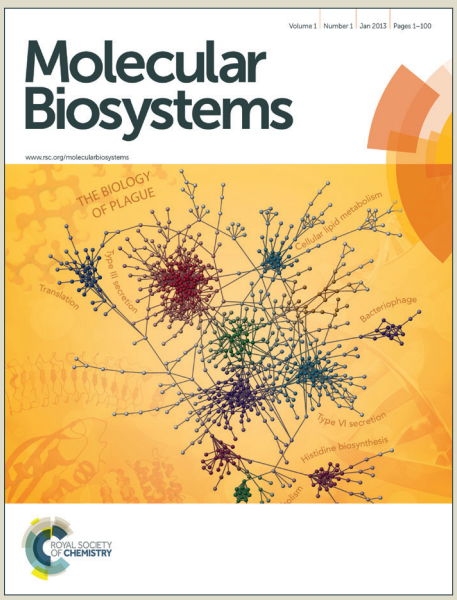
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Molecular BioSystems RSCPublishing

ARTICLE

Genome-scale reconstruction of the metabolic network in *Pseudomonas stutzeri* A1501

Parizad Babaei [a], Sayed-Amir Marashi [a,b], Sedigheh Asad [a,b]

Abstract

Pseudomonas stutzeri A1501 is an endophytic bacterium capable of nitrogen fixation. This strain has been isolated from rice rhizosphere and provides the plant with fixed nitrogen and phytohormones. These interesting features encouraged us to study the metabolism of this microorganism at the systems-level. In this work, we present the first genome-scale metabolic model (*i*PB890) for *P. stutzeri*, involving 890 genes, 1135 reactions, and 813 metabolites. A combination of automatic and manual approaches was used in the reconstruction process. Briefly, using the metabolic networks of *Pseudomonas aeruginosa* and *Pseudomonas putida* as templates, a draft metabolic network of *P. stutzeri* was reconstructed. Then, the draft network was driven through an iterative and curative process of gap filling. In the next step, the model was evaluated using different experimental data such as specific growth rate, Biolog substrate utilization data and other experimental observations. In most of the evaluation cases, the model was successful in correctly predicting the cellular phenotypes. Thus, we posit that the *i*PB890 model serves as a suitable platform to explore the metabolism of *P. stutzeri*.

Molecular BioSystems RSCPublishing

ARTICLE

1. Introduction

1.1. Genome-scale metabolic networks

Recent advances in sequencing techniques and computational tools for genome annotation have led to better understanding of the functions encoded by genomic sequences [1, 2]. It is now possible to automatically extract the information required for metabolic network reconstruction from genomic sequences [3, 4]. During the last decades, a wealth of data from well-controlled biochemical experiments has been provided in scientific literature and collected in databases like KEGG [5], MetaCyc [6] and metaTIGER [7]. Combining the genomic data with the biochemical knowledge leads to the emergence of genome-scale metabolic network models for a wide range of organisms, from bacteria [8-16] to archaea [17-19] to eukaryotic cells [20-26]. Such metabolic network models provide us with the opportunity to explore the physiological properties of different organisms in the context of their environmental and genetic constraints. More specifically, these models proved to be particularly useful for biotechnological applications [27-29].

1.2. *Pseudomonas stutzeri* and its metabolism

Pseudomonas stutzeri is a gram-negative bacterium capable of living in a wide spectrum of natural environments and occupying diverse ecological niches [30] and may even act as an opportunistic pathogen [31, 32]. Many metabolic capabilities such as degradation of aromatic compounds [33-37], denitrification , and nitrogen fixation [38] have been identified in different strains of this species. Various strains of *P. stutzeri* have been used in biotechnological processes such as biocatalysis production of pyruvate[39] and 2-oxobutyrate[40, 41], water quality control[42] and degradation of soil cyanide[43]. The entire genome sequence of *Pseudomonas stutzeri* A1501 has been published in 2008 [44]. This strain is best known for its role in nitrogen-fixation during symbiosis with plants. The strain was originally isolated in China from the rice paddy rhizosphere and can be used as an inoculant for this plant [45, 46]. The flexible metabolism of this microorganism [30] renders the systemic study of *Pseudomonas stutzeri* crucial to the understanding of the versatile nature and potential biotechnological use of this resourceful bacterium.

