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], and Rajib Saha [a #]

[a] Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska, 68508, USA.

Running Head: Metabolic and expression model of *R. palustris*

[#]Address correspondence to Rajib Saha, rsaha2@unl.edu

ABSTRACT

Rhodopseudomonas palustris CGA009 (*R. palustris*) is a gram negative purple non-sulfur bacteria that grows phototrophically or chemotrophically by fixing or catabolizing a wide array of substrates including lignin breakdown products (e.g., *p*-coumarate) for its carbon and 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 in anaerobic mode, this study, for the first time, 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. Since ME-models include 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 multi-omics perspective.

IMPORTANCE

In this work, we reconstructed the first ME-model for a purple non-sulfur 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 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 understanding on regulatory mechanism of biodegradable plastic production precursors, such as polyhydroxybutyrate (PHB).

KEYWORDS

R. palustris, ME-model, nitrogenase, rubisco, ferredoxin, electron distribution.

INTRODUCTION

R. palustris is an alphaproteobacterium which can grow in diverse metabolic modes such as phototrophic or chemotrophic growth. Besides, it can grow under aerobic or anaerobic conditions by using light and organic (e.g., lignin breakdown products) or inorganic compounds as a source of ATP generation (1,2). Using these metabolic versatilities, *R. palustris* has emerged as a potential biotechnological platform for bioremediation (3–5), bioplastics production (6,7),

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* (*A. 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 as a biotechnological chassis for further metabolic engineering (7). Because of these unique features, *R. palustris* has a highly connected metabolic network which requires a systems-level investigation for better understanding.

One widely accepted systems level investigation tool is the stoichiometric constrain-based M-model (19). Initial efforts of reconstructing M-models of purple non-sulfur 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 wider resolution to capture 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 lignin monomer degradation 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 lacking of quantitative characterization of macromolecular machinery synthesis (MMS) could be problematic and may lead to incorrect predictions of biological scenarios, such as inaccurate reaction flux and multiple equivalent cellular phenotypic states (25,26). These inaccuracies can lead to an erroneous understanding of 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 this is the metabolic-expression (ME) modeling approach. ME-model is a resource allocation based model that includes not only the stoichiometric metabolic reactions, but also quantitative MMS information (27). As input, ME-models require the conditions of a steady-state environment and can then output predictions for maximum growth rate, substrate uptake, byproduct secretion, metabolic fluxes, gene expression levels, and protein expression (27). 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 cellular composition and gene expression of *Thermotoga maritima* (*T. maritima*) (28), fermentation profile of *Clostridium ljungdahlii* (*C. ljungdahlii*) (29), overflow metabolism of *Saccharomyces cerevisiae* (*S. cerevisiae*) (30), and multi-scale phenotype, enzyme abundance, and acid stress of *Escherichia coli* (*E. coli*) (31–33). An ME-model for *R. palustris* can also be very useful in answering fundamental biological questions, such as growth profiling, isozyme expression prediction, regulation on 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 observation from literature (34). Then acetate, succinate, butyrate, and *p*-coumarate were 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 approach demonstrated a bottom-up systems-biology approach that can be used to predict and analyze cellular physiology of *R. palustris*, thereby providing an opportunity to generate experimentally testable hypotheses.

RESULTS AND DISCUSSIONS

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 for the metabolic transformations. To reconstruct the ME-model, gene-protein-reaction (GPR) relationships for all the reactions, specially nitrogen fixation (catalyzed by Mo-, V-, and Fe-Nase) and carbon fixation (catalyzed by rubisco form I and form II) 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 are available. Reactions for which GPR associations are not available, it was assumed that an average bacterial enzyme with 31.09 *kDa* molecular weight (35) catalyzed each individual reaction. Overall, the ME-model contains 1398 reactions, 1483 metabolites, and 751 genes. FIG 1 demonstrates the workflow of the ME-model reconstruction.

In *R. palustris*, form I rubisco (L_8S_8) is comprised of eight large subunits (L_8) and eight small subunits (S_8) (36) and encoded by two genes, *rpa1559* and *rpa1560* (2). On the other hand, form

II rubisco (L_2) is comprised of two large subunits both encoded by *rpa4641* (2). Between the two forms of rubisco, form I has a higher molecular weight compared to form II (37,38) and therefore requires more carbon investment to synthesize. As rubisco is one of the most abundant enzymes in nature (39), the kinetics of this enzyme have been determined for multiple organisms (36,40). For different rubisco enzymes, it was shown that although form I has higher molecular weight and more carbon investment cost, form II has higher catalytic turnover rate (k_{cat}) per active site compared to form I (36). Evolutionary selection has played a major role in this counterintuitive observation (41,42). 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 (36). With increasing amounts of oxygen in earth's atmosphere, form I evolved with a much higher selectivity for carbon dioxide but with a lower k_{cat} (36). Since 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 measurements from other phylogenetically close (43) PNSB strains (*Rhodobacter capsulatus* (40) and *Rhodospirillum rubrum* (36), respectively).

For the three nitrogenase isozymes, each enzyme is encoded by a series of genes (2) (Mo-Nase by *rpa4602* - *rpa4633*, V-Nase by *rpa1370* - *rpa1380*, and Fe-Nase by *rpa1435* - *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 from *E. coli*, as discussed in literature (31) (see materials and methods section). These normalized k_{cat} values were used to define three independent nitrogen fixation reactions.

Metabolic and expression model of *R. palustris*

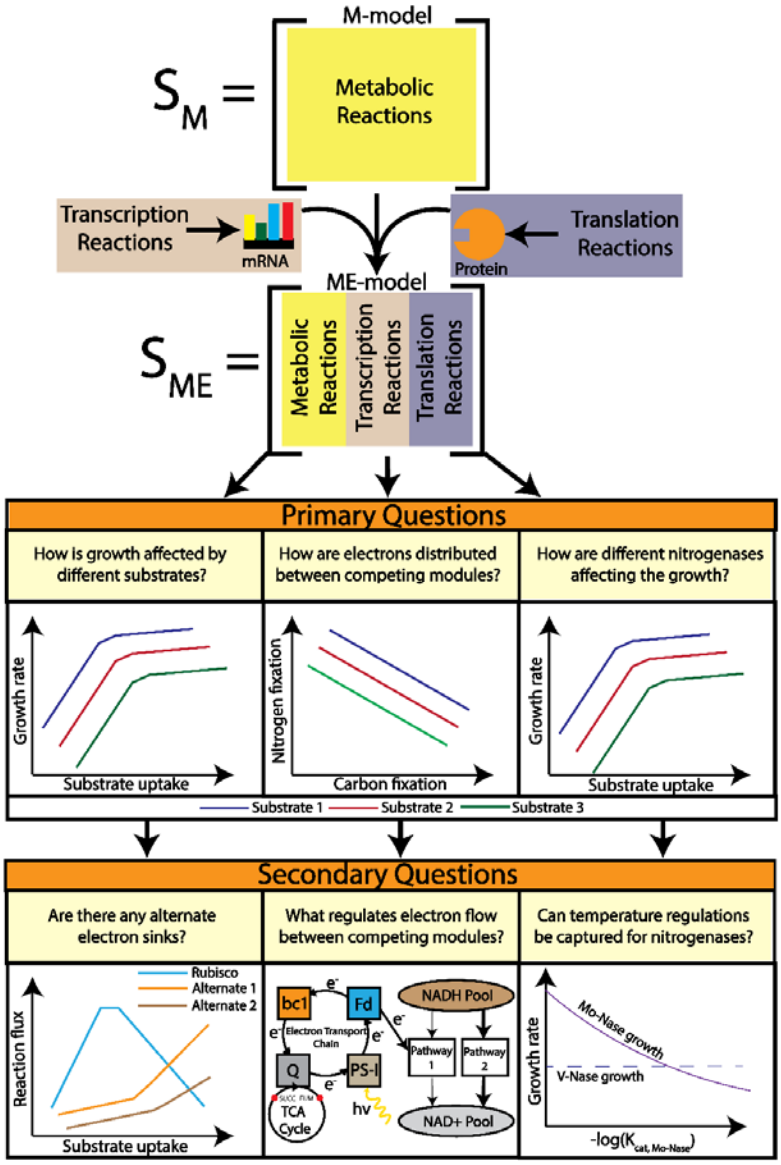


FIG 1. Workflow followed to reconstruct the ME-model from a previously published M-model of *R. palustris*. Transcription and translation reactions were added on top of the metabolic reactions to come up with ME-modeling framework. The ME-modeling framework was used to characterize growth rate profiling, competing metabolic modules, and nitrogenase enzyme activity. From these characterizations, inferences regarding alternate redox balancing, ferredoxin regulation, and temperature regulation of nitrogenase enzymes were gathered.

