Characterizing the Interplay of Rubisco and Nitrogenase Enzymes in Anaerobic-Photoheterotrophically Grown *Rhodopseudomonas palustris* CGA009 through a Genome-Scale Metabolic and Expression Model

Niaz Bahar Chowdhury,[a] Adil Alsiyabi,[a] Rajib Saha[a]

[a]Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

ABSTRACT *Rhodopseudomonas palustris* CGA009 is a Gram-negative purple nonsulfur bacterium that grows phototrophically by fixing carbon dioxide and nitrogen or chemotrophically by fixing or catabolizing a wide array of substrates, including lignin breakdown products for its carbon and fixing nitrogen for its nitrogen requirements. It can grow aerobically or anaerobically and can use light, inorganic, and organic compounds for energy production. Due to its ability to convert different carbon sources into useful products during anaerobic growth, this study reconstructed a metabolic and expression (ME) model of *R. palustris* to investigate its anaerobic-photoheterotrophic growth. Unlike metabolic (M) models, ME models include transcription and translation reactions along with macromolecules synthesis and couple these reactions with growth rate. This unique feature of the ME model led to nonlinear growth curve predictions, which matched closely with experimental growth rate data. At the theoretical maximum growth rate, the ME model suggested a diminishing rate of carbon fixation and predicted malate dehydrogenase and glycerol-3 phosphate dehydrogenase as alternate electron sinks. Moreover, the ME model also identified ferredoxin as a key regulator in distributing electrons between major redox balancing pathways. Because ME models include the turnover rate for each metabolic reaction, it was used to successfully capture experimentally observed temperature regulation of different nitrogenases. Overall, these unique features of the ME model demonstrated the influence of nitrogenases and rubiscos on *R. palustris* growth and predicted a key regulator in distributing electrons between major redox balancing pathways, thus establishing a platform for *in silico* investigation of *R. palustris* metabolism from a multiomics perspective.

IMPORTANCE In this work, we reconstructed the first ME model for a purple nonsulfur bacterium (PNSB). Using the ME model, different aspects of *R. palustris* metabolism were examined. First, the ME model was used to analyze how reducing power entering the *R. palustris* cell through organic carbon sources gets partitioned into biomass, carbon dioxide fixation, and nitrogen fixation. Furthermore, the ME model predicted electron flux through ferredoxin as a major bottleneck in distributing electrons to nitrogenase enzymes. Next, the ME model characterized different nitrogenase enzymes and successfully recapitulated experimentally observed temperature regulations of those enzymes. Identifying the bottleneck responsible for transferring an electron to nitrogenase enzymes and recapitulating the temperature regulation of different nitrogenase enzymes can have profound implications in metabolic engineering, such as hydrogen production from *R. palustris*. Another interesting application of this ME model can be to take advantage of its redox balancing strategy to gain an understanding of the regulatory mechanism of biodegradable plastic production precursors, such as polyhydroxybutyrate (PHB).

*R*hodopseudomonas palustris CGA009 (*R. palustris*) is an alphaproteobacterium with the ability to grow in both metabolic niches, phototrophic, and chemotrophic. Furthermore, it can fix both carbon dioxide and nitrogen and can grow under both aerobic and anaerobic conditions by using light and organic or inorganic compounds to generate ATP (1, 2). Using these metabolic versatilities, *R. palustris* has emerged as a potential biotechnological platform for bioremediation (3–5), bioplastics production (6, 7), and bioelectricity generation (8, 9), wastewater treatment (10–12), and hydrogen production (13–17). Furthermore, *R. palustris* is the only known bacteria to encode all three known nitrogenase enzymes (2) besides *Azotobacter vinelandii* (18). *R. palustris* also encodes both form I and form II of rubisco. These unique features make *R. palustris* an ideal microorganism to be considered a biotechnological chassis for further metabolic engineering (7). However, the complex interplay among these exclusive features and other metabolic pathways of *R. palustris* is still not properly understood and a systems-level analysis is required to understand such interplay.

One widely accepted systems-level investigation tool is the genome-scale M model (19). Initial efforts of reconstructing M models of purple nonsulfur bacteria (PNSB) were limited to the specific metabolic pathways of interest, such as central carbon metabolism (20), and electron transport chain (21). However, those pathway-specific M models did not have a wider resolution to capture the overall metabolic landscape of PNSBs. To overcome that, comprehensive M models were reconstructed for PNSB strains, including *Rhodobacter sphaeroides* (22) and *R. palustris* (23). Recently we further refined the *R. palustris* M model by integrating the annotated metabolic pathways for *p*-coumarate and coniferyl alcohol and validated it by using the experimental data on gene essentiality and metabolic flux analysis for growth under different carbon sources (24). Although these M models were useful to study different metabolic features of PNSB, the inherent lack of quantitative characterization of macromolecular machinery synthesis (MMS) (25), also known as the ribosome, may lead to incorrect predictions of biological scenarios, such as inaccurate reaction flux and multiple equivalent cellular phenotypic states (26, 27). These inaccuracies can lead to an erroneous understanding of the overall metabolic and regulatory features of an organism and can negatively impact the design-build-test-learn cycle for metabolic engineering application.

One of the ways to overcome incorrect predictions of biological scenarios is to use resource allocation-based models, such as whole-cell modeling (28), resource balance analysis (29), and the metabolic and expression (ME) modeling approach (30). Recently the ME modeling approach was used to answer fundamental biological questions regarding the genotype-phenotype relationship. ME model is a resource allocation-based model that includes not only the stoichiometric metabolic reactions but also quantitative MMS information (31). As input, ME models require the conditions of a steady-state environment and can then output everything which can be predicted through an M model, such as predictions for maximum growth rate, substrate uptake, by-product secretion, and metabolic fluxes. In addition, the ME model can exclusively predict gene expression levels, protein levels, and nonlinear growth curves (31). A ME model utilizes a growth optimization function along with coupling constraints that tie flux to transcriptional and translational reactions in the model. These constraints are functions of the growth rate. By including these constraints, ME models set limitations on fluxes based on transcription as well as translation reactions. Thus far, ME models were developed only for a few organisms. These models were used to accurately predict the cellular composition and gene expression of *Thermotoga maritima* (32), fermentation profile of *Clostridium ljungdahlii* (33), overflow metabolism of *Saccharomyces cerevisiae* (34), and multiscale phenotype, enzyme abundance, and acid stress of *Escherichia coli* (35–37). An ME model for *R. palustris* can also be very useful in answering fundamental biological questions, such as growth profiling, isozyme

expression prediction, regulation of electron distribution between competing metabolic modules, and temperature regulation of different enzymes.

In this work, the first-ever ME model was reconstructed for *R. palustris*. The ME model was able to satisfactorily recapitulate the experimental transcriptomics and proteomics observations from the literature (38). Acetate, succinate, butyrate, and *p*-coumarate were then used as carbon sources to characterize the growth profile of *R. palustris*, which closely matched with experimental growth rate data. In addition, it predicted a diminishing rate of carbon fixation at the theoretical maximum growth rate and consequently predicted malate dehydrogenase and glycerol-3 phosphate dehydrogenase as alternate electron sinks. Furthermore, the ME model identified ferredoxin as a key regulator in distributing electrons between major redox balancing pathways, such as carbon and nitrogen fixation. Later, the modeling framework was able to capture experimentally observed temperature regulation of different nitrogenase enzymes by varying turnover rate of nitrogen fixation reactions. Overall, this modeling framework demonstrated a bottom-up systems-biology approach that can be used to predict and analyze the cellular physiology of *R. palustris*, thereby providing an opportunity to generate experimentally testable hypotheses.

RESULTS AND DISCUSSION

Metabolic and expression model development. To reconstruct the ME model, our previously reconstructed M model of *R. palustris, iRpa940* (24), was used as a template. To reconstruct the ME model, gene-protein-reaction (GPR) relationships for all the reactions were manually curated from the complete genome sequence of *R. palustris* (2). Transcription and translation reactions were added for reactions for which GPR relationships were available. These transcription and translation reactions in the ME model did not quantify the amount of mRNA and proteins. However, these predict the rate of production for each of the transcription and translation reactions, which can be considered the proxy of the amount of mRNA and proteins. Reactions for which GPR associations are not available, it was assumed that an average bacterial enzyme with 31.09 kDa molecular weight (39) catalyzed each reaction. Overall, the ME model contains 1397 reactions, 1483 unique metabolites, and 826 genes. A comparison between these matrices between the M model and ME model can be found in Table S1. From the comparison, the M model could only predict the growth rate where more nutrient results in more growth, also known as the exponential growth phase. However, the ME model could predict the exponential growth phase like the M model. In addition, it could predict the stationary growth phase where the growth rate was limited by the nutrient uptake and resource allocation. Furthermore, the M model is in good agreement with the ME model for the exponential growth phase. A similar observation of growth rate prediction between the M model and the ME model was also made in the previous ME model of *E. coli* (35). Fig. 1 shows the workflow of the ME model reconstruction.

In *R. palustris*, there are two different rubisco enzymes, form I and form II. Form I rubisco (L_8S_8) is comprised of eight active large subunits (L_8) and active eight small subunits (S_8) (40) and is encoded by two genes, *rpa1559* and *rpa1560* (2). On the other hand, form II rubisco (L_2) is comprised of two active large subunits both encoded by *rpa4641* (2). Of the two forms of rubisco, form I has a higher molecular weight compared to form II (41, 42) and, therefore, requires more carbon investment to synthesize. Because rubisco is one of the most abundant enzymes in nature (43), the kinetics of this enzyme has been determined for multiple organisms (40, 44). For different rubisco enzymes, it was shown that although form I has higher molecular weight and more carbon investment cost, form II has a higher catalytic turnover rate (k_{cat}) per active site compared to form I (40). Evolutionary selection has played a major role in this counterintuitive observation (45, 46). Early in the earth's history, the concentration of carbon dioxide was higher in the atmosphere and as a result form, II rubisco evolved with a lower selectivity and higher k_{cat} for carbon dioxide (40). With increasing amounts of oxygen in the earth's atmosphere, form I evolved with a much higher selectivity for carbon dioxide but with a lower k_{cat} (40). Because k_{cat} values for *R. palustris* are not available, to account for these evolutionary selections, the k_{cat} values were set to 3.7 s^{-1} active site^{-1} (form I) and 6.6 s^{-1} active site^{-1} (form II) based on the
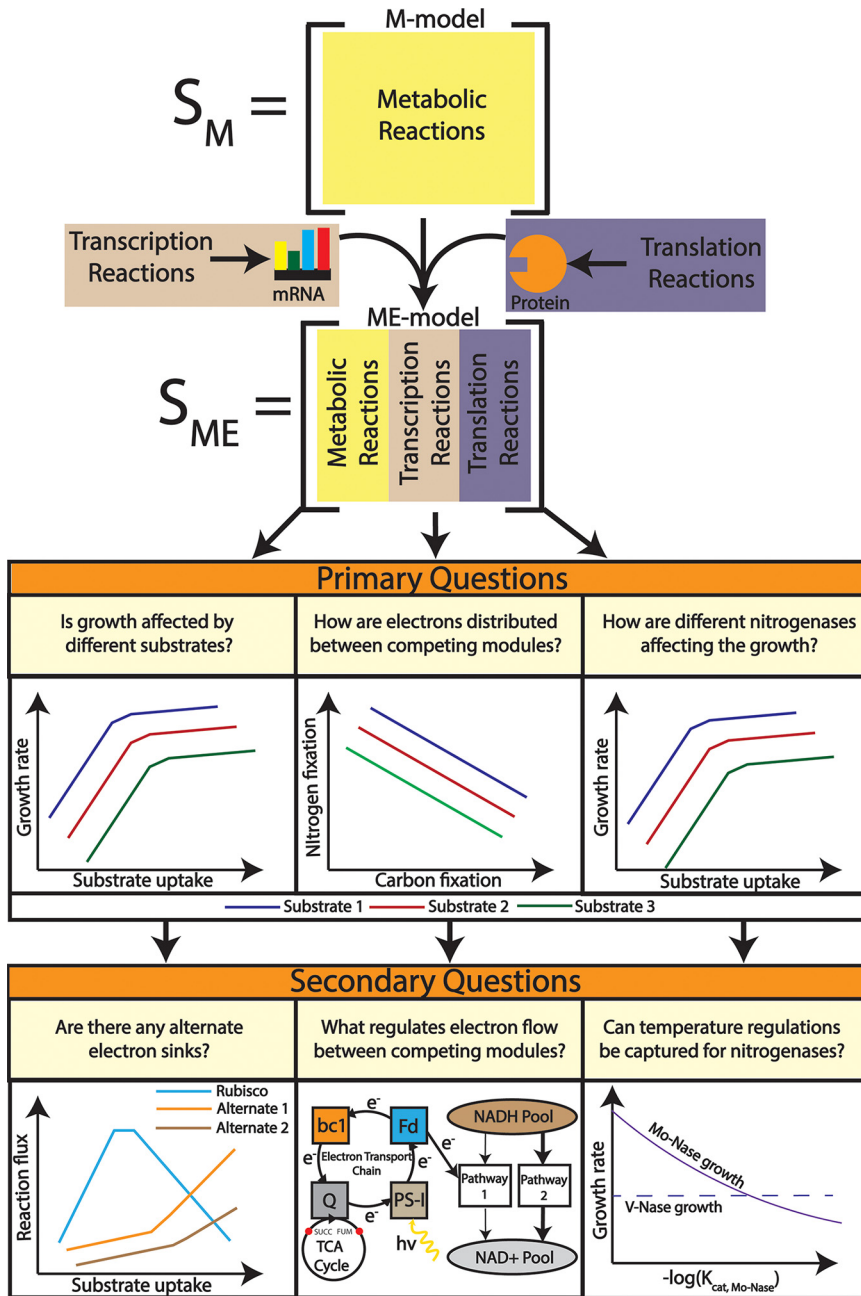
FIG 1 Workflow followed to reconstruct the ME model from a previously published M model of *R. palustris*. Transcription and translation reactions were added on top of the metabolic reactions to come up with an ME modeling framework. The ME modeling framework was used to characterize growth rate profiling, competing for metabolic modules, and nitrogenase enzyme activity. From these characterizations, inferences regarding alternate redox balancing, ferredoxin regulation, and temperature regulation of nitrogenase enzymes were gathered.