One typical systems biology framework for analyzing cell metabolism is the constraint-based modeling. This framework involves the application of stoichiometric and reversibility constraints on reactions [47]. Until now, very few studies have focused on constraint-based reconstruction and analysis of metabolism in agriculturally important bacteria (as plant symbionts [48, 49] or pathogens [50]). In this paper, we present a genome-scale metabolic network model, called *i*PB890, for *Pseudomonas stutzeri* A1501. Then, by using experimental data, we show that the model can correctly predict the metabolic capabilities of this strain. Thus, we posit that the *i*PB890 model serves as a suitable platform to explore the metabolism of *P. stutzeri*.

2. Experimental

2.1. Metabolic network reconstruction

The process of metabolic network reconstruction consists of four main stages. In the first stage, a draft network is created, using genome annotation and available metabolic information from biochemical databases and scientific literature. This stage can be done manually following the well-established protocols [51], or by using different available semi-automated methods. In this study, we used the RAVEN toolbox [52] in order to generate the draft model. This toolbox is able to use previously published models as templates. Briefly, in the first stage, a bi-directional BLAST [53] is run using protein FASTA sequences of the template organism(s) and the target organism. The "getBlast" function of this toolbox was used for this purpose. Then, based on protein homology, a draft metabolic network is reconstructed. The function "GetModelFromHomology" generates the draft network based on the results of the "getBlast" function.

In this work, the draft metabolic network was reconstructed using two *Pseudomonas* metabolic network models as templates, namely *i*MO1086 [11, 54] for *Pseudomonas aeruginosa* PAO1, and *i*JP962 [12, 54] for *Pseudomonas putida* KT2440. The process is schematically represented in Figure 1. These organism-specific models are reliable sources of information, since they have previously undergone manual refinements during the process of reconstruction and validation against different databases, literature and experimental data [54]. Moreover, we have previously shown that these two networks are successful at predicting the experimentally-observed biochemical capabilities of these two species [55]. Template-based model reconstruction is particularly useful when the template organism(s) and the organism of interest are phylogenetically close because many metabolic reactions are shared between them.

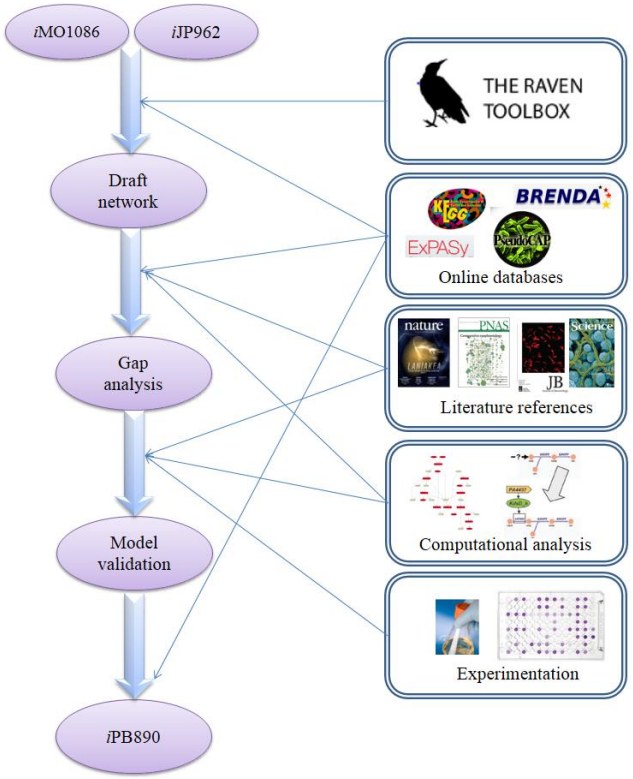


Figure 1. Schematic representation of the reconstruction process used for the metabolic network model of *Pseudomonas stutzeri* A1501.

In the second stage of reconstruction, refinement of the draft model was done in an iterative process. More precisely, we added the reactions in two steps:

Firstly, we tried to add reactions from the template models to the draft network. These reactions were either essential non-gene-associated reactions or associated to gene(s) which are not found in the bidirectional BLASTp search. At the end of this step, the model was functional, i.e., it was able to carry nonzero flux through its biomass producing reaction. The biomass producing reaction is an unreal "auxiliary" reaction added to the network to enable the simulation of cell growth [56]. The reactants of this reaction are the building blocks of biomass (e.g. nucleotides, amino acids, lipids, and cofactors).

In the next step in the second stage, from KEGG database, we downloaded the metabolic reactions of *P. stutzeri* A1501. In order to improve the network and fill in the gaps, in an iterative manual way, we added a set of reactions to the model such that they remain connected to the network (see below).

