, CcpA, CodY, and Translation, the current model contains 76 total iModulons, each of them rich with information about transcriptional regulation and physiology of *S. aureus*. We thus provide a conceptual framework for overall coordination of metabolism in *S. aureus* and approaches to systematically expand and detail the model proposed.

MATERIALS AND METHODS

Strains and growth conditions. The *S. aureus* USA300 isolate, LAC, and its derivative, JE2, were used to collect the new RNA sequencing data in this study. The complete description and condition for each of the samples can be found in Table S2 in the supplemental material. For RNA sequencing from knockout samples, isolates from the Nebraska Transposon Mutant Library were utilized (49). Unless specified otherwise, samples were grown in duplicates in 20 mL of respective media until they reached the optical density at 600 nm (OD_{600}) of 0.5. Three milliliters of culture was harvested and immediately mixed with 6 mL of Qiagen RNAprotect bacteria reagent and incubated at room temperature for 5 min. The supernatant was decanted after the samples were centrifuged for 10 min and 17,500 rpm. The remaining cell pellets were stored at −80°C until they were prepared for RNA extraction.

RNA extraction and sequencing. Total RNA was isolated from the cell pellet in the Qiagen RNeasy minikit columns by following vendor procedures. An on-column DNase treatment was performed for 30 min at room temperature. The rRNA was removed using RiboRid protocol, as described before (50). RNA was quantified using a NanoDrop and quality assessed by running an RNA nanochip on a bioanalyzer (Agilent, CA). A Swift RNA library kit was used following the manufacturer's protocol to create sequencing libraries.

Preparing RNA sequencing data for iModulon calculation. The iModulons were calculated from publicly available RNA sequencing data from SRA and the newly collected data in this study using PyModulon python package (51). The steps used to calculate the iModulons described here were all completed using this package. All RNA sequencing data labeled with *S. aureus* taxonomic ID were downloaded and manually curated to obtain only the samples that were from USA300 isolates. Raw fastq files from curated samples were downloaded, trimmed with TrimGalore, and then aligned to the USA300 TCH1516 genome (GenBank accession nos. NC_010079, NC_012417, and NC_010063) using Bowtie2 (52). QC/QA stats were collected on each sample using MultiQC, and samples that did not pass the QC thresholds (e.g., low read depth, low correlation between replicates, missing metadata) were discarded (53). Transcripts per million (TPM) were calculated from the remaining high-quality RNA sequencing samples. TPM were log transformed and normalized to a control condition within the same bioproject.

Calculating iModulons from RNA sequencing data. Scipy's implementation of FastICA was applied to log-transformed and normalized TPM data to generate independent components (ICs) and their activities (54, 55). Unlike other decomposition methods, ICA requires the number of dimensions to be calculated as an input. Therefore, various models with different dimensionality were created, and the one that maximized regulatory iModulons and minimized single-gene iModulon was chosen (56). The iModulons were then automatically annotated if they overlapped significantly with a curated list of known or predicted regulons and genomic features (e.g., prophages, SCCMec, ACME, etc.) in *S. aureus*. Other iModulons, such as "Translation" or "Autolysin," were manually annotated, as all genes contained within the iModulons have a single function. The gene content of iModulon can be found in Table S2.

Genomic-scale modeling of *S. aureus* USA300 metabolism. The USA300-specific genome-scale model (GEM) iYS854 was used for all metabolic simulations in the paper. The exchange rates of amino acids, glucose, ammonium, and acetate were adjusted to constrain the model to CDM- or CDMG-specific conditions as described in detail before (29). Briefly, the uptake or secretion rate for each metabolite from Halsey et al. was normalized by growth rate to get the growth-adjusted solute uptake rate (10). The exchange rates were then constrained to $\pm15\%$ of uptake and exchange rate to account for variance in the data.

Once constrained, the model was then used to calculate the flux of each medium using pFBA as implemented in the COBRApy package (30, 31). To get the CodY iModulon-specific flux, genes in the CodY iModulon were first mapped to metabolic reactions using the gene product rule (GPR). The absolute values of fluxes from the pFBA solution for the CodY reactions were then summed to get the final CodY iModulon flux.