measurements from other phylogenetically close (47) PNSB strains (*Rhodobacter capsulatus* (44) and *Rhodospirillum rubrum* (40), respectively).

R. palustris has three different nitrogenase enzymes, Mo-Nase, V-Nase, and Fe-Nase. Each enzyme is encoded by a series of genes (2) (Mo-Nase by *rpa4602* to *rpa4633*, V-Nase by *rpa1370* to *rpa1380*, and Fe-Nase by *rpa1435* to *rpa1439*). Unlike rubisco, k_{cat} values of different nitrogenase are not available for *R. palustris* or any other PNSBs. Therefore, the calculated surface accessible surface area (SASA) of each nitrogenase enzyme was used to normalize the mean k_{cat} value, as discussed elsewhere (35) (see

Materials and Methods). These normalized k_{cat} values were used to define three independent nitrogen fixation reactions.

Both above-mentioned enzymes, nitrogenase and rubisco, play a pertinent role in maintaining the cellular redox balance during the photoheterotrophic growth of *R. palustris* by regenerating oxidized cofactors (48). When the ME model was used to simulate the photoheterotrophic growth of *R. palustris*, among three different nitrogenase enzymes, it predicted the expression of Mo-Nase only, which is consistent with the literature (49, 50). For the same photoheterotrophic growth conditions, between two different forms of rubisco enzymes, the model predicted only the expression of form II rubisco. Although expression of only rubisco form II was expected based on its lower carbon cost and higher efficiency, literature evidence suggested coexpression of both forms of rubisco during the photoheterotrophic growth of *R. palustris* (51). The same work suggested that rubisco form I is responsible for providing cellular carbon and dominates under carbon dioxide limiting conditions, whereas rubisco form II balances the intracellular redox potential under carbon and electron abundant conditions (51). In addition, it was also found that the expression of the cbb*cbb* operons (responsible for coding both forms of rubisco) during phototrophic growth is highly dependent on the cellular carbon dioxide level (51). To incorporate these findings, a constraint was added to the ME model to coexpress both forms of rubisco based on the total carbon dioxide produced by *R. palustris* during photoheterotrophic growth (see Materials and Methods).

Model validation using experimental transcriptomics and proteomics data. To validate the prediction accuracy of the model, experimental transcriptomics and proteomics data were used to qualitatively verify whether the model can predict the direction of these experimental fold changes in different conditions. A previous study, which characterized the anaerobic growth of *R. palustris* by comparing the transcriptomic and proteomic profiles of cultures grown in the presence of *p*-coumarate and succinate as the sole carbon source, was used for the validation study (38). The study tested fold change of 4810 genes for *p*-coumarate catabolism considering succinate catabolism as the baseline condition using Affymetrix GeneChip Operating Software Version 1.4 and Cyber-T program. The transcriptomic analysis, which is based on relative transcript abundance, resulted in 369 differentially expressed genes, among which 61 were metabolic genes. Similarly, the proteomic analysis resulted in 341 differentially expressed proteins, among which 67 can act as enzymes. In both transcriptomics and proteomics data sets, nonmetabolic genes/proteins have functions, such as signaling, chromosomal replication, and circadian rhythm (Table S2).

To generate both gene and protein expression information for the same two conditions of the above-mentioned study (38), the ME model was simulated for two points where total rubisco flux was maximal for the *p*-coumarate and succinate uptake, respectively. It was previously reported (48) that carbon fixation is required to maintain redox balance in *R. palustris*. Therefore, a higher growth rate is associated with higher reduced cofactor production, leading to higher rates of carbon fixation. As a result, the decreasing carbon fixation flux with increasing growth (Fig. 2) is a theoretical feature predicted by the ME model. All the experimentally observed and differentially expressed genes and proteins are available in the model. However, for reactions catalyzed by multiple isozymes, the ME model only predicted the most efficient isozyme based on the k_{cat} and molecular weight. As a result, out of these experimentally observed 61 genes and 67 enzymes with metabolic functions, we could only compare 25 genes and 37 enzymes with the experimental data. To eliminate the possibility of alternate solution space, we used the minimization of the total sum of fluxes as an outer objective (52). Hence, the variability of each of the reactions was very tight and the fluxes of the reactions were distributed in a way such that the total amount of enzyme requirements are minimized. Similar tight variability of reaction fluxes was also noticed in the previously published ME model for *E. coli* (35).

As part of the transcriptomics data validation, out of 25 genes, the ME model was able to predict correct gene expression fold change for 23 genes. The model could not
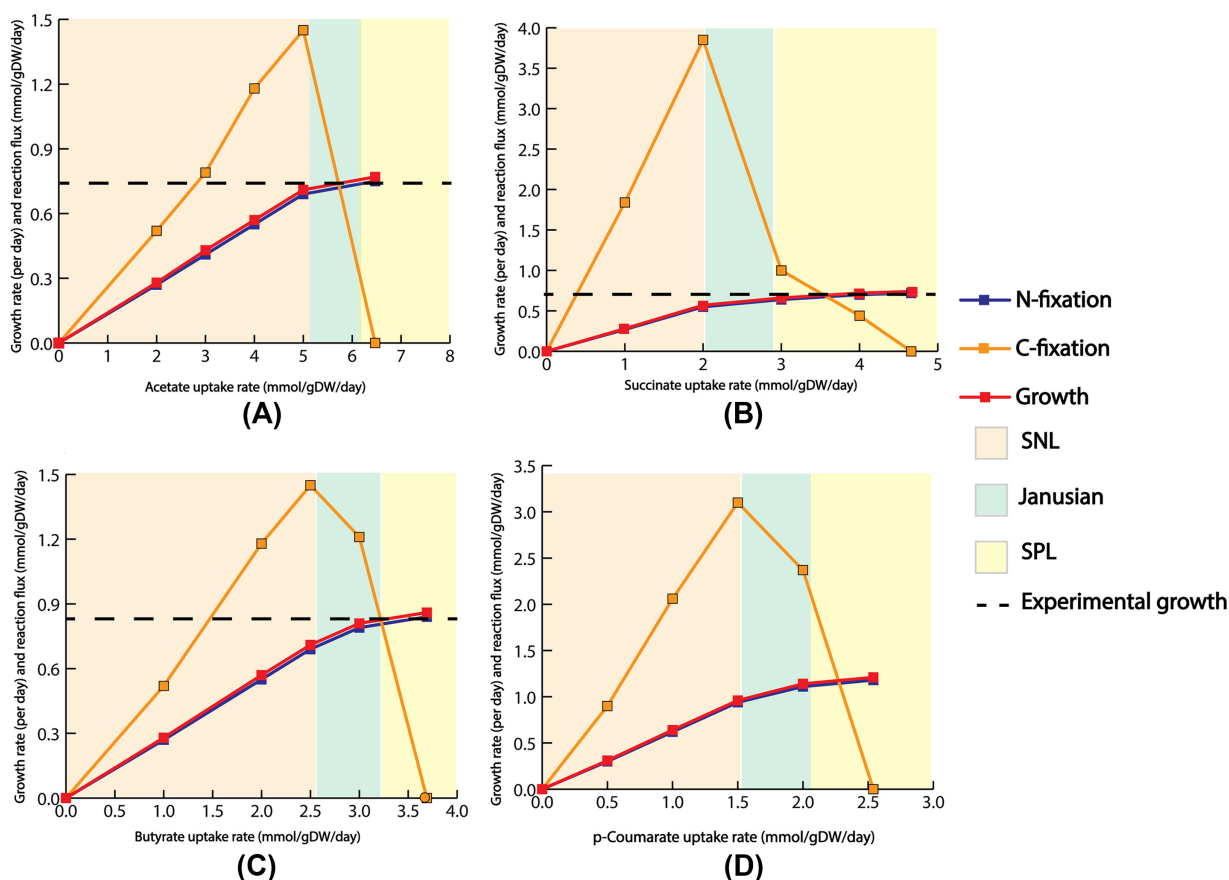
FIG 2 Strictly nutrient-limited (SNL), Janusian, and strictly proteome-limited (SPL) regions for (A) acetate, (B) succinate, (C) butyrate, and (D) *p*-coumarate. The growth rate with respect to different substrate uptakes follows a nonlinear pattern. Flux through nitrogen fixation reaction also followed a similar pattern to the growth rate. Carbon fixation reached a peak in the Janusian region and then diminished in the theoretical maximal growth.

predict the downward fold change of 3-oxoacyl-acyl carrier protein reductase (*rpa3304*) and the 50S ribosomal protein (*rpa0918*). *rpa3304* is one of the genes to convert malonyl-CoA to biotin (53). Biotin is a part of the *R. palustris* cell membrane (54). In Fig. 2, it can be seen that p-coumarate supported a higher growth rate than succinate. Thus, the ME model predicted an upward fold change of *rpa3304* for *p*-coumarate catabolism compared to succinate catabolism. The composition of biotin in the cell membrane may be different in different conditions. However, in the ME model, only protein and nucleotide compositions change with different conditions, while those of cell wall components remain constant (30). This may have caused the mismatch. For incorrect fold change prediction of 50S ribosomal protein, missing reactions, the lack of regulatory mechanisms, and inaccurate k_{cat} data may have played a role. Besides these factors, uncertainty arising from the different parameters in the ME model, such as the mass of rRNA per ribosome (m_{rr}), the molecular weight of average amino acid (m_{aa}), and the fraction of RNA that was rRNA (f_{rRNA}), which were used from the previously published model ME model of *E. coli* (35) and may have contributed to these incorrect predictions.