In this study, we used COBRA toolbox [47, 57] in order to detect gaps in the draft network and then we used biochemical reactions in order to fill the gaps. We used both "detectDeadEnds" and "GapFind" functions of the COBRA toolbox to find the existing gaps in the draft network. Then, we tried to add reactions in order to fill the gaps. The reactions which are used for gap filling were either taken from the template models according to the manual BLAST or from databases like KEGG and ExPASy [58].

In the third stage of reconstruction, we used the "SBMLFromExcel" function from the RAVEN toolbox, which converts excel files into the SBML format. In this stage, the "stoichiometric matrix" is formulated by converting the list of reactions into a machine-readable format. This stage can be done automatically.

The fourth stage of reconstruction is the evaluation of the model. In this stage, the accuracy of the model is measured by comparing the results of simulations (e.g., the growth phenotype) to experimental data. In this stage, the inconsistencies of the model with the biological data are identified. Then, the inconsistencies should be minimized through a series of manual refinements (Stage 2). In the present work, we used different experimental data such as aerobic and anaerobic growth rates and carbon consumption profile to evaluate and validate our model.

2.2. Flux balance analysis

Flux balance analysis (FBA) is a computational method based on linear programming for calculating the flux distribution of a metabolic network under specified conditions [59]. FBA optimizes an objective function (typically, biomass production rate, $v_{biomass}$, or sometimes ATP synthetase flux, $v_{ATP\text{-}synthetase}$) based on the steady-state assumption. In steady-state conditions, it is assumed that the concentrations of the cellular metabolites do not vary during the analysis. For a flux distribution vector **v**, this assumption is applied by the constraint $\mathbf{S}.\mathbf{v}=\mathbf{0}$, whereas S is the stoichiometric matrix and **0** is the zero vector, representing the vector of concentration changes. Furthermore, each flux distribution **v** is constrained by the "capacity" of reactions in the form $\mathbf{a}\leq\mathbf{v}\leq\mathbf{b}$. For example, an irreversible reactions *i* can only carry non-negative flux values, i.e., $0\leq v_i$.

As mentioned, in FBA, flux through a biomass-producing reaction is maximized, based on the evolutionary assumption of the cellular tendency to maximize its growth. The biomass reaction combines the substrates required for cell growth with appropriate stoichiometric coefficients. These coefficients are determined based on the ratios of the components forming the cellular dry weight. In the present work, the "optimizeCbModel" function of the COBRA toolbox was used to solve FBA problems using glpk solver version 4.47.

2.3. Flux variability analysis

The result of FBA is a flux distribution vector, **v**, for which the objective function is maximized. However, for real-world metabolic networks, there are often multiple optimal solutions. In other words, different flux distribution vectors may maximize the objective function, due to existence of alternative pathways. Flux variability analysis (FVA) is a method for determining the theoretical maximum and minimum values of all reaction fluxes when the objective function has its optimal value and the fluxes satisfy the same constraints as in FBA [60]. It is also possible to consider the idea of analyzing flux variability at suboptimal growth

rates, which is sometimes referred to as expanded flux variability analysis [61]. FVA can be used to determine blocked reactions, i.e., those reactions which cannot carry any nonzero flux under steady-state conditions. We used the "fluxVariability" function of the COBRA toolbox in order to perform this analysis.

2.4. Biomass formulation

In FBA, flux through the biomass producing reaction is typically considered as the objective function of the model. The reactants of this reaction are the major components of cellular dry matter. Accurate experimental measurement of the biomass components is a time-consuming and laborious task. Therefore, in the absence of species-specific biomass composition data, the biomass composition of *E. coli* (or, a closely related species) is commonly used for modeling. Since growth yield is not highly dependent on the biomass composition, such an assumption is reasonable [12, 62, 63].

In this work, we used the biomass composition of *P. aeruginosa* with some adjustments. Firstly, the amount of putrescine and spermidine in 1 gram of cellular dry weight was reported for *P. stutzeri* [30]. Secondly, we used high-performance liquid chromatography (HPLC) to measure the amino acid content of the total protein in *P. stutzeri*, and the results of this experiment were integrated in the biomass producing reaction. Here, we briefly mention the methodology used to determine the amino acid composition.