To calculate the valid amino acid biosynthesis pathway-specific flux distribution, the solution spaces of CDM- and CDMG-specific models were sampled 10,000 times using the artificial centering hit-and-run algorithm (57). Next, the reactions in each amino acid biosynthetic pathway were determined with the MinSpan algorithm (58). MinSpan calculates the set of shortest metabolic pathways that are linearly independent of one another and span the null space of the input model. Each independent pathway defines a mass-balanced set of reactions and therefore enables unbiased modularization of metabolism into biologically meaningful pathways. The sampled fluxes (v) can therefore be represented as linear weightings (α) of minspan pathways (P), where

$$v = P \times \alpha$$

The sampled fluxes were converted to pathway-specific weightings (pathway fluxes) using the minspan matrix. Pathways containing amino acid biosynthesis were manually curated, and only amino acid biosynthesis pathways that did not appear in multiple MinSpan pathways were used for analysis, as they can be easily interpreted and do not require analyzing linear combinations of multiple pathways.

Last, the interface metabolites were determined by comparing all metabolites that were involved in at least one CodY and one CcpA reaction. The common metabolites ADP, ATP, CO_2, coenzyme A, H_2O, hydrogen atom, sodium ion, NAD, NADH, NADP, NADPH, ammonium (NH4), and phosphate were excluded from this designation.

Motif enrichment. The 150 base pairs upstream of the *hpf* gene (USA300HOU_RS04065) were scanned for the CcpA motif (WTGNNARCGNWWWCAW) using find individual motif occurrence (FIMO) within the MEME suite (39, 59).

Data and code availability. All RNA sequencing data used in this work are available publicly at Sequence Reads Archive (SRA). The accession numbers for individual samples can be found in Table S2. The code used to create the model and generate all the figures in the paper can be found on GitHub (https://github.com/sapoudel/metabs-paper-code). An interactive version of the iModulon model is also available at iModulonDB (https://imodulondb.org/).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TEXT S1, DOCX file, 0.01 MB.
FIG S1, EPS file, 0.8 MB.
FIG S2, EPS file, 0.8 MB.
FIG S3, EPS file, 4.6 MB.
FIG S4, EPS file, 0.6 MB.
FIG S5, EPS file, 1.1 MB.
FIG S6, EPS file, 1 MB.

FIG S7, EPS file, 1.1 MB.
TABLE S1, CSV file, 0 MB.
TABLE S2, CSV file, 0.3 MB.