For proteomics data validation, out of 37 enzymes, the ME model was able to correctly predict the fold change for 25 enzymes. The ME model could not correctly predict the downward fold change of 12 different enzymes (Table S2). The number of incorrect predictions is higher in the validation of the proteomics data compared to the transcriptomics data. One of the inherent weaknesses of the ME model is unless additional constraints are imposed, the amount of protein is proportional to the amount of transcript, which may not be true in some instances (55). These incorrectly

TABLE 1 Normalized growth rate for different substrate uptakes

Substrate	Experimentally observed growth rate (day^{-1})	The growth rate from the ME model (day^{-1})	Substrate uptake for expt growth rate from model (mmol. gDW^{-1} day^{-1})
Succinate	0.70	0.74	4.66
Acetate	0.74	0.77	6.47
Butyrate	0.82	0.86	3.69
p-coumarate	–[a]	1.21	2.54

[a]- indicates that no experimentally observed growth rate data is available for *p*-coumarate uptake.

predicted enzymes are mainly associated with purine and pyrimidine metabolism, fatty acid metabolism, and lipopolysaccharide metabolism. These pathways are closely associated with the *R. palustris* biomass growth. Because *p*-coumarate supports more growth than succinate, the ME model allocated more proteins for these pathways to sustain biomass growth. There may be unannotated alternate metabolic pathways with less enzyme investment for producing purine, pyrimidine, fatty acid, and lipopolysaccharide when *p*-coumarate is utilized as the carbon source, thus causing these discrepancies. Because the ME model maximizes the biomass growth rate, such incorrect prediction can be considered an inherent weakness of the ME model.

Overall, despite these incorrect fold change predictions, the ME model was able to satisfactorily recapitulate the subgroup of transcriptomics and proteomics observations which matched with the experimental data with 92% and 68% accuracy, respectively (see Materials and Methods for accuracy calculation). The details of experimental and model predictions are in Table S2.

Growth rate versus substrate uptake and alternate redox balancing strategies. Upon the validation with available gene expression and protein abundance data, the model was used to examine how growth, carbon fixation, and nitrogen fixation rates varied with different substrate uptake rates. The goal of this analysis was to investigate how reducing power entering the cell through organic carbon sources gets partitioned into biomass, carbon dioxide fixation, and nitrogen fixation. To perform the analysis, acetate, succinate, butyrate, and *p*-coumarate were used as the substrates. Previous studies have shown that photoheterotrophic growth of *R. palustris* on acetate, succinate, and butyrate is associated with increased cellular redox stress based on the oxidation state of different substrates (56). Hence, these substrates were chosen as they cover a wide range of oxidation states. Here, succinate (+0.5) and acetate (0) have higher oxidation states compared to *R. palustris'* biomass (−0.13) (49), whereas butyrate (−1) and *p*-coumarate (−0.22) have lower oxidation states (49). These oxidation states are based on neutral molecules.

In the ME model, the growth rate is a nonlinear function of substrate uptake rate and eventually reaches a theoretical maximum growth rate (Fig. 2). This behavior is consistent with known microbial empirical growth models, such as Monod growth kinetics (57) and microbial slow growth kinetics (58). Previous work has suggested three distinct growth regions as a function of substrate uptake rate. Strictly nutrient-limited (SNL), Janusian, and strictly proteome-limited (SPL) (35). Growth in the SNL region depends heavily on nutrient uptake and adding more nutrients results in more growth. In this region, the relationship between growth rate and substrate uptake is similar to the prediction made from M models. Contrary to the SNL region, growth in the SPL region (also known as the nutrient excess condition) is limited by the physiological constraint of protein production and catalysis. Janusian growth is the region where a transition from SNL to SPL takes place. A recent experimental study (49) had characterized the growth of wild-type (WT) *R. palustris* for acetate, succinate, and butyrate, respectively, under nitrogen-fixing conditions. Table 1 compares experimentally observed growth rates and those predicted by the model. The growth rate and order predicted by the ME model for succinate, acetate, and butyrate closely followed the experimental growth rate and order. Compared to other substrates, the ME model predicted a significantly higher growth rate on *p*-coumarate. One of our

previous works (7), which experimentally examined different strategies for PHB production under non-nitrogen fixing conditions, also showed significantly higher growth in p-coumarate compared to butyrate and acetate. It was previously reported (7) that, *p*-coumarate consumption led to more ATP production compared to acetate, succinate, and butyrate and thus was able to support more growth.

Theoretical growth rates predicted by the ME model were slightly higher than the experimental growth rates for all tested substrates (6% for succinate, 5% for butyrate, and 4% for acetate). The cell has many more layers of physiological regulations, such as transcriptional regulation, allosteric regulation, posttranscriptional regulations, and single nucleotide polymorphisms, which were not captured in the ME modeling framework. Overall, a growth rate comparison between the ME model prediction and experimental study reveals that similar to *E. coli* (35), optimum resource allocation dictates metabolic activities for *R. palustris*. Table S3 has the theoretical maximum growth rates for different amounts of substrate uptakes.

After characterizing the growth rate with different substrate uptakes, the ME model was used to characterize nitrogen and carbon fixation rates as a function of substrate uptake. For nitrogen fixation, the reaction's activity followed a similar trajectory as growth versus substrate uptake (Fig. 2). Different studies have shown that during photoheterotrophic growth, among three different nitrogenase (Mo-, V-, and Fe-Nase) isozymes encoded in *R. palustris'* genome, Mo-Nase is exclusively expressed (49, 50). *A. vinelandii*, which has three different nitrogenases, also exclusively expresses the Mo-Nase (59). The ME model predicted exclusive expression of Mo-Nase during growth on all four carbon sources. The expression of nitrogenase may be dictated by its ATP requirements because Mo-Nase required the least amount of ATP among the three nitrogenases. In addition, the temperature of the assay played a role in the expression of different nitrogenases as discussed later.

Next, carbon fixation was also characterized with respect to substrate uptake. Unlike nitrogenase, which closely followed the trajectory of the growth rate, carbon fixation reached a peak flux at the start of the Janusian region. In the SPL region, when growth is proteome limited, *R. palustris* optimized protein production to sustain the growing biomass demand. As the cell approached the theoretical maximal growth, more ribose-5 phosphate was needed to sustain the increasing demand for nucleotides and lipopolysaccharides. To meet that demand at the theoretical maximum growth, the ME model predicted that *R. palustris* decreased the expression of phosphoribulokinase (*rpa4645*) and redirected flux toward ribose-5 phosphate production (Fig. 3).

During photoheterotrophic growth under nitrogen-fixing conditions, carbon and nitrogen fixation played a major role in maintaining cellular redox balance. However, in the SPL region, because the reaction flux of carbon fixation diminished at the theoretical maximum growth, the ME model predicted two potential candidates to maintain cellular redox balance: malate dehydrogenase and glycerol-3 phosphate dehydrogenase, in addition to nitrogen-fixing reaction. Malate dehydrogenase uses NAD+/NADH as cofactors and was encoded by *rpa0192*. Similarly, glycerol-3 phosphate dehydrogenase used NAD+/NADH as cofactors and was encoded by *rpa4410*. During the switch from the SNL to the SPL region, at the point where carbon fixation started to diminish, both malate dehydrogenase and glycerol-3 phosphate dehydrogenase fluxes started to increase (Fig. 4). At the theoretical maximum growth, flux through malate dehydrogenase and glycerol-3 phosphate dehydrogenase reached its maximum. Malate dehydrogenase also plays a role in maintaining redox balance in several other Gram-negative bacteria, such as organisms, including *E. coli* (60), and *Corynebacterium glutamicum* (61). Glycerol-3 phosphate dehydrogenase was one of the key enzymes in fatty acid biosynthesis. It was suggested that other biosynthetic pathways, such as fatty acid biosynthesis, could offer flexibility contributing to the redox balance for photoheterotrophically grown *R. rubrum* (62). In addition, several other organisms, such as *S. cerevisiae* (63) and *Kluyveromyces lactis* (64), showed evidence of using glycerol-3 phosphate dehydrogenase to maintain redox balance.
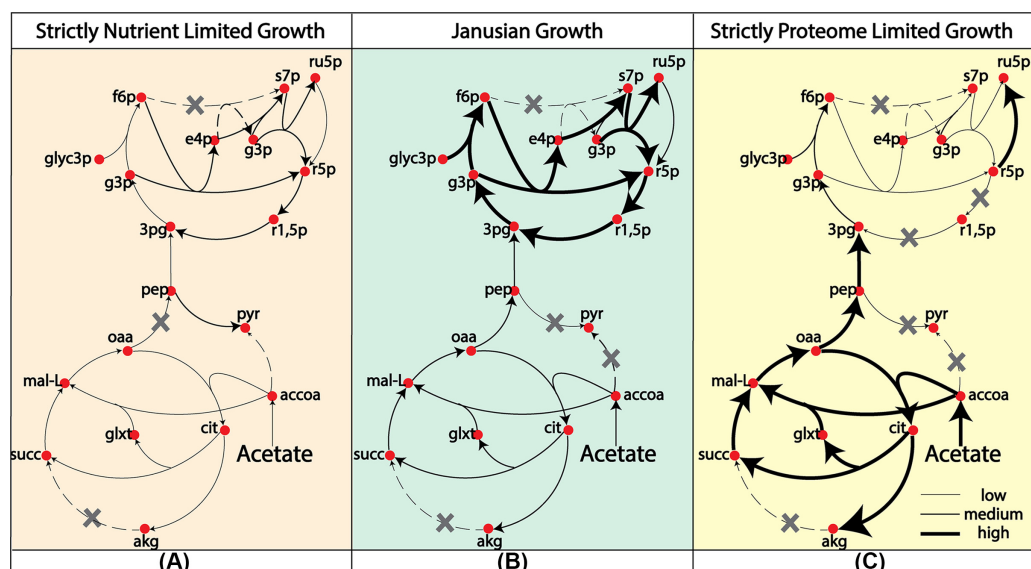
FIG 3 Metabolic activities in the (A) strictly nutrient-limited (SNL) growth, (B) Janusian growth, and (C) strictly proteome limited (SPL) growth. In the theoretical maximum growth, at the SPL region, flux through carbon fixation diminished and reaction flux from ribulose-5 phosphate (cpd00171) to ribose-5 phosphate (cpd00101) significantly increased. The increased biomass growth demand can be met by the precursors from the TCA cycle, which showed a significant increase in reaction flux compared to Janusian growth and SNL. Gray crosses indicate zero reaction flux.

Carbon fixation versus nitrogen fixation – competing metabolic modules for redox balance. During photoheterotrophic growth, *R. palustris* performs cyclic photophosphorylation (2, 21), which meant that electrons from photosystem I (PSI) got transported through ferredoxin, and the bc_1 complex and recycled back to PSI through the oxidation and reduction of quinones (24) (Fig. 5). Because there were no terminal electron acceptors, this could cause an accumulation of reduced cofactors, resulting in impeded growth of the bacterium. To resolve this, *R. palustris* employed various electron sinks to maintain a cellular redox balance. Calvin–Benson–Bassham (CBB) cycle pathway (48), nitrogen fixation pathway (65), and PHB production pathway only under nitrogen starvation condition (7) were some of the prominent pathways. During photoheterotrophic growth, the redox-balancing mechanism consisted primarily of the CBB pathway (48) and nitrogen fixation pathway (65). The nitrogen fixation module became active when *R. palustris* was placed in a nitrogen-limiting environment. Experimen-tal studies have suggested a link between carbon and nitrogen fixation that is intimately associated with the control of intracellular redox balance for different PNSBs, such as *R. palustris* (48), *R. capsulatus* (66), *R. sphaeroides* (65, 67), and *R. rubrum* (65). However, it is still not properly understood what factors decide the distribution of electrons in these two competing metabolic modules. Our previous study (24) indicated that with the increasing light uptake, the quinol oxidation rate also increases and the oxidation state of quinone acts as a feed-forward controller of the CBB cycle. Here, the ME model was used to further analyze the metabolic factors deciding the distribution of electron flux between carbon and nitrogen fixation in maintaining cellular redox balance.