2.5. Determination of amino acid content using HPLC

P. stutzeri cells were cultivated in LB medium for 24 hours at 37 °C and 150 rpm. Then, the medium was centrifuged at 10000 rpm for 15 minutes. The cellular pellet was washed with NaCl solution (8%) twice. The total bacterial protein was extracted using TRIzol reagent [64] by following the manufacturer's protocol. The extracted protein was hydrolyzed using acidic and alkaline hydrolysis protocols explained elsewhere [65, 66]. Standard solutions were obtained by dissolving amino acids in water/methanol. The standard solutions and the sample were dried in warm air and derivatized according to the method of Bidlingmeyer [67]. 100 µl of derivatization reagent (methanol: water: TEA: PITC 7:1:1:1) was added to the tube sample and the mixture was kept for 30 minutes at room temperature before drying. The derivatized sample was dissolved in acetate buffer (pH 7.1) and 10 µl of the sample was injected to an HPLC C8 column (Eruspher, 250×4.6 mm, 5 µm, 100 Å). For analytical separation of PITC-amino acids, a binary mobile phase was used. The flow rate was set to 0.7 ml/min. Detection was carried out by a UV detector using the wavelengths 254 nm and 210 nm. The EZChrom software was used for chromatography data analysis.

2.6. Growth media and chemicals

For growth measurement experiments, the M9 medium was used. This growth medium is chemically defined, and therefore, bacterial growth in this medium can be simulated *in silico*. The M9 medium includes "basic" and "trace" elements. Na_2HPO_4, KH_2PO_4, NH_4Cl, and NaCl are added to the M9 medium as the basic components. Trace elements of the M9 medium consists of $MnCl_2$, $ZnCl_2$, $CuCl_2$, $CoCl_2$, and Na_2MoO_4. Other components include $CaCl_2$, $MgSO_4$, and $FeCl_3$. Glucose was added to the medium as the carbon source.

In order to measure the growth rate of *P. stutzeri*, 20 mL of M9 medium was inoculated with bacteria to obtain a pre-culture. The pre-culture was incubated overnight in a shaker-incubator at 37 °C and 150 rpm. 1.25 mL of the pre-culture was transferred to 50 mL of the fresh medium in a 250-mL Erlenmeyer flask and incubated. Turbidity of this inoculated medium was measured every 30 minutes using spectrophotometer at 600 nm. Each read was

repeated three times and the average of the three measurements was used to calculate the growth rate.

2.7. Phenotyping data

P. stutzeri A1501 was tested for its ability to grow on different carbon sources using Biolog GN2 MicroPlate™. These MicroPlates have 96 wells and each well consists of tetrazolium dye and a minimal growth medium with a particular carbon source. If the strain is able to grow on the defined carbon source, tetrazolium dye will be reduced to produce formazan, which has a purple color. Therefore, the production of formazan when the bacterial cells are inoculated to a minimal medium with only one source of carbon indicates the capability of the microorganism to respire actively while using that specific carbon source.

For carbon source utilization studies, all procedures were performed as indicated by the manufacturer. A pure culture of a bacterium was grown overnight at 37 °C on a Biolog Universal Growth with 5% Sheep Blood agar plate. The bacteria were swabbed from the surface of the agar plate, and suspended in GN/GP Inoculating Fluid to a specified density. 150 µl of bacterial suspension was pipetted into each well and the plate was incubated overnight at 37 °C. Subsequently, the MicroPlate was read with a MicroPlate reader and the results were analyzed according to the cut-offs provided in the Biolog protocols.

2.8. Further evaluation of the model using experimental data from literature

In order to further evaluate the predictive power of the model, we performed a comprehensive literature search, looking for articles containing experimental observations on the *P. stutzeri* metabolic capabilities. More specifically, we chose those articles concerning (i) growth rate measurements under different environmental conditions, (ii) genetic engineering and gene knockouts, and (iii) cellular response to environmental perturbations.

In general changes in the environmental conditions were simulated by manipulating the upper/lower bounds of the exchange reactions in a way that they would reflect the aforementioned conditions. In case of gene deletion, the upper and the lower bounds of the associated reaction were set to zero. Then, by running FBA or FVA, we were able to compare the computational results with the experimental observations.

3. Results

3.1. Biomass equation

In this work, the biomass composition of *P. stutzeri* was assumed to be similar to the biomass composition of *P. aeruginosa* with some modifications. More specifically, the putrescine and spermidine content of the biomass[30] was updated, and additionally, the protein synthesis reaction of the model was updated according to the experimentally determined amino acid composition of *P. stutzeri*. For this purpose, the amino acid composition of *P. stutzeri* was analyzed by HPLC. Then, the stoichiometric coefficients of different amino acids were adjusted accordingly.