Both of the 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 (44). 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 literature (45,46). 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 a co-expression of both forms of rubisco during the photoheterotrophic growth of *R. palustris* (47). 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 (47). In addition, it was also found that expression of the *cbb* operons (responsible for coding both forms of rubisco) during phototrophic growth is highly dependent on the cellular carbon dioxide level (47). To incorporate these findings, a constraint was added to the ME-model to co-express both forms of rubisco based on the total carbon dioxide produced by *R. palustris* during photoheterotrophic growth (see materials and methods section).

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 sole carbon source, was used for the validation study (34). The study tested fold change of 4810 genes for *p*-coumarate catabolism considering succinate catabolism as the baseline condition. The transcriptomic analysis resulted

in 369 differentially expressed genes, among which 61 were metabolic genes. Similarly, proteomic analysis resulted in 341 differentially expressed proteins, among which 67 can act as enzymes. In both transcriptomics and proteomics data sets, non-metabolic genes/proteins have functions such as signaling, chromosomal replication, and circadian rhythm. (see supplemental material Table S1 for more details).

To generate both gene and protein expression information for the same two conditions of the above-mentioned study (34), 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 (44) that carbon fixation is required to maintain redox balance in *R. palustris*. Therefore, 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 61 metabolic genes and 67 metabolic enzymes, 23 genes and 34 enzymes were expressed in the model.

As part of the transcriptomics data validation, out of 23 genes, the ME-model was able to predict correct gene expression fold change for 21 genes. The model could not 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 (48). Biotin is a part of *R. palustris* cell membrane and from FIG 2 it can be seen that *p*-coumarate supports more growth than succinate. Thus, the ME-model predicted an upward fold change of *rpa3304* for *p*-coumarate catabolism compared to succinate catabolism. Composition of biotin in 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 (49). 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.

For proteomics data validation, out of 34 enzymes, the ME-model was able to correctly predict the fold change for 21 enzymes. The ME-model could not correctly predict the downward fold change of 13 different enzymes (see supplemental material Table S1 for more details). These 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. As *p*-coumarate supports more growth than succinate, the ME-model allocated more proteins for these pathways to sustain the 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. As ME-model maximizes the biomass growth rate, such incorrect prediction can be considered as an inherent weakness of the ME-model.

Overall, despite these incorrect fold change predictions, the ME-model was able to satisfactorily recapitulate the aggregate experimental transcriptomics and proteomics observations with 91% and 62% accuracy, respectively (see materials and methods section for accuracy calculation). The details of experimental and model predictions can be found in the supplemental material Table S1.

Growth rate vs. substrate uptake and alternate redox balancing strategies

Metabolic and expression model of *R. palustris*

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 rate. 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 substrates. Previous studies have shown that photoheterotrophic growth of *R. palustris* on acetate, succinate, and butyrate is associated with increasing cellular redox stress based on the oxidation state of different substrates (50). 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) (45), whereas butyrate (-1) and *p*-coumarate (-0.22) have lower oxidation states (45).

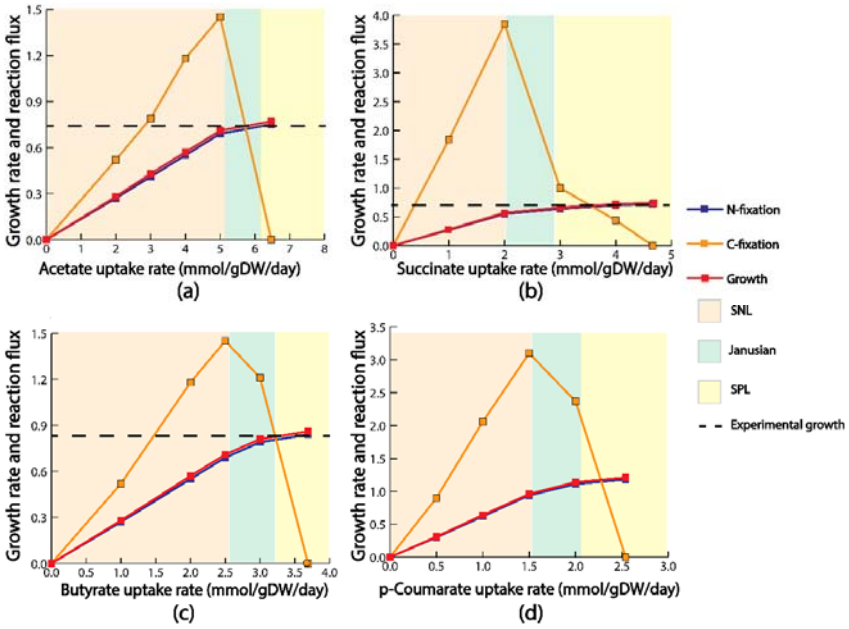


FIG 2 Strictly Nutrient-Limited (SNL), Janusian, and Strictly Proteome-Limited (SPL) regions for (a) acetate (b) succinate (c) butyrate and (d) *p*-coumarate. The growth rate with respect to different substrate uptakes

follows a non-linear pattern. Flux through nitrogen fixation reaction also follows the similar pattern to growth rate. Carbon fixation reached a peak in the Janusian region and then diminished in the theoretical maximal growth.

In the ME-model, 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 (51) and microbial slow growth kinetics (52). 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) (31). Growth in the SNL region depends heavily on nutrient uptake and adding more nutrient 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 nutrient excess condition) is limited by 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 (45) had characterized the growth of wild-type (WT) *R. palustris* for acetate, succinate, and butyrate, respectively, under nitrogen-fixing conditions. Table 2 compares between 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 condition, also showed a significantly higher growth on p-coumarate comparing to butyrate and acetate. It was previously reported (7) that, *p*-coumarate consumption lead to more ATP production compared to acetate, succinate, and butyrate and thus was able to support more growth.

Metabolic and expression model of *R. palustris*

Theoretical growth rates predicted by the ME-model were slightly higher compared to the experimental growth rates for all tested substrates (6% for succinate, 5% for butyrate, and 4% for acetate). It was expected as the cell has many more layers of physiological regulations, such as signaling pathways, allosteric regulation, and polymorphism, which were not captured in the ME-modeling framework. Overall, growth rate comparison between the ME-model prediction and experimental study reveals that, like *E. coli* (31), optimum resource allocation dictates metabolic activities for *R. palustris*. Supplemental material Table S2 records all the theoretical maximum growth rates for different amount of substrate uptakes.

Table 2: Normalized growth rate for different substrate uptakes.

Substrate	Experimentally observed growth rate day^{-1})	Growth rate from the ME-model (day^{-1})	Substrate Uptake for experimental 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	-	1.21	2.54

After characterizing the growth rate with different substrate uptakes, the ME-model was used to characterize nitrogen and carbon fixation rates as a function substrate uptake. For nitrogen fixation, the reaction's activity followed a similar trajectory as growth vs. substrate uptake (FIG 2). Different studies have shown that during WT photoheterotrophic growth, among three

different nitrogenase (Mo-, V-, and Fe-Nase) isozymes encoded in *R. palustris'* genome, Mo-Nase is exclusively expressed (45,46). *A. vinelandii* which has three different nitrogenases also exclusively express the Mo-Nase in the WT (53). The ME-model predicted exclusive expression of Mo-Nase during growth on all four carbon sources. Expression of nitrogenase may be dictated by its ATP requirements, as Mo-Nase requires the least amount of ATP among three nitrogenases. In addition, the temperature of the assay plays a role in the expression of different nitrogenases as discussed later.