To understand the electron distribution, a previous study eliminated rubisco activity in *R. palustris* by knocking out the relevant genes and found that the rate of nitrogen fixation did not vary significantly (48). Because CBB and nitrogen fixation pathways are two major redox balancing mechanisms, when rubisco was eliminated, the nitrogen fixation pathway was likely to carry an additional flux load to maintain cellular redox balance. Because this was not the case in the previous experimental study (48), it was suspected that there exists a metabolic bottleneck preventing additional reaction flux through the nitrogen fixation pathway. In the cyclic photophosphorylation of *R. palust-ris*, electrons get transported in a cyclical manner, and reduced ferredoxin supplies electrons to the nitrogen fixation pathway. Based on the availability of electrons, the nitrogen fixation pathway uses reduced cofactors to fix nitrogen. The more electrons supplied by ferredoxin, the more reduced cofactors will be used by nitrogen fixation
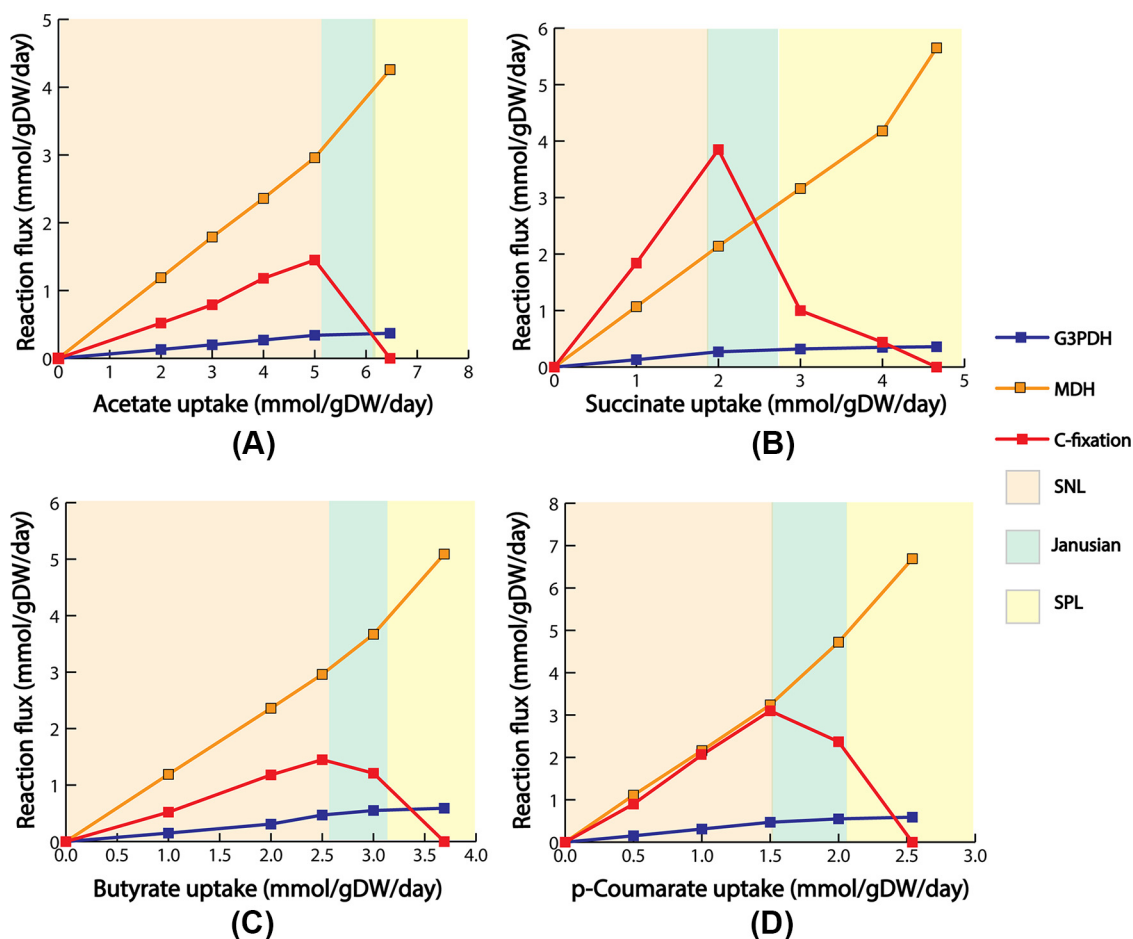
FIG 4 Alternate electron sink for different substrates (A) acetate, (B) succinate, (C) butyrate, and *p*-coumarate (D). In the Janusian regions, flux through carbon fixation reaction started to diminish. With the diminishing carbon fixation flux, the ME model predicted two alternate electrons, malate dehydrogenase, and glycerol-3 phosphate dehydrogenase. Reaction flux through these alternate electron sinks reached its peak when flux through carbon fixation completely diminished at the theoretical maximum growth.

pathways. Hence, less reduced cofactors were available for the carbon fixation pathway to use. Hence, electron transport through ferredoxin (ETFD) could be a potential candidate for the previously discussed bottleneck. Another element that transports electrons to the nitrogen fixation pathway is flavodoxin. A previous study (68) indicated that flavodoxin is a prominent electron donor only when *R. palustris* is under iron starvation. Flavodoxin is an isofunctional flavoprotein present in *R. palustris*, which is induced and replaces ferredoxin under stress conditions. In this study, we did not study the growth of *R. palustris* under any starvation situation. Hence, it is expected the role of flavodoxin in the redox balance will be minimum.

To explore if ETFD was indeed the hypothesized bottleneck, the biomass growth, and substrate uptake rate were kept constant, and only flux through carbon fixation reaction was varied for increasing flux of electron transport through the ferredoxin reaction (ETFD). At first, flux through ETFD was fixed to the solution found by the ME model (indicated by the red line in Fig. 6). The flux through the nitrogen-fixing reaction remained constant with changing flux through the carbon fixation reaction. This finding confirmed the presence of the previously hypothesized bottleneck. Increasing flux through ETFD had varying effects on the rate of nitrogen fixation depending on the utilized carbon substrate. When the reaction flux through ETFD was set to values higher than the ME model solution (indicated by the yellow and blue lines in Fig. 6), a very small change in flux through nitrogenase was noticed for the growth of acetate. For the other carbon sources, when the reaction flux through ETFD was set to values
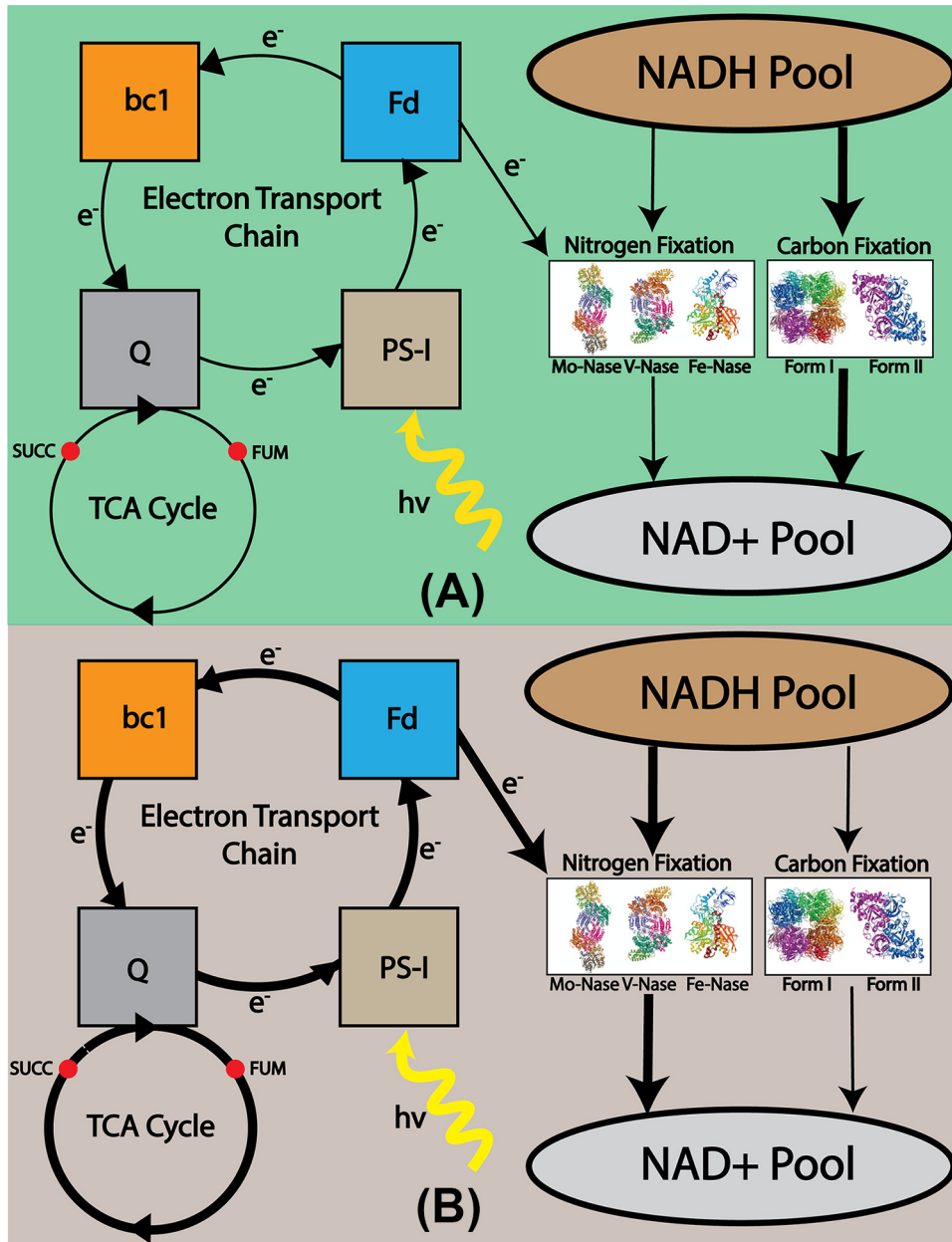
FIG 5 Relation between cyclic photophosphorylation and electron distribution between carbon and nitrogen fixation. (A) Less electron through ferredoxin indicates less flux through nitrogen fixation and more flux through the carbon fixation pathway. As a result, NADH will be more oxidized through a carbon fixation reaction. (B) More electron through ferredoxin indicates more flux through nitrogen fixation and less flux through the carbon fixation pathway. As a result, NADH would be more oxidized through a nitrogen fixation reaction.

higher than the ME model solution (indicated by the yellow and blue lines in Fig. 6), a negative correlation was observed between the carbon and nitrogen fixation reaction flux. When the metabolite pool size (see Text S1 for metabolite pool size calculation) was calculated for different cofactors, acetate produced less reduced cofactors per unit of substrate uptake compared to other substrates. As in this case, the fixed nitrogen is the sole source of nitrogen, cell prioritizes electron transport to nitrogenase rather than rubisco, whose primary function was to maintain the redox balance in the cell. Thus, the relation between carbon and nitrogen fixation was less visible for acetate. However, for succinate, butyrate, and *p*-coumarate, more reduced cofactors are produced per unit of substrate uptake. Thus, more electrons are available for the carbon
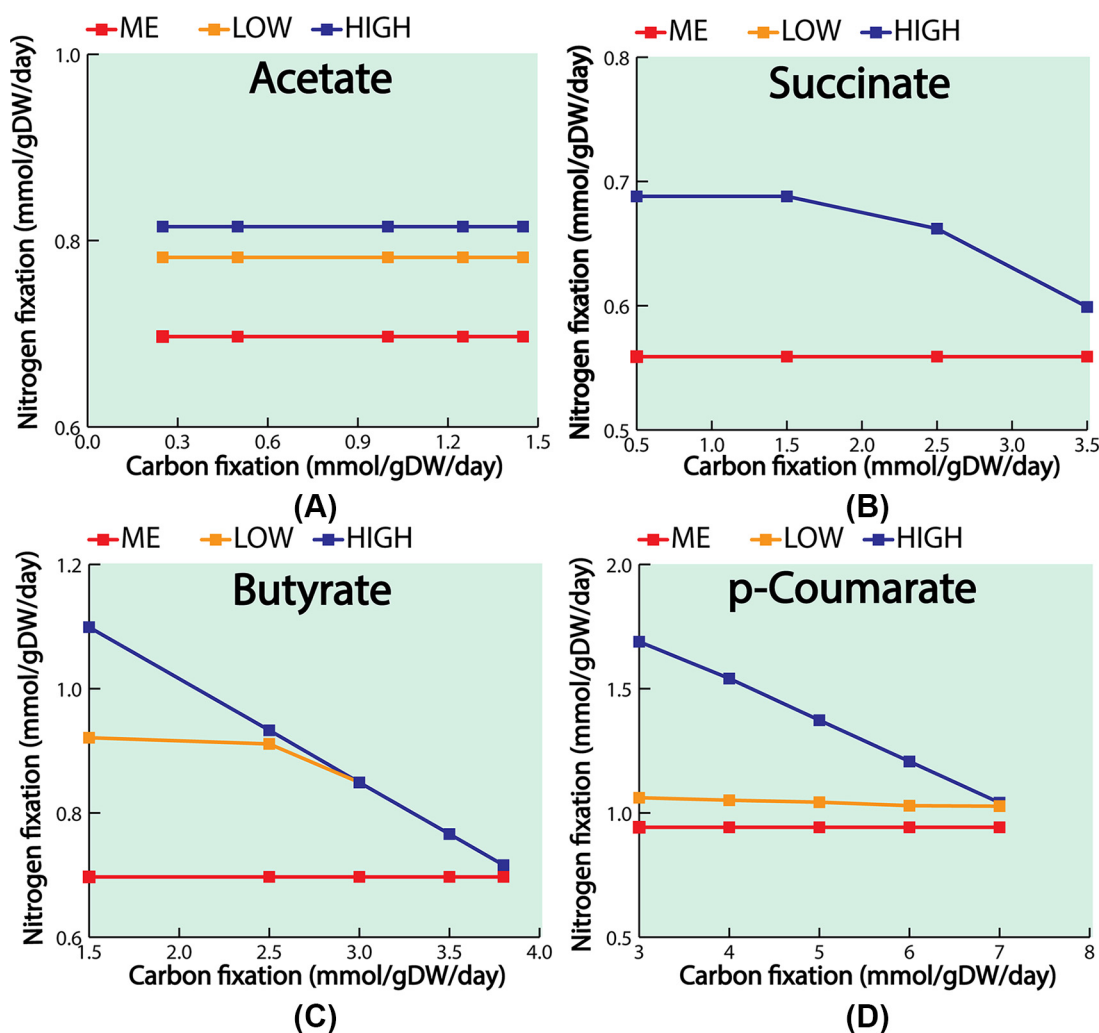
FIG 6 Relation between carbon fixation and nitrogen fixation with different fluxes via electron transport through ferredoxin (ETFD) for different substrates: (A) acetate, (B) succinate, (C) butyrate, and (D) *p*-coumarate. Red color lines indicate the relation between carbon fixation and nitrogen fixation when flux through ETFD was set to the solution found in the ME model. Blue color lines indicate the relation between carbon fixation and nitrogen fixation when ETFD flux was set to a high value. Yellow color lines indicate the relation between carbon fixation and nitrogen fixation when the ETFD flux values were set between ME and high.