3.2. Genome-scale metabolic network of *Pseudomonas stutzeri* A1501

The goal of the present work was to generate a constraint-based genome-scale metabolic network model of *Pseudomonas stutzeri* A1501. The final reconstructed network accounts for 890 genes, 813 metabolites, and 1135 reactions. This network includes the necessary anabolic reactions required for producing biomass components.

Reconstruction of the draft metabolic network model was carried out using the RAVEN toolbox [52]. Two previously published models (i.e., metabolic networks of *P. aeruginosa* and *P. putida*)

were used as templates for draft reconstruction. Therefore, our model shares a part of its reactions with these two models. Briefly, in the first step, 761 reactions were added by the RAVEN toolbox based on the bi-directional BLASTp search for enzyme homologs. In the next step, we manually added 173 reactions from the template models so that the draft model would be functional. Then, in an iterative manual procedure, 188 reactions from KEGG database were added to expand the model. More specifically, from KEGG, we downloaded the metabolic information of *P. stutzeri* A1501 and tried to add reactions as long as they were connected to the network, i.e., their reactants/products were already present in the model or they could be connected using other reactions in the downloaded data. Finally, 12 more reactions were added manually to the model for simulating the experimental data.

The number of shared reactions, as well as the number of unique reactions, is depicted in Figure 2. Obviously, our model includes a large number of reactions (i.e., 200 reactions) which are unique to *P. stutzeri* and are not reported in the previously published models of *P. aeruginosa* and *P. putida*. Among these 200 reactions, one can observe that reactions of lipid metabolism, energy metabolism and xenobiotics' metabolism are over-represented (see Supplementary file S1), which represent the unique features of the metabolism of *P. stutzeri*.

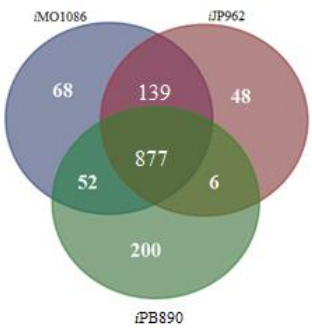


Figure 2. The number of shared and unique reactions in metabolic network models of *P. aeruginosa*, *P. putida* and *P. stutzeri* (*i*MO1086, *i*JP962 and *i*PB890, respectively).

In the next step, this draft network was manually curated by filling the network gaps as much as possible. The resulting model was able to simulate cell growth. In other words, it was able to carry flux through its biomass producing reaction. From this point on, we tried to further expand the model so that it covers a wider range of the metabolic reactions and pathways. In the next step, the model was converted to the standard SBML format (See Supplementary file S2). Finally, the *in silico* predictions of the model were compared with experimental data.

3.2.1. Blocked and unblocked reactions in *i*PB890

When simulating growth in the M9 minimal medium, 516 reactions were found to be blocked (Figure 3). This number remained the same even when suboptimal growth conditions were considered (as small as 80% of the original biomass production rate). One reason for the high number of blocked reactions is that the medium was considered to be minimal and only a limited set of exchange reactions were allowed to carry flux. As depicted in Figure 3, by simulating growth on the LB (rich) medium, the number of blocked reactions decreased to 482. Finally we set the lower bound of all the exchange reactions to -1000 $mmol.g_{DW}^{-1}.h^{-1}$ and performed FVA. In this situation, 400 reactions were found to be blocked. Therefore, the simulated medium can change the flux distribution in a metabolic network, and by choosing richer media a higher number of reactions will be able to carry metabolic flux. We also performed the same analysis on two template models. Figure 3 summarizes the results.

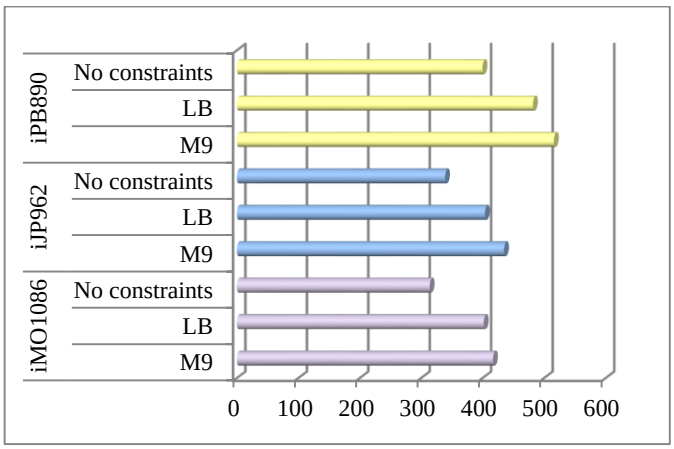


Figure 3. Number of blocked reactions during growth simulation on M9 (minimal) medium and rich (LB) medium. No constraints represents the situation in which the lower bound of all of the exchange reactions were set to -1000 mmol.g_{DW}^{-1}.h^{-1}.