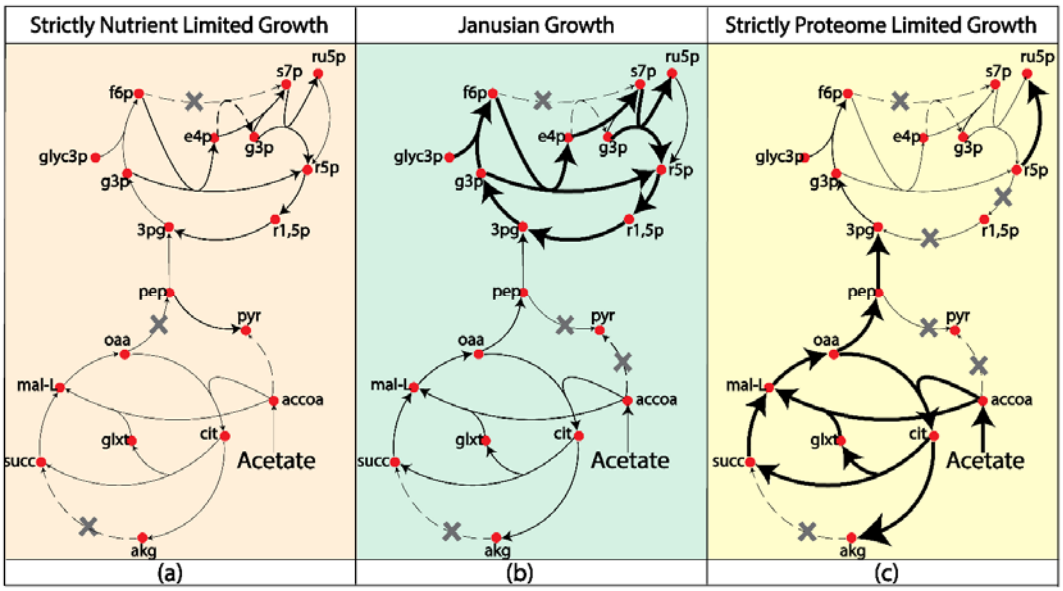


FIG 3. Metabolic activities in the (a) strictly nutrient limited growth (SNL), (b) Janusian growth, and (c) strictly proteome limited growth (SPL). In the theoretical maximum growth, at SPL region, flux through carbon fixation diminished and reaction flux from ribulose-5 phosphate to ribose-5 phosphate significantly increased. The increased biomass growth demand can be met by the precursors from the TCA cycle, which showed significant increase in reaction flux comparing to Janusian growth and SNL. Here gray crosses indicate zero rection flux through that reaction.

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 approaches the theoretical maximal growth, more ribose-5 phosphate is needed to sustain the increasing demand of nucleotides and lipopolysaccharides. To meet that demand at the theoretical maximum growth, the ME-model predicted that *R. palustris* decreases the expression of phosphoribulokinase (*rpa4645*) and redirects flux towards ribose-5 phosphate production (FIG 3).

During photoheterotrophic growth under nitrogen fixing condition, carbon and nitrogen fixation plays a major role in maintaining cellular redox balance. However, in the SPL region, as 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 is encoded by *rpa0192*. Similarly, glycerol-3 phosphate dehydrogenase uses NAD+/NADH as cofactors and is encoded by *rpa4410*. During the switch from the SNL to the SPL region, at the point where carbon fixation starts to diminish, both malate dehydrogenase and glycerol-3 phosphate dehydrogenase fluxes start 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* (54), and *Corynebacterium glutamicum* (*C. glutamicum*) (55). Glycerol-3 phosphate dehydrogenase is one of the key enzymes in the fatty acid biosynthesis. It was suggested that for photoheterotrophically grown *R. rubrum*, it is possible that other biosynthetic pathways such as fatty acid biosynthesis could offer flexibility contributing to the redox balance (56). In addition,

several other organisms such as *S. cerevisiae* (57) and *Kluyveromyces lactis* (*K. lactis*) (58) showed evidence of using glycerol-3 phosphate dehydrogenase to maintain redox balance.

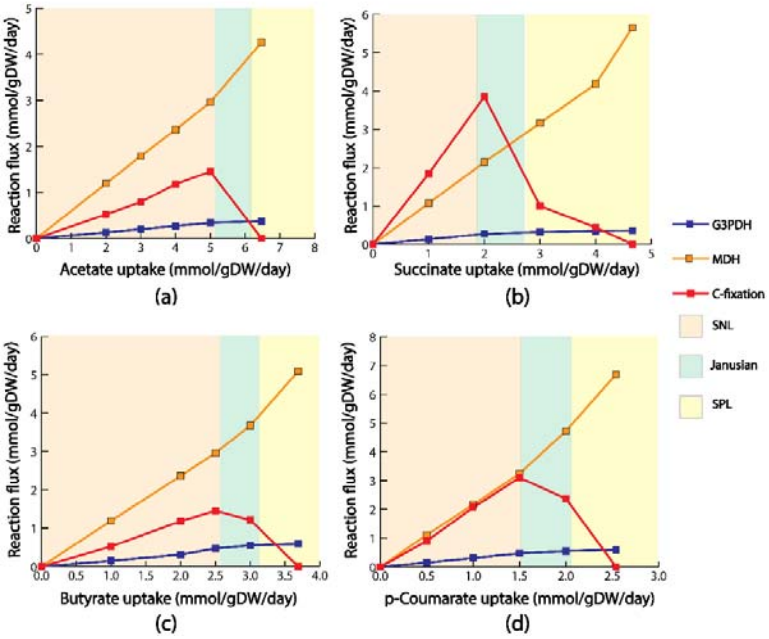


FIG 4 Alternate electron sink for different substrates (a) acetate (b) succinate (c) butyrate and (d). In the Janusian regions, flux through carbon fixation reaction started to diminish. With the diminishing carbon fixation flux, ME-model predicted two alternate electron, malate dehydrogenase and glycerol-3 phosphate dehydrogenase. Reaction flux through these alternate electron sinks reached its peak when flux through carbon fixation completely diminished at the theoretical maximum growth.

Carbon fixation vs. Nitrogen fixation – competing metabolic modules for redox balance

During photoheterotrophic growth, *R. palustris* performs a cyclic photophosphorylation (2,21) which means that electrons from photosystem I (PSI) get transported through ferredoxin and the bc_1 complex and recycled back to PSI through the oxidation and reduction of quinones (59) (FIG 5). As there are no terminal electron acceptors, this can cause an accumulation of reduced cofactors resulting in impeded growth of the bacterium. To resolve this, *R. palustris* employs various electron acceptors to maintain a cellular redox balance. During photoheterotrophic

growth, the redox-balancing mechanism consists primarily of the CBB pathway (44) and nitrogen fixation pathway (60). The nitrogen fixation module becomes active when *R. palustris* is placed in a nitrogen-limiting environment. Experimental 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* (44), *R. capsulatus* (61), *R. sphaeroides* (60,62), and *R. rubrum* (60). However, it is still not properly understood what factors decide the distribution of electrons in these two competing metabolic modules. 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.

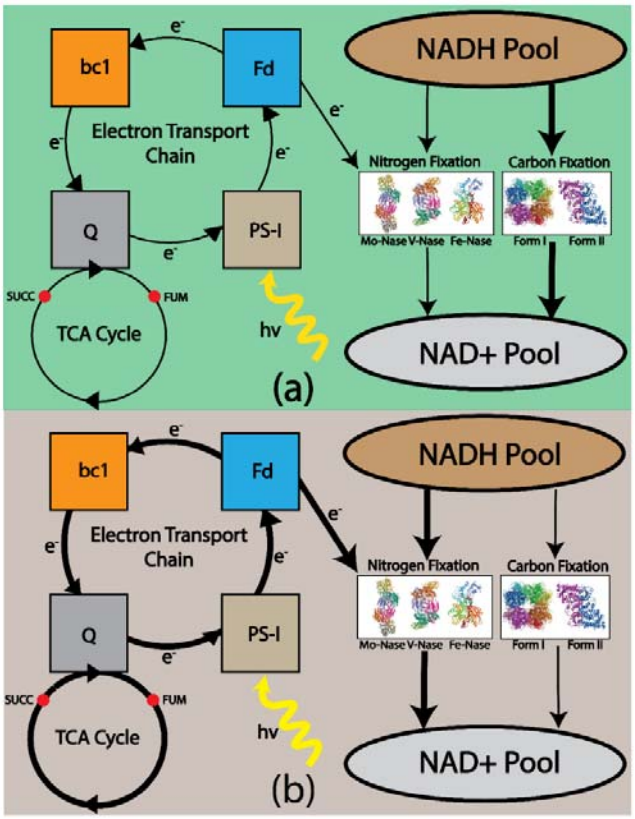
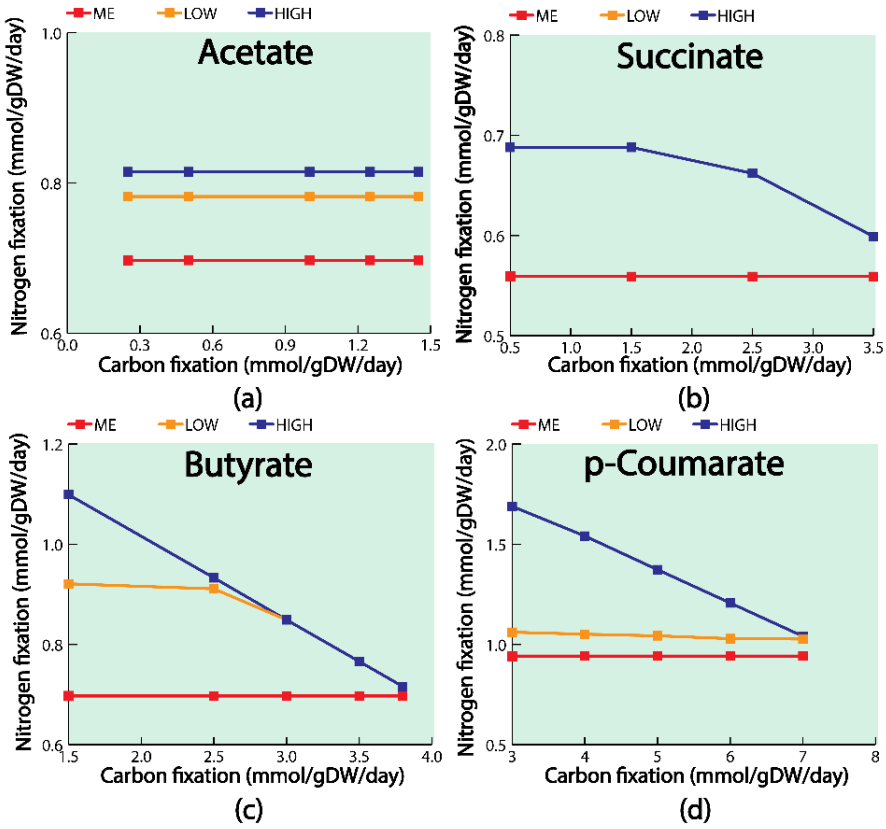