fixation pathway and the regulation is more visible when ETFD flux is higher for these substrates. These results indicated that reaction flux through ETFD may play a regulatory role in distributing electron flux between carbon and nitrogen fixation.

Similar regulation in electron transport between competing metabolic modules, such as respiratory pathways and electron transport, can be observed in model bacteria *E. coli* (69). A highly organized network of overlapping transcriptional regulatory elements regulated the flow of electrons by controlling the expression of different genes in *E. coli*, including genes involved in substrates uptake, control of mixed-acid fermentation pathways, and controlling cofactor biosynthesis. Further experimentation is required to establish a similar molecular level mechanism for ETFD regulation of electron distribution in competing pathways of *R. palustris*. The ETFD regulation, hypothesized in this study, could have profound implications for future metabolic engineering efforts of *R. palustris*. Specifically, this regulation could be exploited to increase hydrogen production from *R. palustris* to achieve energy sustainability goals.

Characterization of Mo-, V-, and Fe-Nase nitrogenase enzymes. Because ETFD was postulated to play a regulatory role in distributing electrons to the nitrogen-fixing pathway, the ME model was next used to characterize how these electrons were used
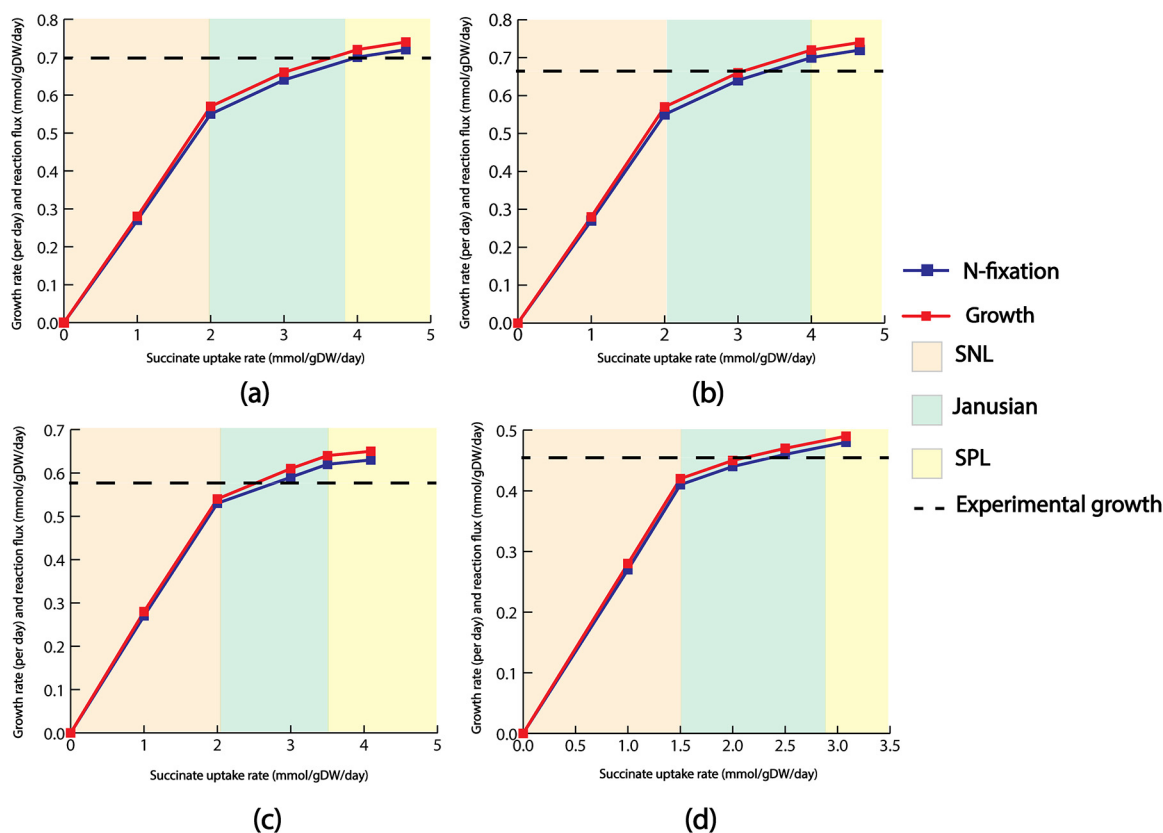
FIG 7 Growth rate and nitrogen fixation rate for (A) WT *R. palustris*, (B) Mo-only mutant, (C) V-only mutant, and (D) Fe-only mutant for succinate uptake. For each case, growth rate and nitrogen fixation closely followed each other. The dotted line in each graph indicates the experimentally observed growth.

by different nitrogenase enzymes. First, growth was simulated for the WT *R. palustris* with succinate as the substrate. In this case, only Mo-Nase was expressed and the growth versus substrate uptake curve (Fig. 7A) followed the pattern identified in the literature (35). Exclusively expressing the Mo-Nase in the WT was also consistent with previous literature findings (49, 50). Next, the growth versus substrate uptake graphs (Fig. 7B to D) was developed for three different mutants of *R. palustris*, each expressing a single nitrogenase isozyme. When the theoretical maximum growths for these mutants were compared with WT, it was found that WT and the Mo-only mutant had the highest growth rate followed by the growth rate of V-only and Fe-only mutants. Compared with the experimental growth rate data from the literature (49) for WT, Mo-only, V-only, and Fe-only growths followed a similar pattern as predicted by the ME model (Fig. S1). Theoretically, the growth of the WT, V-only, and Fe-only mutant strains of *R. palustris* can be coupled with the ATP requirement because Mo-nase required the least and Fe-nase required the most amount of ATP for nitrogen fixation.

Contrary to the pattern observed for succinate uptake, when other carbon sources (acetate, butyrate, and *p*-coumarate) were used as the substrates, V-Nase exhibited higher growth compared to Mo-Nase, Fe-Nase, and even WT (Fig. S2 to S4). To explain this mismatch, a previous study observed that the Mo-Nase was less expressed at a lower temperature compared to the other isozymes, such as V-Nase (49). Fig. 8A qualitatively summarized this idea. Because the experimental values used in that study were generated at 19°C, it is possible that Mo-Nase has less selectivity toward fixing nitrogen rate than other substrates, such as nitragin (N_2H_2), methane, ethane, etc. The effect of decreasing assay temperature on the activity of nitrogenase is complex. It was reported (70) that for the Mo-Nase of *A. vinelandii*, the rate of nitrogen reduction at 10°C is very low despite continued hydrolysis of ATP. In the
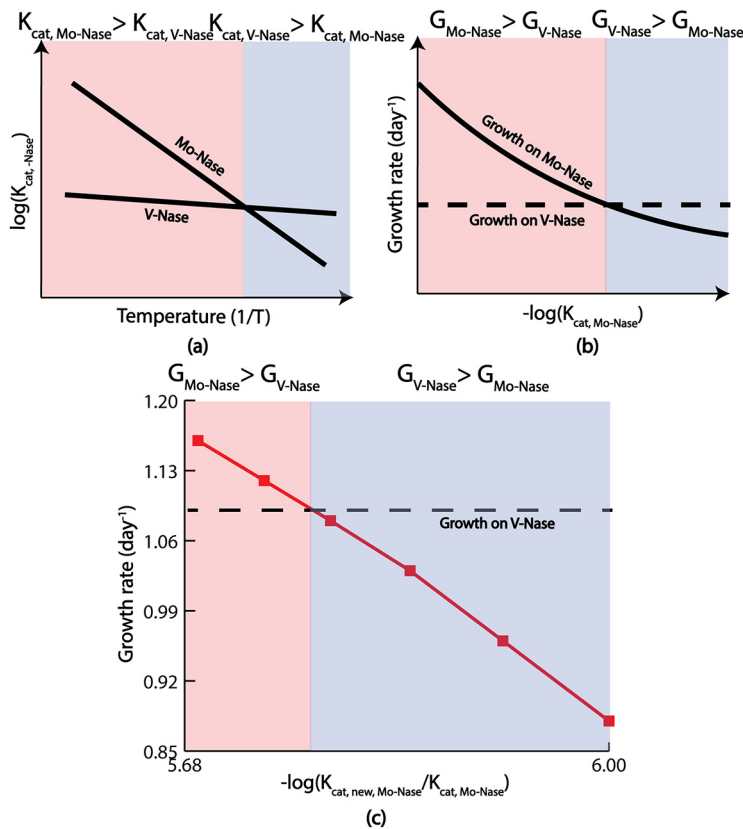
FIG 8 A representation of temperature regulation of Mo-Nase and V-Nase, and the effect of k_{cat} on the growth of Mo-only and V-only mutant. (A) From the literature, it was known that V-Nase had less sensitivity with respect to temperature compared to the Mo-Nase. The prediction from this ME model corroborated that finding. (B) Because k_{cat} is a parameter that is a function of temperature, from the Arrhenius equation, we knew that by reducing the temperature, k_{cat} was also reduced. With reducing k_{cat} at one stage, growth for the Mo-only mutant fell below the growth of the V-only mutant, capturing the experimentally observed temperature regulation of Mo- and V-Nase. (C) A quantitative representation of the idea mentioned in (B).

case of Mo-Nase of *Klebsiella pneumoniae*, decreasing the temperature not only curtailed electron flux but also resulted in the preferential loss of activity toward nitrogen as a substrate compared with H^+ or ethyne (C_2H_2) (71).

The ME model was further used to investigate the decreased growth rate of Mo-Nase at a lower temperature. From the Arrhenius equation (72), it is known that the turnover rate of an enzyme, k_{cat}, increases exponentially with the increasing temperature. Because k_{cat} is one of the temperature-sensitive parameters in this study, k_{cat} values of Mo- and V-Nase were varied to see at what point the V-Nase growth rate exceeded that of Mo-Nase and Mo-Nase. At first, the k_{cat} of V-Nase was increased to a very high value, but the growth rate of *R. palustris* fixing nitrogen through V-Nase was still lower than the WT and Mo-Nase. This tuning of k_{cat} indicated that the sensitivity of V-Nase activity with respect to temperature was low. This finding is consistent with previously published experimental work on another Gram-negative bacteria, *Azotobacter chroococcum* (73). Later, the k_{cat} of Mo-Nase was decreased to a very low value, and at that low k_{cat}, the growth rate of *R. palustris* fixing nitrogen through Mo-Nase was lower than the V-Nase and higher than the Fe-Nase, which is similar to the finding from literature (49). Therefore, by tuning the k_{cat}, the ME model was able to capture the experimentally observed temperature sensitivity of different nitrogenase enzymes. Fig. 8B qualitatively and Fig. 8C quantitatively summarized the effect of k_{cat} on the growth of Mo-only and V-only strains of *R. palustris*.