3.3. Prediction of growth phenotypes

The growth rate of *Pseudomonas stutzeri* A1501 in the M9 medium (with glucose as the carbon source) was measured as described in the Materials and Methods section. Under those growth conditions, the growth rate was determined as 0.569 h^{-1} (Figure 4). On the other hand, the *in silico* growth rate of *P. stutzeri* on the M9 medium was determined using FBA. In order to simulate the M9 medium, the lower/upper bounds of the exchange reactions were set in a way that only the components of the medium (as described in the Materials and Methods section) were allowed to enter the system. The objective function of the FBA was considered to be the biomass production rate, $v_{biomass}$. The calculated *in silico* growth rate by FBA was 0.546 h^{-1}. (For more details on the *in silico* simulations, see Table S1 of the Supplementary file S3).

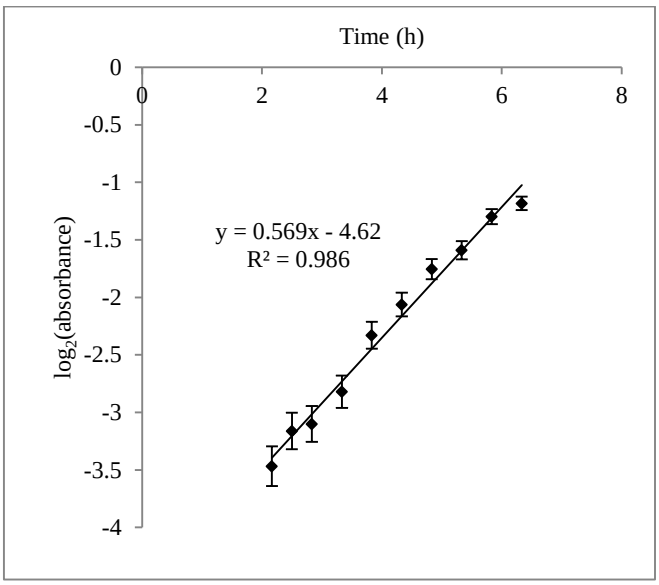


Figure 4. Experimental growth curve of *Pseudomonas stutzeri* A1501. Every read was repeated three times and the average of these three values were used for growth rate measurement. Vertical axis is $\log_2$ of absorbance at 600 nm and horizontal axis is time represented in hour. The experimentally measured growth rate of *P. stutzeri* was determined 0.569 h^{-1}.

3.4. Prediction of carbon source utilization phenotypes

P. stutzeri A1501 was tested for its ability to utilize various carbon sources using a high-throughput Biolog phenotypic assay. In this assay, a certain concentration of bacterial cells is inoculated in a 96-well Biolog MicroPlate™. Each well of the MicroPlate consists of tetrazolium dye and a minimal medium with a particular carbon source (except one well which only contains water as a negative control). If during incubation, the bacteria respire actively, the dye will turn into its reduced form and produce a visible purple color. The carbon consumption profile was determined based on the absorbance measurement of the produced color in each well. On the other hand, by computational growth simulation on each carbon source using FBA, we evaluated the capability of the network model for the utilization of carbon sources. Both $v_{biomass}$ and, $v_{ATP\text{-}synthetase}$ were tested as the objective function of FBA, but the results were found to be essentially the same.

Among the 95 different carbon sources available on the Biolog MicroPlate™, 25 had a corresponding transport reaction in our model. Therefore, these carbon sources were appropriate for the evaluation of the *in silico* predictions. The complete list of these compounds, together with their experimental and computational growth phenotypes, is presented in Table 1. The *in silico* growth simulations were performed by using FBA, as described in Materials and Methods. As shown in this table, in the majority of the cases (i.e., 20 out of 25), modeling results were consistent with experimental data, confirming the accuracy of our model.

Table 1. Experimental growth phenotype and computational growth simulation results for carbon sources whose transport reactions are accounted for in the model. Positive means "growth" and negative means "no growth" in both columns. In 22 cases of a total of 26, the model was able to successfully predict the growth phenotype.