FIG 5 Relation between cyclic photophosphorylation and electron distribution between carbon and nitrogen fixation. (a) Less electron through ferredoxin indicates less flux through nitrogen fixation and more flux through carbon fixation pathway. As a result, NADH will be more oxidized through carbon fixation reaction (b) More

348 electron through ferredoxin indicates more flux through nitrogen fixation and less flux through carbon fixation
349 pathway. As a result NADH will be more oxidized through nitrogen fixation reaction.



350

351 **FIG 6 Relation between carbon fixation and nitrogen fixation with different fluxes via electron transport**
352 **through ferredoxin (ETFD) for different substrates (a) acetate (b) succinate (c) butyrate and (d) *p*-coumarate.**
353 Red color lines indicate the relation between carbon fixation and nitrogen fixation when flux through ETFD is set to
354 the solution found from the ME-model. Blue color lines indicate the relation between carbon fixation and nitrogen
355 fixation when ETFD flux is set to a very high value. Yellow color lines indicate the relation between carbon
356 fixation and nitrogen fixation when the ETFD flux values is set between ME and High.

357 To understand the electron distribution, a previous study eliminated rubisco activity in *R.*
358 *palustris* and found that the rate of nitrogen fixation did not vary significantly (44). As CBB and
359 nitrogen fixation pathways are two major redox balancing mechanisms, when rubisco was
360 eliminated, nitrogen fixation pathway was likely to carry additional flux load to maintain cellular

redox balance. As this was not the case in the previous experimental study (44), it was suspected that there exists a metabolic bottleneck preventing additional reaction flux through the nitrogen fixation pathway. FIG 5 shows the cyclic photophosphorylation of *R. palustris*. Here, electrons get transported in a cyclical manner and reduced ferredoxin supplies electrons to nitrogen fixation pathway. Based on the availability of electrons, 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 pathways. Hence, less reduced cofactors will be available for carbon fixation pathway to use. So, electron transport through ferredoxin (ETFD) can be a potential candidate of the previously discussed bottleneck.

In order to explore if ETFD is 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 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 growth on acetate. For the other carbon sources, 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 negative correlation was observed between the carbon and nitrogen fixation reaction flux. When the metabolite pool size (See supplemental material Text S1 for metabolite pool size calculation

detail) 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 prioritize electron transport to nitrogenase rather than rubisco, whose primary function is to maintain the redox balance in the cell. Thus the relation between carbon and nitrogen fixation is 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 carbon 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* (63). A highly organized network of overlapping transcriptional regulatory elements regulates 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, can have profound implications in future metabolic engineering efforts of *R. palustris*. Specially, this regulation can be exploited to increase hydrogen production from *R. palustris* to achieve energy sustainability goals.

Characterization of Mo-, V-, and Fe-Nase nitrogenase enzymes

Since ETFD was postulated to play a regulatory role in distributing electron to the nitrogen fixing pathway, the ME-model was next used to characterize how these electrons were used 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 vs. substrate uptake curve (FIG 7 a) followed the pattern identified from the literature (31). Exclusively expressing the Mo-Nase in the WT was also consistent with previous literature findings (45,46). Next, the growth vs. substrate uptake graphs (FIG 7 b, c, and d) were 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. When compared with the experimental growth rate data from literature (45) for WT, Mo-only, V-only, and Fe-only growths followed a similar pattern as predicted by the ME-model (supplemental material FIG. S1). Theoretically, growth of the WT and -only mutant strains of *R. palustris* can be coupled with the ATP requirement, as Mo-nase requires the least and Fe-nase requires the most amount of ATP for nitrogen fixation.

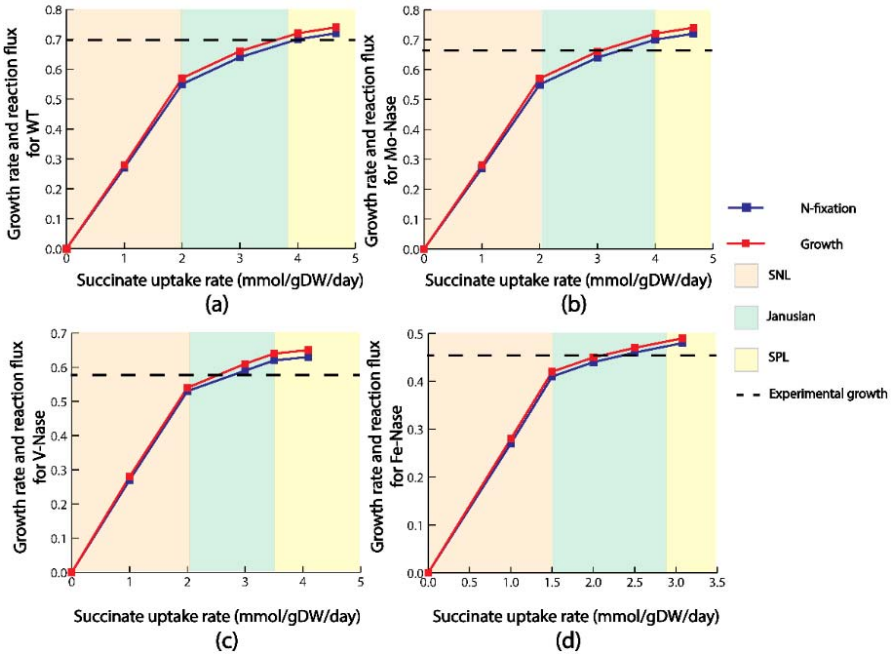


FIG 7 Growth rate and nitrogen fixation rate for (a) WT *R. palustris* (b) Mo-only mutant (c) V-only mutant (d) Fe-only mutant for succinate uptake. For each of the case, growth rate and nitrogen fixation closely follow each other. Dotted line in each of the graph indicates the experimentally observed growth.

Contrary to the pattern observed for succinate uptake, when other carbon sources were used as substrates, V-Nase exhibited higher growth comparing to Mo-Nase, Fe-Nase, and even WT (supplemental material FIG. S2-S4). Previous studies have observed that the Mo-Nase is more sensitive towards decreasing temperature compared to the other isozymes, such as V-Nase (45). FIG 8a qualitatively summarized this idea. Since the experimental values used in that study were generated at 19 ^0C, it is possible that Mo-Nase may have less selectivity towards fixing nitrogen rate than other substrates. The effect of decreasing assay temperature on the activity of nitrogenase is complex. It was reported (64) that for the Mo-Nase of *A. vinelandii*, the rate of nitrogen reduction at 10 ^0C is very low despite continued hydrolysis of ATP. In the case of Mo-Nase of *Klebsiella pneumoniae* (*K. pneumoniae*), decreasing the temperature not only curtails electron flux, but also results in the preferential loss of activity towards nitrogen as a substrate compared with H^+ or ethyne (C_2H_2) (65).

This modeling framework was further used to investigate the decreased growth rate of Mo-Nase at lower temperature. From Arrhenius equation (66) it is known that turnover rate of an enzyme, k_{cat}, increases exponentially with the increasing temperature. As 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 V-Nase growth rate exceeds that of Mo-Nase or Mo-Nase does the same compared to V-nase. At first, the k_{cat} of V-Nase was increased to a very high value, but the growth rate of V-Nase was still lower than the WT and Mo-Nase. It indicates that the sensitivity of V-Nase activity with respect to temperature is very low. This finding is consistent with previously

published work on another gram negative bacteria, *Azotobacter chroococcum* (67). Later, the k_{cat} of Mo-Nase was decreased to a very low value, and at that low k_{cat}, Mo-nase growth was actually lower than the V-Nase and higher than the Fe-Nase, which is similar to the finding from literature (45). Therefore, by tuning the k_{cat}, ME-model was able to capture the experimentally observed temperature sensitivity of different nitrogenase enzyme. FIG 8b qualitatively summarized the effect of k_{cat} on growth of Mo-only and V-only strains of *R. palustris*.