Conclusions. In this work, an ME model of *R. palustris* was developed. Growth rates predicted by the ME model for different substrates were closely matched
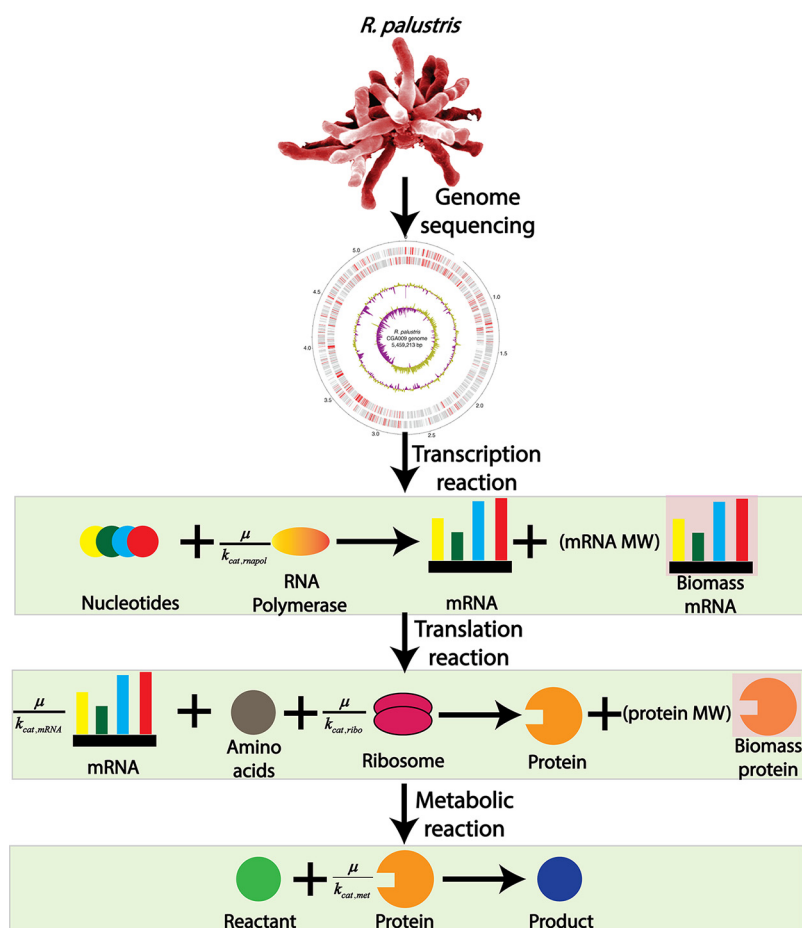
FIG 9 *R. palustris* ME model reconstruction. In the M model, only metabolic reactions were incorporated to perform genome-scale metabolic modeling. However, in the ME modeling framework, transcription and translation processes were also incorporated, adding two separate layers of regulation for metabolic reaction. Each layer of regulation was coupled with the biomass growth through catalytic turnover rate and biomass growth. This process is known as coupling and the coupling parameter is in the form of μ/k_{cat}, where μ indicates the growth rate and k_{cat} indicates the catalytic turnover rate for that.

with experimental growth rate data. The ME model also predicted a diminishing carbon fixation at the theoretical maximum growth and subsequently malate dehydrogenase and glycerol-3 phosphate dehydrogenase as alternate electron sinks. Furthermore, the ME model postulated electron transport through ferredoxin as a key regulatory feature to distribute reduced cofactor pools between carbon and nitrogen fixation pathways. Finally, the ME modeling framework successfully captured experimentally observed temperature regulation of different nitrogenase enzymes.

Going forward, this ME model could be used as a powerful platform to further characterize different features of *R. palustris* metabolism. Specially characterizing a complete profile of environment-specific isozyme expressions and optimal protein allocation. Furthermore, this ME model can be used to design and fine-tune mutants of *R. palustris* for metabolic engineering purposes. One such application would be to produce PHB, a bioplastic precursor, which has the potential to replace petroleum-based plastics. Under anaerobic-photoheterotrophic growth of *R. palustris*, PHB can work as an electron sink (7). Our previous effort (7) successfully established three design strategies to select the ideal lignin breakdown products (LBPs) for commercial PHB production from *R. palustris*. This ME modeling framework could be further used to gain similar regulatory insights, as discussed here, on how electrons are distributed in PHB-producing pathways when different LBPs are used as the substrates.

MATERIALS AND METHODS

ME model of *R. palustris*. In addition to the metabolic reactions from the M model, the ME model consisted of translation and transcription reactions (Fig. 9). To model transcription and translation reactions, a GPR association of each reaction was required. The initial GPR association was collected from the literature (24). Later, the GPR association was manually curated using the detailed genome annotation from literature (2).

This GPR association is in Table S4. The overall ME model reconstruction procedure was conducted in accordance with the COBRAme protocol (30), which is summarized in Fig. 2. The ME model is a multi-scale model (30). Hence, it required the addition of coupling constraints to relate different cellular processes to each other. The coupling constraints are in the form of μ/k_{cat}. Here, μ was the growth rate and k_{cat} approximated the effective turnover rate for the different macromolecules. Detailed mathematical descriptions for the μ/k_{cat} of different macromolecular processes and values of different parameters are in Text S1 and the original COBRAme protocol (30).

In terms of including tRNA charging reactions in the ME model (74), initiation was the rate-limiting step in the translation reaction rather than the elongation process. A similar result was obtained for the other prokaryotic and eukaryotic organisms (75, 76). Hence, we did not include the tRNA charging reactions in the ME model because it was not the rate-limiting step in the translation reaction. For protein translocation, the *R. palustris* model had only one compartment, cytosol, apart from the extracellular space. Hence, the need for translocation reactions was absent. Because DNA replication only ensures that the chromosome was copied at each cell doubling, the rate of this "reaction" would only be dependent on the growth rate and not on the growth condition. For posttranslational modifications, there is evidence of posttranslational modification for nitrogenase enzymes (77). But for most other proteins, posttranslational modifications were not known. If we had added posttranslational modifications only for nitrogenases, that would have introduced bias against nitrogenase enzymes compared to other enzymes from a resource allocation perspective. Hence, we chose not to add that bias to the model.

To calculate k_{cat} for different enzymes, a mean k_{cat} value of 234,000 day^{-1} was used, which was reported for the photosynthetic cyanobacteria (78). This k_{cat} was further modified for each enzyme based on the solvent-accessible surface area (SASA), following the same ME modeling framework (35). SASA was defined as the surface area of an enzyme that was accessible to a solvent. In addition, a previous study (79) reported a correlation between SASA and the molecular weight of the enzyme as follows: SASA = (molecular weight of the enzyme)$^{3/4}$. Overall, the following equation was used to calculate $k_{cat,enzyme}$ for each enzyme, based on the mean turnover rate ($k_{cat,\ mean}$), mean SASA ($SASA_{mean}$), and SASA for the specific enzyme ($SASA_{enzyme}$): $k_{cat,enzyme} = k_{cat,\ mean} \times SASA_{enzyme}/SASA_{mean}$. For transcription reactions, RNA polymerase was needed to produce the required mRNA for protein production. RNA polymerase of *R. palustris* consisted of five subunits: two alpha (α) subunits, a beta (β) subunit, a beta prime subunit (β'), and a small omega (ω) subunit (80). In the model, each subunit was synthesized to form the RNA polymerase. Later, these RNA polymerases transformed different nucleotides into mRNA.

For translation reactions, rRNA was required to transform amino acids into different proteins. *R. palustris* utilized 70S ribosomes, each consisting of a small (30S) and a large (50S) subunit (81). The large subunit was composed of a 5S RNA subunit (120 nucleotides), a 23S RNA subunit (2900 nucleotides), and 31 proteins. The small subunit was composed of a 16S RNA subunit (1542 nucleotides) and 21 proteins (81).

For each transcription or translation reaction in the ME model, an amount of a biomass protein and biomass mRNA were produced with a stoichiometry equal to the molecular weight (in kDa) of the protein or mRNA being made. Fig. 9 shows an example of this where the translation reaction produces both the catalytic protein and the biomass protein. Similarly, the transcription reaction produced mRNA required for the protein synthesis and biomass mRNA. The biomass protein and mRNA participated in the ME model biomass dilution reaction, restricting the total biomass components production equal to the rate of biomass dilution.

Transcription and translation reactions were included for all reactions for which GPR was available. For the remaining pathways, an enzyme was used with an average length of 283 amino acids and a molecular weight of 31.09 kDa based on the literature (39).

To capture the differential expression of the carbon fixing isozymes, a constraint was added to the ME model to account for the coexpression of both rubisco form I and form II as follows:

$$v_{rubisco\,I} = \left[-\sum v_{CO_2} + \frac{\mu}{\mu_{max}} \sum v_{CO_{2,max}} \right] \times \frac{k_{cat,\,rubisco\,I}}{k_{cat,\,rubisco\,II}} \tag{1}$$

In equation (1), $v_{rubisco\ I}$ represent the expression of rubisco form I, which is a function of carbon dioxide generation ($\sum v_{CO_2}$), growth rate (μ), theoretical maximum growth rate (μ_{max}), carbon dioxide generation at the theoretical maximum growth rate ($\sum v_{CO_{2,max}}$), and effective catalytic rate of rubisco form I ($k_{cat,\ rubisco\ I}$) and rubisco form II ($k_{cat,\ rubisco\ II}$).

The main source of ATP production in *R. palustris* is photosynthesis (24). Hence, for each the substrate, the maximum ATP production by the ME model was capped by its photosynthetic yield according to the following equation proposed in the literature (7):

$$v_{PSII}^{S} = v_{PSII}^{ace} \frac{\varnothing_{PSII}^{S}}{\varnothing_{PSII}^{ace}} \tag{2}$$

Here, "S" and "ace" refer to different substrates and acetate, respectively. Also, $\varnothing_{PSII}^{S}$ and $\varnothing_{PSII}^{ace}$ refer to the photosynthetic yield of different substrates and acetate, respectively. Photosynthetic yields for different substrates are calculated from literature (7) and provided in Table 2 below.

TABLE 2 Maximum ATP production rate calculated from photosynthetic yield for substrates

Substrates	ATP production rate (mmol. gDW^{-1} day^{-1})
Acetate	54.0
Succinate	45.7
Butyrate	56.7
p-coumarate	85.4

Accuracy calculation in the validation study. In the validation study, using the ME model, aerobic growth of *R. palustris* was simulated with *p*-coumarate and succinate as sources of carbon and $(NH_4)_2SO_4$ as a sole source of nitrogen. From the ME model, fluxes of transcriptomics and proteomics reactions were calculated for both carbon sources. Considering the transcriptomics and proteomics reaction fluxes for succinate uptake as the baseline condition, fold changes for all the gene expressions and protein were calculated for *p*-coumarate uptake. If the fold change is greater than 1, it was noted as upregulated. If the fold change is less than 1, it was noted as downregulated. Once the upregulated/downregulated fold changes of transcription and translation reactions were calculated and fold changes were compared with the literature (38). If both fold changes, from the ME model and the experimental study, showed the same direction (upregulated or downregulated) of fold change, then the prediction is correct. Otherwise, the prediction is incorrect. Accuracy was then calculated as a percentage between correct prediction and total predictions.

Simulation tools and software. Due to the large variability of scaling of different coefficients, the ME model was solved using the General Algebraic Modeling System (GAMS) version 24.7.4 with IBM CPLEX solver using the pFBA (52) algorithm. The GAMS files were run on a Linux-based high-performance cluster computing system at the University of Nebraska-Lincoln. For community use, an interface between GAMS Studio and NEOS Server is available to run the ME model without having to buy any license. Moreover, a python implementation is also available for interested users. Note that, python users will at least need to run the GAMS model once to generate the necessary input files. Text S2 has further instructions on how to run the GAMS and Python files using GAMS Studio and NEOS Server, without having to buy any license.