REFERENCES

1. Potter AD, Butrico CE, Ford CA, Curry JM, Trenary IA, Tummarakota SS, Hendrix AS, Young JD, Cassat JE. 2020. Host nutrient milieu drives an essential role for aspartate biosynthesis during invasive *Staphylococcus aureus* infection. Proc Natl Acad Sci U S A 117:12394–12401. https://doi.org/10.1073/pnas.1922211117.
2. Bosi E, Monk JM, Aziz RK, Fondi M, Nizet V, Palsson BØ. 2016. Comparative genome-scale modelling of *Staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity. Proc Natl Acad Sci U S A 113:E3801–E3809. https://doi.org/10.1073/pnas.1523199113.
3. Diep BA, Stone GG, Basuino L, Graber CJ, Miller A, des Etages S-A, Jones A, Palazzolo-Ballance AM, Perdreau-Remington F, Sensabaugh GF, DeLeo FR, Chambers HF. 2008. The arginine catabolic mobile element and staphylococcal chromosomal cassette *mec* linkage: convergence of virulence and resistance in the USA300 clone of methicillin-resistant *Staphylococcus aureus*. J Infect Dis 197:1523–1530. https://doi.org/10.1086/587907.
4. Lade H, Park JH, Chung SH, Kim IH, Kim JM, Joo HS, Kim JS. 2019. Biofilm formation by *Staphylococcus aureus* clinical isolates is differentially affected by glucose and sodium chloride supplemented culture media. J Clin Med Res 8:1853. https://doi.org/10.3390/jcm8111853.
5. DeMars Z, Bose JL. 2018. Redirection of metabolism in response to fatty acid kinase in *Staphylococcus aureus*. J Bacteriol 200:e00345-18. https://doi.org/10.1128/JB.00345-18.
6. Melter O, Radojevič B. 2010. Small colony variants of *Staphylococcus aureus*–review. Folia Microbiol (Praha) 55:548–558. https://doi.org/10.1007/s12223-010-0089-3.
7. Yang JH, Wright SN, Hamblin M, McCloskey D, Alcantar MA, Schrübbers L, Lopatkin AJ, Satish S, Nili A, Palsson BO, Walker GC, Collins JJ. 2019. A white-box machine learning approach for revealing antibiotic mechanisms of action. Cell 177:1649–1661.e9. https://doi.org/10.1016/j.cell.2019.04.016.
8. Lopatkin AJ, Stokes JM, Zheng EJ, Yang JH, Takahashi MK, You L, Collins JJ. 2019. Bacterial metabolic state more accurately predicts antibiotic lethality than growth rate. Nat Microbiol 4:2109–2117. https://doi.org/10.1038/s41564-019-0536-0.
9. Gaupp R, Lei S, Reed JM, Peisker H, Boyle-Vavra S, Bayer AS, Bischoff M, Herrmann M, Daum RS, Powers R, Somerville GA. 2015. *Staphylococcus aureus* metabolic adaptations during the transition from a daptomycin susceptibility phenotype to a daptomycin nonsusceptibility phenotype. Antimicrob Agents Chemother 59:4226–4238. https://doi.org/10.1128/AAC.00160-15.
10. Halsey CR, Lei S, Wax JK, Lehman MK, Nuxoll AS, Steinke L, Sadykov M, Powers R, Fey PD. 2017. Amino acid catabolism in *Staphylococcus aureus* and the function of carbon catabolite repression. mBio 8:e01434-16. https://doi.org/10.1128/mBio.01434-16.
11. Ding Y, Liu X, Chen F, Di H, Xu B, Zhou L, Deng X, Wu M, Yang C-G, Lan L. 2014. Metabolic sensor governing bacterial virulence in *Staphylococcus aureus*. Proc Natl Acad Sci U S A 111:E4981–E4990.
12. Thomas VC, Sadykov MR, Chaudhari SS, Jones J, Endres JL, Widhelm TJ, Ahn J-S, Jawa RS, Zimmerman MC, Bayles KW. 2014. A central role for carbon-overflow pathways in the modulation of bacterial cell death. PLoS Pathog 10:e1004205. https://doi.org/10.1371/journal.ppat.1004205.
13. Orth JD, Thiele I, Palsson BØ. 2010. What is flux balance analysis? Nat Biotechnol 28:245–248. https://doi.org/10.1038/nbt.1614.
14. Lewis NE, Nagarajan H, Palsson BO. 2012. Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. Nat Rev Microbiol 10:291–305. https://doi.org/10.1038/nrmicro2737.
15. You C, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, Wang Y-P, Lenz P, Yan D, Hwa T. 2013. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. Nature 500:301–306. https://doi.org/10.1038/nature12446.
16. Basan M, Honda T, Christodoulou D, Hörl M, Chang Y-F, Leoncini E, Mukherjee A, Okano H, Taylor BR, Silverman JM, Sanchez C, Williamson JR, Paulsson J, Hwa T, Sauer U. 2020. A universal trade-off between growth and lag in fluctuating environments. Nature 584:470–474. https://doi.org/10.1038/s41586-020-2505-4.