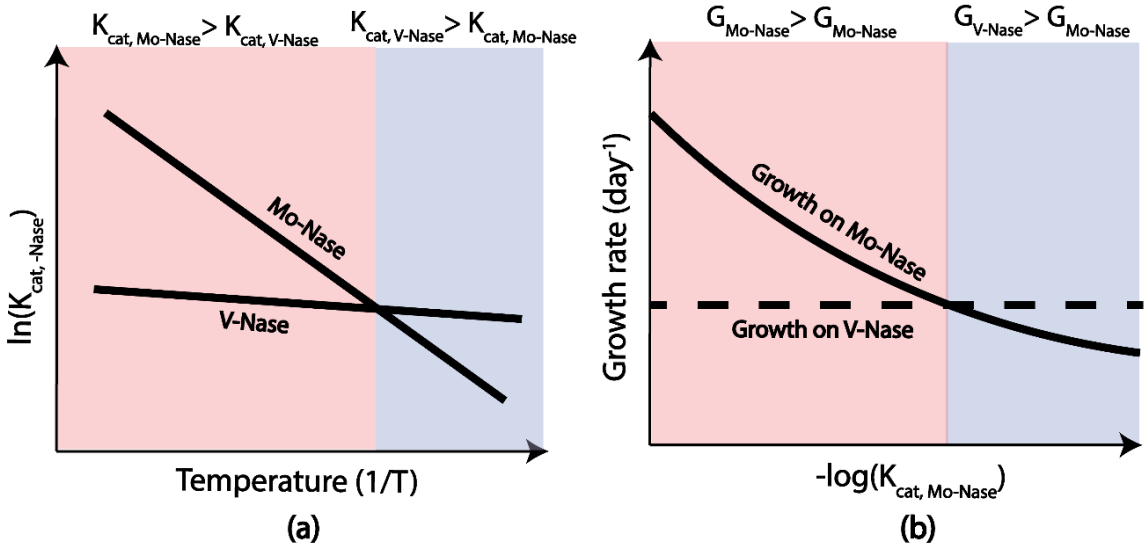


FIG 8 A qualitative representation of temperature regulation of Mo-Nase and V-Nase, and effect of k_{cat} on the growth of Mo-only and V-only mutant. (a) From literature, it is known that V-Nase has less sensitivity with respect to temperature comparing to the Mo-Nase. The prediction from this ME-model corroborates that finding. (b) As k_{cat} is a parameter which is a function of temperature, from Arrhenius equation, we know that with reducing temperature, k_{cat} also reduces. With reducing k_{cat}, at one stage, growth for Mo-only mutant falls below the growth of V-only mutant, capturing the experimentally observed temperature regulation of Mo- and V-Nase.

Conclusion. In this work, the first ever ME-model of *R. palustris* was developed. Growth rates predicted by the ME-model for different substrates closely matched 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, ME-modeling framework successfully captured experimentally observed temperature regulation of different nitrogenase enzymes.

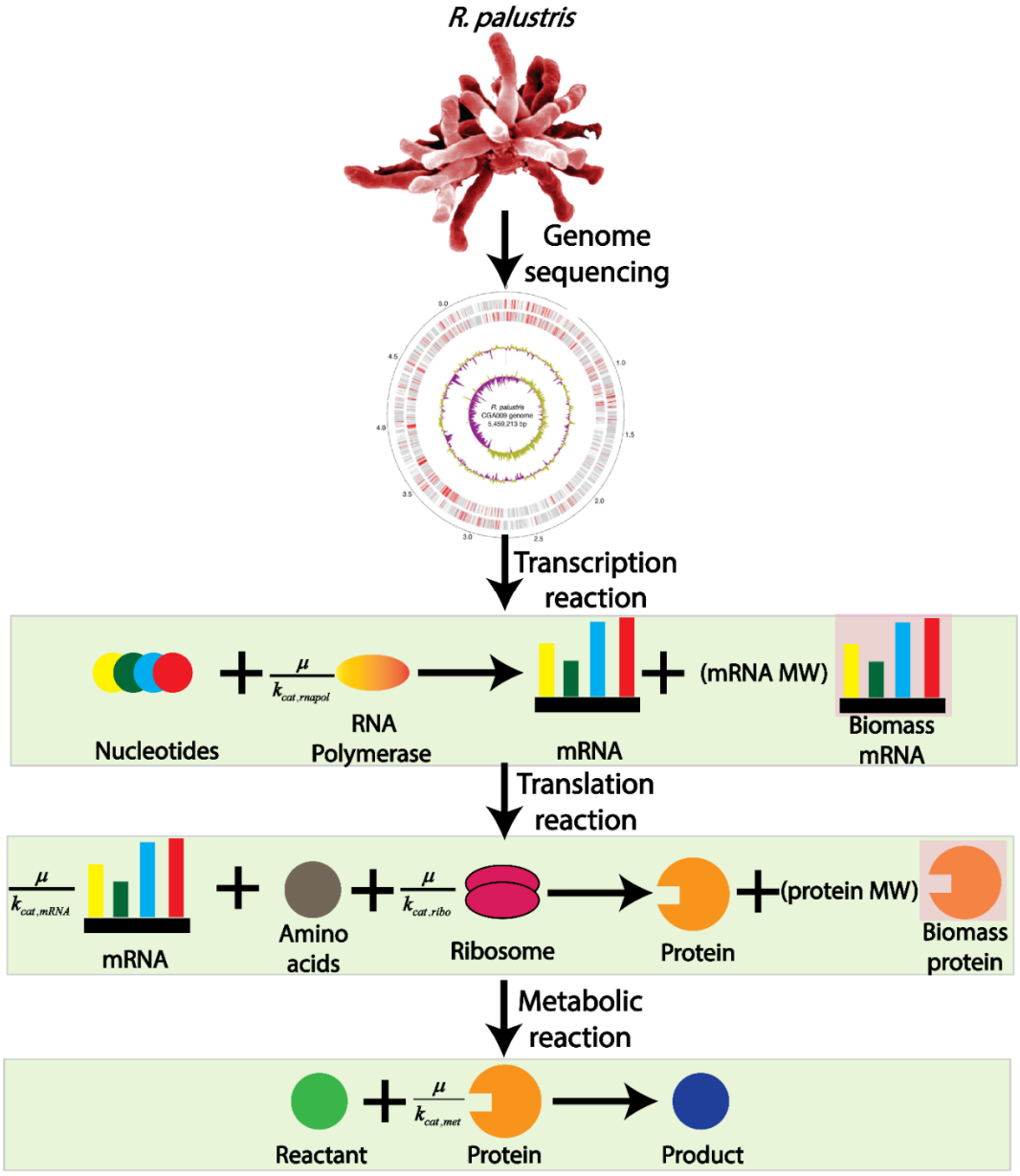
Going froward, this ME-model can be used as a powerful platform to further characterize different features of *R. palustris* metabolism. Specially characterizing a complete profile of environment specific isozyme expressions and optimal protein allocation. Furthermore, this ME-model can be used to design and fine-tune mutants of *R. palustris* for metabolic engineering purpose. One such application can be to produce PHB, a bioplastic precursor, which has 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 can be further used to gain similar regulatory insights, as discussed in this paper, on how electrons are distributed in PHB producing pathways when different LBPs are used as substrates.

MATERIALS AND METHODS:

ME-model of *R. palustris*:

In addition to the metabolic reactions from the M-model, ME-model consists of translation and transcription reactions along with metabolic reactions (FIG 9). In order to model transcription and translation reaction, GPR association of each reaction is required. The initial GPR

480 association was collected from literature (59). Later that GPR association was manually curated
481 using the detail genome annotation from literature (2).



482

483 **FIG 9** ***R. palustris*** **ME-model reconstruction.** In the M-model, only metabolic reactions are incorporated to
484 perform genome-scale metabolic modeling. However, in the ME-modeling framework, transcription and translation
485 process are also incorporated, adding two separate layer of regulation for metabolic reaction. Each layer of
486 regulation are coupled with the biomass growth through catalytic turnover rate and the biomass growth. This process

is known as coupling and the coupling parameter is in the form of $\frac{\mu}{k_{cat}}$, where μ indicates the growth rate and k_{cat} indicates the catalytic turnover rate for that.