Because the ME model contains coupling coefficients. Hence, it is a nonlinear optimization problem. Thus, to find the optimal growth rate from the ME model, we parameterized the optimal biomass growth rate. To solve the optimization problem, we first set the biomass growth rate to a fixed value to make the problem linear. For a given substrates (acetate, succinate, butyrate, and *p*-coumarate) uptake rate, different biomass growth rates were simulated until the optimization problem became infeasible. Then the biomass growth rate just before the optimization problem become infeasible was reported as the optimal biomass growth rate.

Data availability. All code used in this work can be found in the following GitHub directory: https://github.com/ssbio/palustris_ME_model.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.2 MB.
SUPPLEMENTAL FILE 2, XLSX file, 0.04 MB.
SUPPLEMENTAL FILE 3, XLSX file, 0.1 MB.
SUPPLEMENTAL FILE 4, XLSX file, 0.1 MB.
SUPPLEMENTAL FILE 5, XLSX file, 1.4 MB.
SUPPLEMENTAL FILE 6, PDF file, 0.1 MB.
SUPPLEMENTAL FILE 7, PDF file, 0.3 MB.

ACKNOWLEDGMENTS

REFERENCES

1. VerBerkmoes NC, Shah MB, Lankford PK, Pelletier DA, Strader MB, Tabb DL, McDonald WH, Barton JW, Hurst GB, Hauser L, Davison BH, Beatty JT, Harwood CS, Tabita FR, Hettich RL, Larimer FW. 2006. Determination and comparison of the baseline proteomes of the versatile microbe Rhodopseudomonas palustris under its major metabolic states. J Proteome Res 5:287–298. https://doi.org/10.1021/pr0503230.

2. Larimer FW, Chain P, Hauser L, Lamerdin J, Malfatti S, Do L, Land ML, Pelletier DA, Beatty JT, Lang AS, Tabita FR, Gibson JL, Hanson TE, Bobst C, Torres J. L T y, Peres C, Harrison FH, Gibson J, Harwood CS. 2004. Complete genome sequence of the metabolically versatile

photosynthetic bacterium Rhodopseudomonas palustris. Nat Biotechnol 22:55–61. https://doi.org/10.1038/nbt923.

3. Sakpirom J, Kantachote D, Nunkaew T, Khan E. 2017. Characterizations of purple non-sulfur bacteria isolated from paddy fields, and identification of strains with potential for plant growth-promotion, greenhouse gas mitigation and heavy metal bioremediation. Res Microbiol 168:266–275. https://doi.org/10.1016/j.resmic.2016.12.001.

4. Nookongbut P, Kantachote D, Megharaj M. 2016. Arsenic contamination in areas surrounding mines and selection of potential As-resistant purple nonsulfur bacteria for use in bioremediation based on their detoxification

mechanisms. Ann Microbiol 66:1419–1429. https://doi.org/10.1007/s13213-016-1229-z.

5. Lu H, Yuan Y, Campbell DE, Qin P, Cui L. 2014. Integrated water quality, energy and economic evaluation of three bioremediation treatment systems for eutrophic water. Ecol Eng 69:244–254. https://doi.org/10.1016/j.ecoleng.2014.04.024.

6. Brown B, Immethun C, Wilkins M, Saha R. 2020. Rhodopseudomonas palustris CGA009 polyhydroxybutyrate production from a lignin aromatic and quantification via flow cytometry. Bioresour Technol Rep 11:100474. https://doi.org/10.1016/j.biteb.2020.100474.

7. Alsiyabi A, Brown B, Immethun C, Long D, Wilkins M, Saha R. 2021. Synergistic experimental and computational approach identifies novel strategies for polyhydroxybutyrate overproduction. Metab Eng 68:1–13. https://doi.org/10.1016/j.ymben.2021.08.008.

8. Wang X, Feng Y, Wang H, Qu Y, Yu Y, Ren N, Li N, Wang E, Lee H, Logan BE. 2009. Bioaugmentation for electricity generation from corn stover biomass using microbial fuel cells. Environ Sci Technol 43:6088–6093. https://doi.org/10.1021/es900391b.

9. Sun J, Yang P, Li N, Zhao M, Zhang X, Zhang Y, Yuan Y, Lu X, Lu X. 2020. Extraction of photosynthetic electron from mixed photosynthetic consortium of bacteria and algae towards sustainable bioelectrical energy harvesting. Electrochim Acta 336:135710. https://doi.org/10.1016/j.electacta.2020.135710.

10. Wu P, Chen Z, Zhang Y, Wang Y, Zhu F, Cao B, Wu Y, Li N. 2019. Rhodopseudomonas palustris wastewater treatment: cyhalofop-butyl removal, biochemicals production and mathematical model establishment. Bioresour Technol 282:390–397. https://doi.org/10.1016/j.biortech.2018.11.087.

11. Kim MK, Choi K-M, Yin C-R, Lee K-Y, Im W-T, Lim JH, Lee S-T. 2004. Odorous swine wastewater treatment by purple non-sulfur bacteria, Rhodopseudomonas palustris, isolated from eutrophicated ponds. Biotechnol Lett 26:819–822. https://doi.org/10.1023/B:BILE.0000025884.50198.67.

12. Lu H, Zhang G, He S, Zhao R, Zhu D. 2021. Purple non-sulfur bacteria technology: a promising and potential approach for wastewater treatment and bioresources recovery. World J Microbiol Biotechnol 37:161. https://doi.org/10.1007/s11274-021-03133-z.

13. Mabutyana L, Pott RWM. 2021. Photo-fermentative hydrogen production by Rhodopseudomonas palustris CGA009 in the presence of inhibitory compounds. Int J Hydrogen Energy 46:29088–29099. https://doi.org/10.1016/j.ijhydene.2020.12.189.

14. Du Toit J-P, Pott RWM. 2021. Heat-acclimatised strains of Rhodopseudomonas palustris reveal higher temperature optima with concomitantly enhanced biohydrogen production rates. Int J Hydrogen Energy 46:11564–11572. https://doi.org/10.1016/j.ijhydene.2021.01.068.

15. Ji Y, Sultan MA, Kim DY, Meeks N, Hastings JT, Bhattacharyya D. 2021. Effect of silica-core gold-shell nanoparticles on the kinetics of biohydrogen production and pollutant hydrogenation via organic acid photofermentation over enhanced near-infrared illumination. Int J Hydrogen Energy 46:7821–7835. https://doi.org/10.1016/j.ijhydene.2020.11.257.

16. Gosse JL, Engel BJ, Hui JC-H, Harwood CS, Flickinger MC. 2010. Progress toward a biomimetic leaf: 4,000 h of hydrogen production by coating-stabilized nongrowing photosynthetic Rhodopseudomonas palustris. Biotechnol Prog 26:907–918. https://doi.org/10.1002/btpr.406.

17. Hongyuan L, Jiahua C, Yangyang J, Mingwei C, Lee PKH. 2016. Transcriptomic responses of the interactions between Clostridium cellulovorans 743B and Rhodopseudomonas palustris CGA009 in a cellulose-grown co-culture for enhanced hydrogen production. Appl Environ Microbiol 82:4546–4559. https://doi.org/10.1128/AEM.00789-16.

18. Setubal JC, dos Santos P, Goldman BS, Ertesvåg H, Espin G, Rubio LM, Valla S, Almeida NF, Balasubramanian D, Cromes L, Curatti L, Du Z, Godsy E, Goodner B, Hellner-Burris K, Hernandez JA, Houmiel K, Imperial J, Kennedy C, Larson TJ, Latreille P, Ligon LS, Lu J, Maerk M, Miller NM, Norton S, O'Carroll IP, Paulsen I, Raulfs EC, Roemer R, Rosser J, Segura D, Slater S, Stricklin SL, Studholme DJ, Sun J, Viana CJ, Wallin E, Wang B, Wheeler C, Zhu H, Dean DR, Dixon R, Wood D. 2009. Genome sequence of Azotobacter vinelandii, an obligate aerobe specialized to support diverse anaerobic metabolic processes. J Bacteriol 191:4534–4545. https://doi.org/10.1128/JB.00504-09.

19. Orth JD, Thiele I, Palsson BO. 2010. What is flux balance analysis? Nat Biotechnol 28:245–248. https://doi.org/10.1038/nbt.1614.

20. Klamt S, Schuster S, Gilles ED. 2002. Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria. Biotechnol Bioeng 77:734–751. https://doi.org/10.1002/bit.10153.

21. Klamt S, Grammel H, Straube R, Ghosh R, Gilles ED. 2008. Modeling the electron transport chain of purple non-sulfur bacteria. Mol Syst Biol 4:156. https://doi.org/10.1038/msb4100191.

22. Imam S, Yilmaz S, Sohmen U, Gorzalski AS, Reed JL, Noguera DR, Donohue TJ. 2011. iRsp1095: a genome-scale reconstruction of the Rhodobacter sphaeroides metabolic network. BMC Syst Biol 5:116. https://doi.org/10.1186/1752-0509-5-116.

23. Navid A, Jiao Y, Wong SE, Pett-Ridge J. 2019. System-level analysis of metabolic trade-offs during anaerobic photoheterotrophic growth in Rhodopseudomonas palustris. BMC Bioinformatics 20:233. https://doi.org/10.1186/s12859-019-2844-z.

24. Alsiyabi A, Immethun CM, Saha R. 2019. Modeling the interplay between photosynthesis, CO_2 fixation, and the quinone pool in a purple non-sulfur bacterium. Sci Rep 9:1–9. https://doi.org/10.1038/s41598-019-49079-z.

25. Arenz S, Wilson DN. 2016. Bacterial protein synthesis as a target for antibiotic inhibition. Cold Spring Harb Perspect Med 6:a025361. https://doi.org/10.1101/cshperspect.a025361.

26. Reed JL, Palsson BØ. 2004. Genome-scale in silico models of E. coli have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. Genome Res 14:1797–1805. https://doi.org/10.1101/gr.2546004.

27. Schellenberger J, Lewis NE, Palsson BØ. 2011. Elimination of thermodynamically infeasible loops in steady-state metabolic models. Biophys J 100:544–553. https://doi.org/10.1016/j.bpj.2010.12.3707.

28. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW. 2012. A whole-cell computational model predicts phenotype from genotype. Cell 150:389–401. https://doi.org/10.1016/j.cell.2012.05.044.

29. Bulović A, Fischer S, Dinh M, Golib F, Liebermeister W, Poirier C, Tournier L, Klipp E, Fromion V, Goelzer A. 2019. Automated generation of bacterial resource allocation models. Metab Eng 55:12–22. https://doi.org/10.1016/j.ymben.2019.06.001.

30. Lloyd CJ, Ebrahim A, Yang L, King ZA, Catoiu E, O'Brien EJ, Liu JK, Palsson BO. 2018. COBRAme: a computational framework for genome-scale models of metabolism and gene expression. PLoS Comput Biol 14:e1006302. https://doi.org/10.1371/journal.pcbi.1006302.

31. Alsiyabi A, Chowdhury NB, Long D, Saha R. 2022. Enhancing in silico strain design predictions through next generation metabolic modeling approaches. Biotechnol Adv 54:107806. https://doi.org/10.1016/j.biotechadv.2021.107806.

32. Lerman JA, Hyduke DR, Latif H, Portnoy VA, Lewis NE, Orth JD, Schrimpe-Rutledge AC, Smith RD, Adkins JN, Zengler K, Palsson BO. 2012. In silico method for modelling metabolism and gene product expression at genome scale. Nat Commun 3:929. https://doi.org/10.1038/ncomms1928.

33. Liu JK, Lloyd C, Al-Bassam MM, Ebrahim A, Kim J-N, Olson C, Aksenov A, Dorrestein P, Zengler K. 2019. Predicting proteome allocation, overflow metabolism, and metal requirements in a model acetogen. PLoS Comput Biol 15:e1006848. https://doi.org/10.1371/journal.pcbi.1006848.

34. Oftadeh O, Salvy P, Masid M, Curvat M, Miskovic L, Hatzimanikatis V. 2021. A genome-scale metabolic model of Saccharomyces cerevisiae that integrates expression constraints and reaction thermodynamics. Nat Commun 12:4790. https://doi.org/10.1038/s41467-021-25158-6.