17. Erickson DW, Schink SJ, Patsalo V, Williamson JR, Gerland U, Hwa T. 2017. A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. Nature 551:119–123. https://doi.org/10.1038/nature24299.
18. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. 2012. Multidimensional optimality of microbial metabolism. Science 336:601–604. https://doi.org/10.1126/science.1216882.
19. Utrilla J, O'Brien EJ, Chen K, McCloskey D, Cheung J, Wang H, Armenta-Medina D, Feist AM, Palsson BO. 2016. Global rebalancing of cellular resources by pleiotropic point mutations illustrates a multi-scale mechanism of adaptive evolution. Cell Syst 2:260–271. https://doi.org/10.1016/j.cels.2016.04.003.
20. Sastry AV, Gao Y, Szubin R, Hefner Y, Xu S, Kim D, Choudhary KS, Yang L, King ZA, Palsson BO. 2019. The *Escherichia coli* transcriptome mostly consists of independently regulated modules. Nat Commun 10:5536. https://doi.org/10.1038/s41467-019-13483-w.
21. Poudel S, Tsunemoto H, Seif Y, Sastry AV, Szubin R, Xu S, Machado H, Olson CA, Anand A, Pogliano J, Nizet V, Palsson BO. 2020. Revealing 29 sets of independently modulated genes in *Staphylococcus aureus*, their regulators, and role in key physiological response. Proc Natl Acad Sci U S A 117:17228–17239. https://doi.org/10.1073/pnas.2008413117.
22. Rychel K, Sastry AV, Palsson BO. 2020. Machine learning uncovers independently regulated modules in the *Bacillus subtilis* transcriptome. Nat Commun 11:6338. https://doi.org/10.1038/s41467-020-20153-9.
23. Yoo R, Rychel K, Poudel S, Al-Bulushi T, Yuan Y, Chauhan S, Lamoureux C, Palsson BO, Sastry A. 2021. Machine learning of all *Mycobacterium tuberculosis* H37Rv RNA-seq data reveals a structured interplay between metabolism, stress response, and infection. mSphere 7:e00033-22. https://doi.org/10.1128/msphere.00033-22.
24. Chauhan SM, Poudel S, Rychel K, Lamoureux C, Yoo R, Al Bulushi T, Yuan Y, Palsson BO, Sastry AV. 2021. Machine learning uncovers a data-driven transcriptional regulatory network for the crenarchaeal thermoacidophile *Sulfolobus acidocaldarius*. Front Microbiol 12:753521. https://doi.org/10.3389/fmicb.2021.753521.
25. Yuan Y, et al. 2022. Pan-genomic analysis of transcriptional modules across *Salmonella Typhimurium* reveals the regulatory landscape of different strains. bioRxiv. https://doi.org/10.1101/2022.01.11.475931.
26. Kaiser JC, King AN, Grigg JC, Sheldon JR, Edgell DR, Murphy MEP, Brinsmade SR, Heinrichs DE. 2018. Repression of branched-chain amino acid synthesis in *Staphylococcus aureus* is mediated by isoleucine via CodY, and by a leucine-rich attenuator peptide. PLoS Genet 14:e1007159. https://doi.org/10.1371/journal.pgen.1007159.
27. Radin JN, Kelliher JL, Párraga Solórzano PK, Kehl-Fie TE. 2016. The two-component system ArlRS and alterations in metabolism enable *Staphylococcus aureus* to resist calprotectin-induced manganese starvation. PLoS Pathog 12:e1006040. https://doi.org/10.1371/journal.ppat.1006040.
28. Reiss S, Pané-Farré J, Fuchs S, François P, Liebeke M, Schrenzel J, Lindequist U, Lalk M, Wolz C, Hecker M, Engelmann S. 2012. Global analysis of the *Staphylococcus aureus* response to mupirocin. Antimicrob Agents Chemother 56:787–804. https://doi.org/10.1128/AAC.05363-11.
29. Seif Y, Monk JM, Mih N, Tsunemoto H, Poudel S, Zuniga C, Broddrick J, Zengler K, Palsson BO. 2019. A computational knowledge-base elucidates the response of *Staphylococcus aureus* to different media types. PLoS Comput Biol 15:e1006644. https://doi.org/10.1371/journal.pcbi.1006644.
30. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. 2013. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. BMC Syst Biol 7:74. https://doi.org/10.1186/1752-0509-7-74.
31. Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, Adkins JN, Schramm G, Purvine SO, Lopez-Ferrer D, Weitz KK, Eils R, König