Carbon source	Experimental results	*In silico* results	Consistent?
No carbon source	-	-	✓
L-Histidine	-	-	✓
D-Fructose	+	+	✓
L-Leucine	+	+	✓
D-Alanine	-	+	
L-Phenylalanine	-	-	✓
D-Glucose	+	+	✓
L-Alanine	+	+	✓
L-Proline	+	+	✓
Malonic Acid	+	+	✓
L-Aspartic Acid	+	+	✓
Glycerol	+	+	✓
L-Glutamic Acid	+	+	✓
L-Threonine	-	+	
D-Mannitol	-	+	
D, L-Carnitine	-	-	✓
γ-Aminobutyric Acid	+	+	✓
L-Ornithine	-	+	
Putrescine	+	+	✓
Succinic Acid	+	+	✓
Acetate	+	+	✓
Citrate	+	+	✓
D-Mannose	-	-	✓
L-Serine	-	+	
L-Asparagine	+	+	✓
2,3-Butanediol	-	-	✓

The remaining 70 compounds included 19 intracellular metabolites for which no transport reaction was included in the model and 51 metabolites in the Biolog MicroPlate™ which are not included in our metabolic network model (see Supplementary file S4).. We observed that for 38 out of these 51 carbon sources (~75%) no cellular respiration was detected, while in case of 13 carbon sources we observed some respiration activity. These 13 cases should be further investigated in the future in order to check their potential presence in the metabolism of *P. stutzeri*.

The growth phenotype of the intracellular metabolites of the model can be used for further improvement of the metabolic network model. For these metabolites, we simulated *in silico* growth by temporarily adding a hypothetical transport reaction and analyzing the results. If by adding such a reaction, the predictive power of the model increased, we kept the transport reaction in the model. This was the case when bacteria grew on a specific carbon source and the model only lacked the transport reaction. In the case of six carbon sources, namely cis-aconitate, D-gluconate, glycogen, ketoglutarate, maltose and R-lactate, such an improvement was observed. By adding the uptake and transport reaction of these metabolites to the model, the accuracy of the model for predicting the carbon source utilization phenotype reached 84% (which was 81% without considering these metabolites).

As we mentioned before, there are 51 metabolites in the Biolog MicroPlate™ which are not included in our metabolic network model (see Supplementary file S4). We observed that for 38 out of these 51 carbon sources (~75%) no growth is observed, while in case of 13 carbon sources we observe some respiration activity. These 13 cases should be further investigated in the future in order to check their potential presence in the metabolism of *P. stutzeri*.

3.5. Model evaluation based on literature data: insights into nitrogen metabolism

P. stutzeri A1501 is famous for its N_2 fixation capability. However, this strain has the ability to utilize nitrogen in different forms. For example, this strain can reduce nitrate to nitrite by using three distinct nitrate reductases through three physiologically different processes [68]. Firstly, nitrate can be used as the terminal electron acceptor in *nitrate respiration* under anaerobic conditions. In this case, nitrate is used in order to generate proton motive force to synthesize ATP. Secondly, nitrate can serve as a nitrogen source for growth in a process called *nitrate assimilation*. Thirdly, nitrate can be used for redox balancing in a process called *nitrate dissipation* when bacteria are grown on highly reduced carbon sources [69].

In the following sections, we show how the results of the *in silico* simulation of *P. stutzeri* matches the experimentally reported phenotypes of nitrogen metabolism in this species. It should be noted that we were not able to model nitrate dissipation, as we did not find any studies about dissimilatory nitrate reductase activity in *P. stutzeri* A1501, and therefore, we do not know the environmental conditions under which nitrate dissipation occurs.

3.5.1. Modeling anaerobic growth and nitrate respiration

All known *Pseudomonas stutzeri* strains are facultative anaerobes with nitrate as the final electron acceptor [30]. In a previous work by Rediers et al., anaerobic growth of *P. stutzeri* cells on a minimal medium supplemented with nitrate was investigated [68]. Using the observed growth curve (i.e., Figure 1 of Ref. [68]), we estimated the anaerobic growth rate of *P. stutzeri* to be ~0.453 h^{-1}. Then, we simulated anaerobic conditions by removing the oxygen uptake reaction (v_{oxygen_uptake}=0) and allowing nitrate to enter the system. By performing FBA in these conditions, the predicted growth rate by the model was found to be 0.398 h^{-1}, which is comparable to the experimentally measured growth rate.