This GPR association can be accessed in supplemental material Table 3. The overall ME-model reconstruction procedure was conducted in accordance to the COBRAme protocol (49) which is summarized in FIG 2. The ME-model is a multi-scale model; hence it requires the addition of coupling constraints to relate different cellular processes to each other. The coupling constraints are in the form of $\frac{\mu}{k_{cat}}$. Here μ is the growth rate and k_{cat} approximates the effective turnover rate for the different macromolecules. Detailed mathematical description for $\frac{\mu}{k_{cat}}$ of different macromolecular process and values of different parameters can be found in the supplemental material Text S1 and in the original COBRAme protocol (49).

To calculate k_{cat} for different enzymes, a mean k_{cat} value of 65 s^{-1} was used, which was reported for the *E. coli* in another ME-modeling framework (31). This mean k_{cat} was modified for each enzyme based on the solvent accessible surface area (SASA), following the same ME-modeling framework (31). SASA can be defined as is the surface area of an enzyme that is accessible to a solvent. Also, a previous study (68) reported a correlation between SASA and molecular weight of the enzyme as following:

$$SASA = (molecular\ weight\ of\ the\ enzyme)^{\frac{3}{4}} \quad \ldots (1)$$

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 ($SASA_{enzyme}$).

$$k_{cat,enzyme} = k_{cat,mean} \times \frac{SASA_{enzyme}}{SASA_{mean}} \quad \ldots (2)$$

Metabolic and expression model of *R. palustris*

For transcription reactions, RNA polymerase is needed to produce the required mRNA for the protein production. RNA polymerase of *R. palustris* consists of five subunits: two alpha (α) subunits, a beta (β) subunit, a beta prime subunit (β'), and a small omega (ω) subunit (69). In the model each individual subunits were synthesized to form the RNA polymerase. Later, these RNA polymerase transform different nucleotide to mRNA.

For translation reactions, ribosomal RNA is required to transform amino acids into different proteins. *R. palustris* utilizes 70S ribosomes, each consisting of a small (30S) and a large (50S) subunit (70). The large subunit is composed of a 5S RNA subunit (120 nucleotides), a 23S RNA subunit (2900 nucleotides), and 31 proteins. The small subunit is composed of a 16S RNA subunit (1542 nucleotides) and 21 proteins (70). It was also assumed that tRNA charging of amino acid to the ribosome was not a rate limiting process in the translation reaction. Hence no macromolecular synthesis of tRNA was included in the model.

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 as well as the biomass protein. Similarly, the transcription reaction produces mRNA required for the protein synthesis and also biomass mRNA requirements. The biomass protein and mRNA participate 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 are available. For remaining pathways, an enzyme was used with an average length of 283 amino acids and molecular weight of 31.09 *kDA*.

Metabolic and expression model of *R. palustris*

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

$$v_{rubisco\ I} = \left[-\sum v_{CO_2} + \frac{\mu}{\mu_{max}} \sum v_{CO_{2,max}}\right] \times \frac{k_{cat,rubisco\ I}}{k_{cat,rubisco\ II}} \ldots (3)$$

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

For each of the substrate, the total ATP production by the ME-model was capped according to the following equation proposed in the literature (7):

$$v^S_{PSII} = v^{ace}_{PSII} \frac{\emptyset^S_{PSII}}{\emptyset^{ace}_{PSII}} \ldots (4)$$

Here "S" and "ace" refer to different substrates and acetate, respectively. Also, $\emptyset^S_{PSII}$ and $\emptyset^{ace}_{PSII}$ refer to the photosynthetic yield of different substrates and acetate respectively. Photosynthetic yields for different substrates are collected from literature (7).

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 expression and protein was 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, that fold changes were compared with the literature (34). If both fold changes, from the ME-model and the experimental study, showed 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:

The General Algebraic Modeling System (GAMS) version 24.7.4 with IBM CPLEX solver was used to run pFBA algorithm on the model. The algorithm was scripted in GAMS and then run on a Linux-based high-performance cluster computing system at the University of Nebraska-Lincoln.

SUPPLEMENTARY MATERIALS

Supplemental material is available online only.

FIG S1-S4, PDF file, 0.01 MB.

TABLE S1, DOCX file, 1.7 MB

TABLE S2, DOCX file, 0.05 MB

TABLE S3, DOCX file, 0.2 MB

TEXT S1, DOCX file, 0.03 MB

DATA AVAILABILITY

Metabolic and expression model of *R. palustris*

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

ACKNOWLEDGEMENT

We gratefully acknowledge funding support from NSF Career Award grant number 1943310.

REFERENCES

1. VerBerkmoes NC, Shah MB, Lankford PK, Pelletier DA, Strader MB, Tabb DL, et al. Determination and Comparison of the Baseline Proteomes of the Versatile Microbe Rhodopseudomonas palustris under Its Major Metabolic States. J Proteome Res [Internet]. 2006 Feb 1;5(2):287–98. Available from: https://doi.org/10.1021/pr0503230

2. Larimer FW, Chain P, Hauser L, Lamerdin J, Malfatti S, Do L, et al. Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris. Nat Biotechnol [Internet]. 2004;22(1):55–61. Available from: https://doi.org/10.1038/nbt923

3. Sakpirom J, Kantachote D, Nunkaew T, Khan E. 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 [Internet]. 2017;168(3):266–75. Available from: https://www.sciencedirect.com/science/article/pii/S0923250816301528

4. Nookongbut P, Kantachote D, Megharaj M. 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 [Internet].

2016;66(4):1419–29. Available from: https://doi.org/10.1007/s13213-016-1229-z

5. Lu H, Yuan Y, Campbell DE, Qin P, Cui L. Integrated water quality, emergy and economic evaluation of three bioremediation treatment systems for eutrophic water. Ecol Eng [Internet]. 2014;69:244–54. Available from: https://www.sciencedirect.com/science/article/pii/S0925857414001906

6. Brown B, Immethun C, Wilkins M, Saha R. Rhodopseudomonas palustris CGA009 polyhydroxybutyrate production from a lignin aromatic and quantification via flow cytometry. Bioresour Technol Reports [Internet]. 2020;11:100474. Available from: https://www.sciencedirect.com/science/article/pii/S2589014X20300955

7. Alsiyabi A, Brown B, Immethun C, Long D, Wilkins M, Saha R. Synergistic experimental and computational approach identifies novel strategies for polyhydroxybutyrate overproduction. Metab Eng [Internet]. 2021;68:1–13. Available from: https://www.sciencedirect.com/science/article/pii/S1096717621001324

8. Wang X, Feng Y, Wang H, Qu Y, Yu Y, Ren N, et al. Bioaugmentation for Electricity Generation from Corn Stover Biomass Using Microbial Fuel Cells. Environ Sci Technol [Internet]. 2009 Aug 1;43(15):6088–93. Available from: https://doi.org/10.1021/es900391b

9. Sun J, Yang P, Li N, Zhao M, Zhang X, Zhang Y, et al. Extraction of photosynthetic electron from mixed photosynthetic consortium of bacteria and algae towards sustainable bioelectrical energy harvesting. Electrochim Acta [Internet]. 2020;336:135710. Available from: https://www.sciencedirect.com/science/article/pii/S001346862030102X

10. Wu P, Chen Z, Zhang Y, Wang Y, Zhu F, Cao B, et al. Rhodopseudomonas palustris

wastewater treatment: Cyhalofop-butyl removal, biochemicals production and mathematical model establishment. Bioresour Technol [Internet]. 2019;282:390–7. Available from: https://www.sciencedirect.com/science/article/pii/S0960852418316249

11. Kim MK, Choi K-M, Yin C-R, Lee K-Y, Im W-T, Lim JH, et al. Odorous swine wastewater treatment by purple non-sulfur bacteria, Rhodopseudomonas palustris, isolated from eutrophicated ponds. Biotechnol Lett [Internet]. 2004;26(10):819–22. Available from: https://doi.org/10.1023/B:BILE.0000025884.50198.67

12. Lu H, Zhang G, He S, Zhao R, Zhu D. Purple non-sulfur bacteria technology: a promising and potential approach for wastewater treatment and bioresources recovery. World J Microbiol Biotechnol [Internet]. 2021;37(9):161. Available from: https://doi.org/10.1007/s11274-021-03133-z

13. Mabutyana L, Pott RWM. Photo-fermentative hydrogen production by Rhodopseudomonas palustris CGA009 in the presence of inhibitory compounds. Int J Hydrogen Energy [Internet]. 2021;46(57):29088–99. Available from: https://www.sciencedirect.com/science/article/pii/S0360319920348527