35. O'Brien EJ, Lerman JA, Chang RL, Hyduke DR, Palsson BØ. 2013. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. Mol Syst Biol 9:693. https://doi.org/10.1038/msb.2013.52.

36. Liu JK, O'Brien EJ, Lerman JA, Zengler K, Palsson BO, Feist AM. 2014. Reconstruction and modeling protein translocation and compartmentalization in Escherichia coli at the genome-scale. BMC Syst Biol 8:110. https://doi.org/10.1186/s12918-014-0110-6.

37. Du B, Yang L, Lloyd CJ, Fang X, Palsson BO. 2019. Genome-scale model of metabolism and gene expression provides a multi-scale description of acid stress responses in Escherichia coli. PLoS Comput Biol 15:e1007525. https://doi.org/10.1371/journal.pcbi.1007525.

38. Pan C, Oda Y, Lankford PK, Zhang B, Samatova NF, Pelletier DA, Harwood CS, Hettich RL. 2008. Characterization of anaerobic catabolism of p-Coumarate in Rhodopseudomonas palustris by integrating transcriptomics and quantitative proteomics. Mol Cell Proteomics 7:938–948. https://doi.org/10.1074/mcp.M700147-MCP200.

39. Kozlowski LP. 2017. Proteome-pl: proteome isoelectric point database. Nucleic Acids Res 45:D1112. https://doi.org/10.1093/nar/gkw978.

40. Davidi D, Shamshoum M, Guo Z, Bar-On YM, Prywes N, Oz A, Jablonska J, Flamholz A, Wernick DG, Antonovsky N, Pins B, Shachar L, Hochhauser D, Peleg Y, Albeck S, Sharon I, Mueller-Cajar O, Milo R. 2020. Highly active rubiscos discovered by systematic interrogation of natural sequence diversity. EMBO J 39:e104081. https://doi.org/10.15252/embj.2019104081.

41. Consortium TU. 2019. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47:D506. https://doi.org/10.1093/nar/gky1049.

42. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. Nucleic Acids Res 36:D480–4. https://doi.org/10.1093/nar/gkm882.

43. Bar-On YM, Milo R. 2019. The global mass and average rate of rubisco. Proc Natl Acad Sci U S A 116:4738–4743. https://doi.org/10.1073/pnas.1816654116.

44. Horken KM, Tabita FR. 1999. The "Green" form I ribulose 1,5-bisphosphate carboxylase/oxygenase from the nonsulfur purple bacterium Rhodobacter capsulatus. J Bacteriol 181:3935–3941. https://doi.org/10.1128/JB.181.13.3935-3941.1999.

45. Tabita FR, Satagopan S, Hanson TE, Kreel NE, Scott SS. 2008. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. J Exp Bot 59:1515–1524. https://doi.org/10.1093/jxb/erm361.

46. Erb TJ, Zarzycki J. 2018. A short history of RubisCO: the rise and fall (?) of Nature's predominant CO_2 fixing enzyme. Curr Opin Biotechnol 49:100–107. https://doi.org/10.1016/j.copbio.2017.07.017.

47. Higuchi-Takeuchi M, Morisaki K, Numata K. 2016. A screening method for the isolation of polyhydroxyalkanoate-producing purple non-sulfur photosynthetic bacteria from natural seawater. Front Microbiol 7:1509. https://doi.org/10.3389/fmicb.2016.01509.

48. McKinlay JB, Harwood CS. 2010. Carbon dioxide fixation as a central redox cofactor recycling mechanism in bacteria. Proc Natl Acad Sci U S A 107:11669–11675. https://doi.org/10.1073/pnas.1006175107.

49. Luxem KE, Kraepiel AML, Zhang L, Waldbauer JR, Zhang X. 2020. Carbon substrate re-orders relative growth of a bacterium using Mo-, V-, or Fe-nitrogenase for nitrogen fixation. Environ Microbiol 22:1397–1408. https://doi.org/10.1111/1462-2920.14955.

50. Yasuhiro O, Samanta SK, Rey FE, Wu L, Liu X, Yan T. 2005. Functional genomic analysis of three nitrogenase isozymes in the photosynthetic bacterium Rhodopseudomonas palustris. J Bacteriol 187:7784–7794. https://doi.org/10.1128/JB.187.22.7784-7794.2005.

51. Joshi GS, Romagnoli S, Verberkmoes NC, Hettich RL, Pelletier D, Tabita FR. 2009. Differential accumulation of form I RubisCO in Rhodopseudomonas palustris CGA010 under photoheterotrophic growth conditions with reduced carbon sources. J Bacteriol 191:4243–4250. https://doi.org/10.1128/JB.01795-08.

52. 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.

53. Hu Z, Dong H, Ma J-C, Yu Y, Li K-H, Guo Q-Q, Zhang C, Zhang W-B, Cao X, Cronan JE, Wang H. 2018. Novel Xanthomonas campestris long-chain-specific 3-oxoacyl-acyl carrier protein reductase involved in diffusible signal factor synthesis. mBio 9:e00596-18. https://doi.org/10.1128/mBio.00596-18.

54. Çeti²Nkaya DG, Öztürk A, Çakmakçı L. 1999. Properties of the Rhodopseudomonas palustris strains isolated from an alkaline lake in Turkey. Tr J Biol 23:457–463.

55. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, Cheng T-Y, Moody DB, Murray M, Galagan JE. 2009. Interpreting expression data with metabolic flux models: predicting Mycobacterium tuberculosis mycolic acid production. PLoS Comput Biol 5:e1000489. https://doi.org/10.1371/journal.pcbi.1000489.

56. McKinlay JB, Harwood CS, Newman DK. 2011. Calvin cycle flux, pathway constraints, and substrate oxidation state together determine the H2 biofuel yield in photoheterotrophic bacteria. mBio 2:e00323-10. https://doi.org/10.1128/mBio.00323-10.

57. Monod J. 1949. The growth of bacterial cultures. Annu Rev Microbiol 3:371–394. https://doi.org/10.1146/annurev.mi.03.100149.002103.

58. Koch AL. 1997. Microbial physiology and ecology of slow growth. Microbiol Mol Biol Rev 61:305–318. https://doi.org/10.1128/mmbr.61.3.305-318.1997.

59. Jacobitz S, Bishop PE. 1992. Regulation of nitrogenase-2 in Azotobacter vinelandii by ammonium, molybdenum, and vanadium. J Bacteriol 174:3884–3888. https://doi.org/10.1128/jb.174.12.3884-3888.1992.

60. Jantama K, Haupt MJ, Svoronos SA, Zhang X, Moore JC, Shanmugam KT, Ingram LO. 2008. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate. Biotechnol Bioeng 99:1140–1153. https://doi.org/10.1002/bit.21694.

61. Mitsuhashi S, Hayashi M, Ohnishi J, Ikeda M. 2006. Disruption of malate:quinone oxidoreductase increases L-lysine production by Corynebacterium glutamicum. Biosci Biotechnol Biochem 70:2803–2806. https://doi.org/10.1271/bbb.60298.

62. McCully AL, Onyeziri MC, LaSarre B, Gliessman JR, McKinlay JB. 2020. Reductive tricarboxylic acid cycle enzymes and reductive amino acid synthesis pathways contribute to electron balance in a Rhodospirillum rubrum Calvin-cycle mutant. Microbiology (Reading) 166:199–211. https://doi.org/10.1099/mic.0.000877.

63. Lopes M. d B, -Ur-Rehman ATA, Gockowiak H, Heinrich AJ, Langridge P, Henschke PA. 2000. Fermentation properties of a wine yeast over-expressing the Saccharomyces cerevisiae glycerol 3-phosphate dehydrogenase gene (GPD2). Aust J Grape Wine Res 6:208–215. https://doi.org/10.1111/j.1755-0238.2000.tb00181.x.

64. Guadalupe-Medina V, Metz B, Oud B, van Der Graaf CM, Mans R, Pronk JT, van Maris AJA. 2014. Evolutionary engineering of a glycerol-3-phosphate dehydrogenase-negative, acetate-reducing Saccharomyces cerevisiae strain enables anaerobic growth at high glucose concentrations. Microb Biotechnol 7:44–53. https://doi.org/10.1111/1751-7915.12080.

65. Joshi HM, Tabita FR. 1996. A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. Proc Natl Acad Sci U S A 93:14515–14520. https://doi.org/10.1073/pnas.93.25.14515.

66. Tichi MA, Tabita FR. 2000. Maintenance and control of redox poise in Rhodobacter capsulatus strains deficient in the Calvin-Benson-Bassham pathway. Arch Microbiol 174:322–333. https://doi.org/10.1007/s002030000209.

67. Yilei Q, Robert TF. 1998. Expression of glnB and aglnB-Like Gene (glnK) in a Ribulose Bisphosphate Carboxylase/Oxygenase-Deficient Mutant of Rhodobacter sphaeroides. J Bacteriol 180:4644–4649.

68. Fixen KR, Pal Chowdhury N, Martinez-Perez M, Poudel S, Boyd ES, Harwood CS. 2018. The path of electron transfer to nitrogenase in a phototrophic alpha-proteobacterium. Environ Microbiol 20:2500–2508. https://doi.org/10.1111/1462-2920.14262.

69. Gunsalus RP. 1992. Control of electron flow in Escherichia coli: coordinated transcription of respiratory pathway genes. J Bacteriol 174:7069–7074. https://doi.org/10.1128/jb.174.22.7069-7074.1992.

70. Watt GD, Burns A. 1977. Kinetics of dithionite ion utilization and ATP hydrolysis for reactions catalyzed by the nitrogenase complex from Azotobacter vinelandii. Biochemistry 16:264–270. https://doi.org/10.1021/bi00621a017.

71. Thorneley RNF, Eady RR. 1977. Nitrogenase of Klebsiella pneumoniae. Distinction between proton-reducing and acetylene-reducing forms of the enzyme: effect of temperature and component protein ratio on substrate-reduction kinetics. Biochem J 167:457–461. https://doi.org/10.1042/bj1670457.

72. Jensen F. 1985. Activation energies and the arrhenius equation. Qual Reliab Engng Int 1:13–17. https://doi.org/10.1002/qre.4680010104.

73. Miller RW, Eady RR. 1988. Molybdenum and vanadium nitrogenases of Azotobacter chroococcum. Low temperature favours N2 reduction by vanadium nitrogenase. Biochem J 256:429–432. https://doi.org/10.1042/bj2560429.

74. Shah P, Ding Y, Niemczyk M, Kudla G, Plotkin JB. 2013. Rate-limiting steps in yeast protein translation. Cell 153:1589–1601. https://doi.org/10.1016/j.cell.2013.05.049.

75. Kozak M. 1999. Initiation of translation in prokaryotes and eukaryotes. Gene 234:187–208. https://doi.org/10.1016/s0378-1119(99)00210-3.

76. Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI. 2001. Molecular mechanisms of translation initiation in eukaryotes. Proc Natl Acad Sci U S A 98:7029–7036. https://doi.org/10.1073/pnas.111145798.

77. Heiniger EK, Harwood CS. 2015. Posttranslational modification of a vanadium nitrogenase. Microbiologyopen 4:597–603. https://doi.org/10.1002/mbo3.265.

78. Ito S, Koyama N, Osanai T. 2019. Citrate synthase from Synechocystis is a distinct class of bacterial citrate synthase. Sci Rep 9:6038. https://doi.org/10.1038/s41598-019-42659-z.

79. Miller S, Lesk AM, Janin J, Chothia C. 1987. The accessible surface area and stability of oligomeric proteins. Nature 328:834–836. https://doi.org/10.1038/328834a0.

80. A Darst S. 2001. Bacterial RNA polymerase. Curr Opin Struct Biol 11:155–162. https://doi.org/10.1016/S0959-440X(00)00185-8.

81. Strader MB, Verberkmoes NC, Tabb DL, Connelly HM, Barton JW, Bruce BD, Pelletier DA, Davison BH, Hettich RL, Larimer FW, Hurst GB. 2004. Characterization of the 70S ribosome from Rhodopseudomonas palustris using an integrated "Top-Down" and "Bottom-Up" mass spectrometric approach. J Proteome Res 3:965–978. https://doi.org/10.1021/pr049940z.