R, Smith RD, Palsson BØ. 2010. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. Mol Syst Biol 6:390. https://doi.org/10.1038/msb.2010.47.

32. Proctor RA. 2008. Role of folate antagonists in the treatment of methicillin-resistant *Staphylococcus aureus* infection. Clin Infect Dis 46:584–593. https://doi.org/10.1086/525536.

33. Puri P, Eckhardt TH, Franken LE, Fusetti F, Stuart MCA, Boekema EJ, Kuipers OP, Kok J, Poolman B. 2014. *Lactococcus lactis* YfiA is necessary and sufficient for ribosome dimerization. Mol Microbiol 91:394–407. https://doi.org/10.1111/mmi.12468.

34. Gohara DW, Yap M-NF. 2018. Survival of the drowsiest: the hibernating 100S ribosome in bacterial stress management. Curr Genet 64:753–760. https://doi.org/10.1007/s00294-017-0796-2.

35. Basu A, Shields KE, Eickhoff CS, Hoft DF, Yap MN. 2018. Thermal and nutritional regulation of ribosome hibernation in *Staphylococcus aureus*. J Bacteriol 200:e00426-18. https://doi.org/10.1128/JB.00426-18.

36. Basu A, Yap M-NF. 2016. Ribosome hibernation factor promotes *Staphylococcal* survival and differentially represses translation. Nucleic Acids Res 44:4881–4893. https://doi.org/10.1093/nar/gkw180.

37. Lipońska A, Yap M-NF. 2021. Hibernation-promoting factor sequesters *Staphylococcus aureus* ribosomes to antagonize RNase R-mediated nucleolytic degradation. mBio 12:e00334-21. https://doi.org/10.1128/mBio.00334-21.

38. Gao Y, Poudel S, Seif Y, Shen Z, Palsson BO. 2021. Elucidating the CodY regulon in *Staphylococcus aureus* USA300 substrains. bioRxiv. https://doi.org/10.1101/2021.01.08.426013.

39. Schumacher MA, Sprehe M, Bartholomae M, Hillen W, Brennan RG. 2011. Structures of carbon catabolite protein A–(HPr-Ser46-P) bound to diverse catabolite response element sites reveal the basis for high-affinity binding to degenerate DNA operators. Nucleic Acids Res 39:2931–2942. https://doi.org/10.1093/nar/gkq1177.

40. Fujita Y, Satomura T, Tojo S, Hirooka K. 2014. CcpA-mediated catabolite activation of the *Bacillus subtilis ilv-leu* operon and its negation by either CodY- or TnrA-mediated negative regulation. J Bacteriol 196:3793–3806. https://doi.org/10.1128/JB.02055-14.

41. Nuxoll AS, Halouska SM, Sadykov MR, Hanke ML, Bayles KW, Kielian T, Powers R, Fey PD. 2012. CcpA regulates arginine biosynthesis in *Staphylococcus aureus* through repression of proline catabolism. PLoS Pathog 8: e1003033. https://doi.org/10.1371/journal.ppat.1003033.

42. Seidl K, Müller S, François P, Kriebitzsch C, Schrenzel J, Engelmann S, Bischoff M, Berger-Bächi B. 2009. Effect of a glucose impulse on the CcpA regulon in *Staphylococcus aureus*. BMC Microbiol 9:95. https://doi.org/10.1186/1471-2180-9-95.

43. Sonenshein AL. 2007. Control of key metabolic intersections in *Bacillus subtilis*. Nat Rev Microbiol 5:917–927. https://doi.org/10.1038/nrmicro1772.

44. Waters NR, Samuels DJ, Behera RK, Livny J, Rhee KY, Sadykov MR, Brinsmade SR. 2016. A spectrum of CodY activities drives metabolic reorganization and virulence gene expression in *Staphylococcus aureus*. Mol Microbiol 101:495–514. https://doi.org/10.1111/mmi.13404.