14. du Toit J-P, Pott RWM. Heat-acclimatised strains of Rhodopseudomonas palustris reveal higher temperature optima with concomitantly enhanced biohydrogen production rates. Int J Hydrogen Energy [Internet]. 2021;46(21):11564–72. Available from: https://www.sciencedirect.com/science/article/pii/S0360319921001348

15. Ji Y, Sultan MA, Kim DY, Meeks N, Hastings JT, Bhattacharyya D. 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 [Internet]. 2021;46(11):7821–35. Available from: https://www.sciencedirect.com/science/article/pii/S0360319920345080

16. Gosse JL, Engel BJ, Hui JC-H, Harwood CS, Flickinger MC. Progress toward a biomimetic leaf: 4,000 h of hydrogen production by coating-stabilized nongrowing photosynthetic Rhodopseudomonas palustris. Biotechnol Prog [Internet]. 2010 Jul 1;26(4):907–18. Available from: https://doi.org/10.1002/btpr.406

17. Hongyuan L, Jiahua C, Yangyang J, Mingwei C, H. LPK, H. N. Transcriptomic Responses of the Interactions between Clostridium cellulovorans 743B and Rhodopseudomonas palustris CGA009 in a Cellulose-Grown Coculture for Enhanced Hydrogen Production. Appl Environ Microbiol [Internet]. 2016 Aug 1;82(15):4546–59. Available from: https://doi.org/10.1128/AEM.00789-16

18. C. SJ, Patricia dos S, S. GB, Helga E, Guadelupe E, M. RL, et al. Genome Sequence of Azotobacter vinelandii, an Obligate Aerobe Specialized To Support Diverse Anaerobic Metabolic Processes. J Bacteriol [Internet]. 2009 Jul 15;191(14):4534–45. Available from: https://doi.org/10.1128/JB.00504-09

19. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Nat Biotechnol. 2010;28(3):245–8.

20. Klamt S, Schuster S, Gilles ED. Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria. Biotechnol Bioeng [Internet]. 2002 Mar 30;77(7):734–51. Available from: https://doi.org/10.1002/bit.10153

21. Klamt S, Grammel H, Straube R, Ghosh R, Gilles ED. Modeling the electron transport

chain of purple non-sulfur bacteria. Mol Syst Biol [Internet]. 2008 Jan 1;4(1):156. Available from: https://doi.org/10.1038/msb4100191

22. Imam S, Yilmaz S, Sohmen U, Gorzalski AS, Reed JL, Noguera DR, et al. iRsp1095: A genome-scale reconstruction of the Rhodobacter sphaeroides metabolic network. BMC Syst Biol [Internet]. 2011;5(1):116. Available from: https://doi.org/10.1186/1752-0509-5-116

23. Navid A, Jiao Y, Wong SE, Pett-Ridge J. System-level analysis of metabolic trade-offs during anaerobic photoheterotrophic growth in Rhodopseudomonas palustris. BMC Bioinformatics [Internet]. 2019;20(1):233. Available from: https://doi.org/10.1186/s12859-019-2844-z

24. Alsiyabi A, Immethun CM, Saha R. Modeling the Interplay between Photosynthesis, CO2 Fixation, and the Quinone Pool in a Purple Non-Sulfur Bacterium. Sci Rep [Internet]. 2019;9(1):1–9. Available from: http://dx.doi.org/10.1038/s41598-019-49079-z

25. Reed JL, Palsson BØ. Genome-Scale In Silico Models of *E. coli* Have Multiple Equivalent Phenotypic States: Assessment of Correlated Reaction Subsets That Comprise Network States. Genome Res [Internet]. 2004 Dec 14;14(9):1797–805. Available from: http://genome.cshlp.org/lookup/doi/10.1101/gr.2546004

26. Schellenberger J, Lewis NE, Palsson BØ. Elimination of Thermodynamically Infeasible Loops in Steady-State Metabolic Models. Biophys J [Internet]. 2011;100(3):544–53. Available from: https://www.sciencedirect.com/science/article/pii/S0006349510052252

27. Alsiyabi A, Chowdhury NB, Long D, Saha R. Enhancing in silico strain design predictions through next generation metabolic modeling approaches. Biotechnol Adv

[Internet]. 2021 Oct 28;107806. Available from: https://www.sciencedirect.com/science/article/pii/S0734975021001129

28. Lerman JA, Hyduke DR, Latif H, Portnoy VA, Lewis NE, Orth JD, et al. In silico method for modelling metabolism and gene product expression at genome scale. Nat Commun [Internet]. 2012;3(1):929. Available from: https://doi.org/10.1038/ncomms1928

29. Liu JK, Lloyd C, Al-Bassam MM, Ebrahim A, Kim J-N, Olson C, et al. Predicting proteome allocation, overflow metabolism, and metal requirements in a model acetogen. PLOS Comput Biol [Internet]. 2019 Mar 7;15(3):e1006848. Available from: https://doi.org/10.1371/journal.pcbi.1006848

30. Oftadeh O, Salvy P, Masid M, Curvat M, Miskovic L, Hatzimanikatis V. A genome-scale metabolic model of Saccharomyces cerevisiae that integrates expression constraints and reaction thermodynamics. Nat Commun [Internet]. 2021;12(1):4790. Available from: https://doi.org/10.1038/s41467-021-25158-6

31. O'Brien EJ, Lerman JA, Chang RL, Hyduke DR, Palsson BØ. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. Mol Syst Biol [Internet]. 2013 Jan 1;9(1):693. Available from: https://doi.org/10.1038/msb.2013.52

32. Liu JK, O'Brien EJ, Lerman JA, Zengler K, Palsson BO, Feist AM. Reconstruction and modeling protein translocation and compartmentalization in Escherichia coli at the genome-scale. BMC Syst Biol [Internet]. 2014;8(1):110. Available from: https://doi.org/10.1186/s12918-014-0110-6

33. Du B, Yang L, Lloyd CJ, Fang X, Palsson BO. Genome-scale model of metabolism and gene expression provides a multi-scale description of acid stress responses in Escherichia

coli. PLOS Comput Biol [Internet]. 2019 Dec 6;15(12):e1007525. Available from: https://doi.org/10.1371/journal.pcbi.1007525

34. Pan C, Oda Y, Lankford PK, Zhang B, Samatova NF, Pelletier DA, et al. Characterization of Anaerobic Catabolism of p-Coumarate in Rhodopseudomonas palustris by Integrating Transcriptomics and Quantitative Proteomics *. Mol Cell Proteomics [Internet]. 2008 Oct 26;7(5):938–48. Available from: https://www.mcponline.org/article/S1535-9476(20)31181-6/abstract

35. Kozlowski LP. Proteome-pI: proteome isoelectric point database. Nucleic Acids Res [Internet]. 2017 Jan 4;45(D1):D1112–6. Available from: https://doi.org/10.1093/nar/gkw978

36. Davidi D, Shamshoum M, Guo Z, Bar-On YM, Prywes N, Oz A, et al. Highly active rubiscos discovered by systematic interrogation of natural sequence diversity. EMBO J [Internet]. 2020 Sep 15;39(18):e104081. Available from: https://doi.org/10.15252/embj.2019104081

37. Consortium TU. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res [Internet]. 2019 Jan 8;47(D1):D506–15. Available from: https://doi.org/10.1093/nar/gky1049

38. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res [Internet]. 2008 Jan 1;36(suppl_1):D480–4. Available from: https://doi.org/10.1093/nar/gkm882

39. Bar-On YM, Milo R. The global mass and average rate of rubisco. Proc Natl Acad Sci [Internet]. 2019 Mar 5;116(10):4738 LP – 4743. Available from:

http://www.pnas.org/content/116/10/4738.abstract

40. M. HK, Robert TF. The “Green” Form I Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase from the Nonsulfur Purple BacteriumRhodobacter capsulatus. J Bacteriol [Internet]. 1999 Jul 1;181(13):3935–41. Available from: https://doi.org/10.1128/JB.181.13.3935-3941.1999

41. Tabita FR, Satagopan S, Hanson TE, Kreel NE, Scott SS. 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 [Internet]. 2008 May 1;59(7):1515–24. Available from: https://doi.org/10.1093/jxb/erm361

42. Erb TJ, Zarzycki J. A short history of RubisCO: the rise and fall (?) of Nature’s predominant CO2 fixing enzyme. Curr Opin Biotechnol [Internet]. 2018;49:100–7. Available from: https://www.sciencedirect.com/science/article/pii/S095816691730099X

43. Higuchi-Takeuchi M, Morisaki K, Numata K. A Screening Method for the Isolation of Polyhydroxyalkanoate-Producing Purple Non-sulfur Photosynthetic Bacteria from Natural Seawater [Internet]. Vol. 7, Frontiers in Microbiology . 2016. p. 1509. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2016.01509

44. McKinlay JB, Harwood CS. Carbon dioxide fixation as a central redox cofactor recycling mechanism in bacteria. Proc Natl Acad Sci [Internet]. 2010 Jun 29;107(26):11669 LP – 11675. Available from: http://www.pnas.org/content/107/26/11669.abstract

45. Luxem KE, Kraepiel AML, Zhang L, Waldbauer JR, Zhang X. Carbon substrate re-orders relative growth of a bacterium using Mo-, V-, or Fe-nitrogenase for nitrogen fixation. Environ Microbiol [Internet]. 2020 Apr 1;22(4):1397–408. Available from:

https://doi.org/10.1111/1462-2920.14955

46. Yasuhiro O, K. SS, E. RF, Liyou W, Xiudan L, Tingfen Y, et al. Functional Genomic Analysis of Three Nitrogenase Isozymes in the Photosynthetic Bacterium Rhodopseudomonas palustris. J Bacteriol [Internet]. 2005 Nov 15;187(22):7784–94. Available from: https://doi.org/10.1128/JB.187.22.7784-7794.2005

47. S. JG, Simona R, C. VN, L. HR, Dale P, Robert TF. Differential Accumulation of Form I RubisCO in Rhodopseudomonas palustris CGA010 under Photoheterotrophic Growth Conditions with Reduced Carbon Sources. J Bacteriol [Internet]. 2009 Jul 1;191(13):4243–50. Available from: https://doi.org/10.1128/JB.01795-08

48. Zhe H, Huijuan D, Jin-Cheng M, Yonghong Y, Kai-Hui L, Qiao-Qiao G, et al. Novel Xanthomonas campestris Long-Chain-Specific 3-Oxoacyl-Acyl Carrier Protein Reductase Involved in Diffusible Signal Factor Synthesis. MBio [Internet]. 2022 Jan 31;9(3):e00596-18. Available from: https://doi.org/10.1128/mBio.00596-18

49. Lloyd CJ, Ebrahim A, Yang L, King ZA, Catoiu E, O'Brien EJ, et al. COBRAme: A computational framework for genome-scale models of metabolism and gene expression. PLOS Comput Biol [Internet]. 2018 Jul 5;14(7):e1006302. Available from: https://doi.org/10.1371/journal.pcbi.1006302

50. B. MJ, S. HC, K. ND. Calvin Cycle Flux, Pathway Constraints, and Substrate Oxidation State Together Determine the H2 Biofuel Yield in Photoheterotrophic Bacteria. MBio [Internet]. 2011 Oct 25;2(2):e00323-10. Available from: https://doi.org/10.1128/mBio.00323-10

51. Monod J. THE GROWTH OF BACTERIAL CULTURES. Annu Rev Microbiol

[Internet]. 1949 Oct 1;3(1):371–94. Available from: https://doi.org/10.1146/annurev.mi.03.100149.002103

52. Koch AL. Microbial physiology and ecology of slow growth. Microbiol Mol Biol Rev [Internet]. 1997 Sep 1;61(3):305–18. Available from: https://doi.org/10.1128/mmbr.61.3.305-318.1997

53. Jacobitz S, Bishop PE. Regulation of nitrogenase-2 in Azotobacter vinelandii by ammonium, molybdenum, and vanadium. J Bacteriol [Internet]. 1992 Jun 1;174(12):3884–8. Available from: https://doi.org/10.1128/jb.174.12.3884-3888.1992

54. Jantama K, Haupt MJ, Svoronos SA, Zhang X, Moore JC, Shanmugam KT, et al. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate. Biotechnol Bioeng [Internet]. 2008 Apr 1;99(5):1140–53. Available from: https://doi.org/10.1002/bit.21694

55. MITSUHASHI S, HAYASHI M, OHNISHI J, IKEDA M. Disruption of Malate:Quinone Oxidoreductase Increases L-Lysine Production by Corynebacterium glutamicum. Biosci Biotechnol Biochem [Internet]. 2006 Nov 2;70(11):2803–6. Available from: https://doi.org/10.1271/bbb.60298

56. Alexandra L. McCully, Maureen C. Onyeziri, Breah LaSarre, Jennifer R. Gliessman1 JBM. Reductive tricarboxylic acid cycle enzymes and reductive amino acid synthesis pathways contribute to electron balance in a Rhodospirillum rubrum Calvin-cycle mutant. Microbiology [Internet]. 2019;166(2):199–211. Available from: https://www.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.000877

57. Lopes M de B, -ur-REHMAN ATA, Gockowiak H, Heinrich AJ, Langridge P, Henschke

PA. Fermentation properties of a wine yeast over-expressing the Saccharomyces cerevisiae glycerol 3-phosphate dehydrogenase gene (GPD2). Aust J Grape Wine Res [Internet]. 2000 Nov 2;6(3):208–15. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1755-0238.2000.tb00181.x

58. Guadalupe-Medina V, Metz B, Oud B, van Der Graaf CM, Mans R, Pronk JT, et al. Evolutionary engineering of a glycerol-3-phosphate dehydrogenase-negative, acetate-reducing Saccharomyces cerevisiae strain enables anaerobic growth at high glucose concentrations. Microb Biotechnol [Internet]. 2014 Jan 1;7(1):44–53. Available from: https://doi.org/10.1111/1751-7915.12080

59. Alsiyabi A, Immethun CM, Saha R. Modeling the Interplay between Photosynthesis, CO2 Fixation, and the Quinone Pool in a Purple Non-Sulfur Bacterium. Sci Rep [Internet]. 2019;9(1):12638. Available from: https://doi.org/10.1038/s41598-019-49079-z

60. Joshi HM, Tabita FR. A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. Proc Natl Acad Sci [Internet]. 1996 Dec 10;93(25):14515 LP – 14520. Available from: http://www.pnas.org/content/93/25/14515.abstract

61. Tichi MA, Tabita FR. Maintenance and control of redox poise in Rhodobacter capsulatus strains deficient in the Calvin-Benson-Bassham pathway. Arch Microbiol [Internet]. 2000;174(5):322–33. Available from: https://doi.org/10.1007/s002030000209

62. Yilei Q, Robert TF. Expression of glnB and aglnB-Like Gene (glnK) in a Ribulose Bisphosphate Carboxylase/Oxygenase-Deficient Mutant of Rhodobacter sphaeroides. J Bacteriol [Internet]. 1998 Sep 1;180(17):4644–9. Available from:

https://doi.org/10.1128/JB.180.17.4644-4649.1998

63. Gunsalus RP. Control of electron flow in Escherichia coli: coordinated transcription of respiratory pathway genes. J Bacteriol [Internet]. 1992 Nov 1;174(22):7069–74. Available from: https://doi.org/10.1128/jb.174.22.7069-7074.1992

64. Watt GD, Burns A. Kinetics of dithionite ion utilization and ATP hydrolysis for reactions catalyzed by the nitrogenase complex from Azotobacter vinelandii. Biochemistry [Internet]. 1977 Jan 25;16(2):264–70. Available from: https://doi.org/10.1021/bi00621a017

65. Thorneley RNF, Eady RR. 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 [Internet]. 1977 Nov 1;167(2):457–61. Available from: https://doi.org/10.1042/bj1670457

66. Jensen F. Activation energies and the arrhenius equation. Qual Reliab Eng Int [Internet]. 1985 Jan 1;1(1):13–7. Available from: https://doi.org/10.1002/qre.4680010104

67. Miller RW, Eady RR. Molybdenum and vanadium nitrogenases of Azotobacter chroococcum. Low temperature favours N2 reduction by vanadium nitrogenase. Biochem J [Internet]. 1988 Oct 28;256(2):429–32. Available from: https://doi.org/10.1042/bj2560429

68. Miller S, Lesk AM, Janin J, Chothia C. The accessible surface area and stability of oligomeric proteins. Nature [Internet]. 1987;328(6133):834–6. Available from: https://doi.org/10.1038/328834a0

829 69. A Darst S. Bacterial RNA polymerase. Curr Opin Struct Biol [Internet]. 2001;11(2):155–
830 62. Available from:
831 https://www.sciencedirect.com/science/article/pii/S0959440X00001858

832 70. Strader MB, VerBerkmoes NC, Tabb DL, Connelly HM, Barton JW, Bruce BD, et al.
833 Characterization of the 70S Ribosome from Rhodopseudomonas palustris Using an
834 Integrated “Top-Down” and “Bottom-Up” Mass Spectrometric Approach. J Proteome Res
835 [Internet]. 2004 Oct 1;3(5):965–78. Available from: https://doi.org/10.1021/pr049940z

836