45. Samarrai W, Liu DX, White A-M, Studamire B, Edelstein J, Srivastava A, Widom RL, Rudner R. 2011. Differential responses of *Bacillus subtilis* rRNA promoters to nutritional stress. J Bacteriol 193:723–733. https://doi.org/10.1128/JB.00708-10.

46. Kochanowski K, Okano H, Patsalo V, Williamson J, Sauer U, Hwa T. 2021. Global coordination of metabolic pathways in *Escherichia coli* by active and passive regulation. Mol Syst Biol 17:e10064. https://doi.org/10.15252/msb.202010064.

47. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T. 2015. Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. Nature 528:99–104. https://doi.org/10.1038/nature15765.

48. Szenk M, Dill KA, De Graff AMR. 2017. Why do fast-growing bacteria enter overflow metabolism? Testing the membrane real estate hypothesis. Cell Syst 5:95–104. https://doi.org/10.1016/j.cels.2017.06.005.

49. Fey PD, Endres JL, Yajjala VK, Widhelm TJ, Boissy RJ, Bose JL, Bayles KW. 2013. A genetic resource for rapid and comprehensive phenotype screening of nonessential *Staphylococcus aureus* genes. mBio 4:e00537-12. https://doi.org/10.1128/mBio.00537-12.

50. Choe D, Szubin R, Poudel S, Sastry A, Song Y, Lee Y, Cho S, Palsson B, Cho B-K. 2021. RiboRid: a low cost, advanced, and ultra-efficient method to remove ribosomal RNA for bacterial transcriptomics. PLoS Genet 17:e1009821. https://doi.org/10.1371/journal.pgen.1009821.

51. Sastry AV, Poudel S, Rychel K, Yoo R, Lamoureux CR, Chauhan S, Haiman ZB, Al Bulushi T, Seif Y, Palsson BO. 2021. Mining all publicly available expression data to compute dynamic microbial transcriptional regulatory networks. bioRxiv. https://doi.org/10.1101/2021.07.01.450581.

52. Highlander SK, Hultén KG, Qin X, Jiang H, Yerrapragada S, Mason EO, Shang Y, Williams TM, Fortunov RM, Liu Y, Igboeli O, Petrosino J, Tirumalai M, Uzman A, Fox GE, Cardenas AM, Muzny DM, Hemphill L, Ding Y, Dugan S, Blyth PR, Buhay CJ, Dinh HH, Hawes AC, Holder M, Kovar CL, Lee SL, Liu W, Nazareth LV, Wang Q, Zhou J, Kaplan SL, Weinstock GM. 2007. Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. BMC Microbiol 7:99. https://doi.org/10.1186/1471-2180-7-99.

53. Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

54. Hyvärinen A, Oja E. 2000. Independent component analysis: algorithms and applications. Neural Netw 13:411–430. https://doi.org/10.1016/S0893-6080(00)00026-5.

55. Jones E, Oliphant T, Peterson P. 2001. SciPy: open source scientific tools for python.

56. McConn JL, Lamoureux CR, Poudel S, Palsson BO, Sastry AV. 2021. Optimal dimensionality selection for independent component analysis of transcriptomic data. BMC Bioinformatics 22:584. https://doi.org/10.1186/s12859-021-04497-7.

57. Megchelenbrink W, Huynen M, Marchiori E. 2014. optGpSampler: an improved tool for uniformly sampling the solution-space of genome-scale metabolic networks. PLoS One 9:e86587. https://doi.org/10.1371/journal.pone.0086587.

58. Bordbar A, Nagarajan H, Lewis NE, Latif H, Ebrahim A, Federowicz S, Schellenberger J, Palsson BO. 2014. Minimal metabolic pathway structure is consistent with associated biomolecular interactions. Mol Syst Biol 10:737. https://doi.org/10.15252/msb.20145243.

59. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208. https://doi.org/10.1093/nar/gkp335.