The respiratory nitrate reductase enables bacteria to grow in anaerobic conditions in which reduction of nitrate is coupled with the oxidation of the quinol pool [69]. The three subunits of this highly-conserved enzyme is coded by *nar* genes [70]. In the denitrification process, nitrate is reduced to nitrite, and subsequently to the ultimate product of this pathway, N_2. We investigated the ability of our model to simulate nitrate respiration. As mentioned above, we simulated the anaerobic conditions. Additionally, we ran FVA to see the minimum and maximum possible activity of the denitrification pathway. The results suggest that the pathway is active, carrying 33.0 $mmol.g_{DW}^{-1}.h^{-1}$ flux, and essential for the biomass production. Altogether, our observations provide an explanation on the importance of the denitrification pathway and the reason of its high conservation.

3.5.2. Modeling nitrate assimilation

Nitrate assimilation is a competitive advantage for *P. stutzeri* since it enables the organism to use nitrate as an alternative nitrogen source. This process begins with active transport of nitrate into the cell by an ABC-type transporter, continues with the reduction of nitrate to nitrite by the assimilatory nitrate reductase and subsequent reduction of nitrite to ammonia by the assimilatory nitrite reductase [71, 72]. Since this pathway exists in *P. stutzeri* A1501, the model was tested to evaluate the activity of this pathway qualitatively. In order to simulate a condition in which the only nitrogen source is nitrate, except for nitrate for all other potential nitrogen sources, the lower bounds of their exchange reactions were set to zero. Then, by running FBA, the aforementioned pathway was found to be active, carrying 5.69 $mmol.g_{DW}^{-1}.h^{-1}$ flux. We also ran FVA to see its possible range of fluxes. FVA results also confirmed that this pathway is active, as the minimum and the maximum fluxes of the

reaction were equal to 5.69 mmol.g_{DW}^{-1}.h^{-1}. Therefore, our model could successfully reflect the activity of the nitrate assimilation pathway in the absence of all other nitrogen sources.

3.5.3. Mutation in *narG* gene

The α subunit of the nitrate reductase is encoded by *narG* gene. In a previous study, a *narG* mutant of *Pseudomonas stutzeri* A1501 has been constructed and some of its metabolic features have been studied [68]. Here, we report the results of the *in silico* simulation of these analyses.

3.5.3.1. Growth characteristics

The essentiality of the *narG* gene has been studied, and it has been shown that it is crucial for bacterial growth in anaerobic conditions in a medium containing nitrate [68]. It has also been shown that the mutant strain in *narG* could grow anaerobically in a medium supplemented with nitrite with a growth rate similar to the wild type strain. However, the wild type strain grew faster in the presence of nitrate in comparison to when nitrite was present. We performed four sets of *in silico* experiments in which the biomass production rate was calculated using FBA. In the four aforementioned computational tests, we simulated the conditions in which the wild type strain and the *narG* mutant grew anaerobically in a nitrate- or nitrite-containing medium. We expected to see the highest growth rate in the case of growth in the presence of nitrate for the wild type strain and after that for the case of growth in the presence of nitrite, for both the wild type and the mutant strains. We also expected to see no growth prediction for the mutant strain growing in the medium supplemented with nitrate. The experimental and computational results are shown in Table 2.

Table 2. Experimental observations and computational results concerning anaerobic growth of the wild type and *narG* mutant strain of *Pseudomonas stutzeri* A1501 in the presence of nitrate and nitrite. The qualitative consistency is observable in all four cases.

	Nitrate-containing medium		Nitrite-containing medium	
	Experimental	Computational	Experimental	Computational
Wild type strain	Growth	Growth	Growth	Growth
NarG mutant strain	No growth	No growth	Growth	Growth

As shown in Table 2, the model could reflect the experimental observations successfully. The observed experimental difference could not be seen *in silico* since the reactions of the denitrification pathway are fully coupled. Consequently, in case that the path starts with nitrate, the result would be the same as the case that the path starts with nitrite (FBA results are shown in Table S3). On the other hand, in the experimental data, the bacterium had a higher growth rate on nitrate compared to nitrite. This could be due to bacteria's lower affinity to nitrite in comparison to nitrate (because a prolonged lag phase was observed when the bacteria were grown on the nitrite-supplemented medium [68]). Since the uptake rate was uncharacterized in all cases, we did not set any constraints on the uptake reactions of neither nitrate nor nitrite.

3.5.3.2. Denitrification activity

The process of nitrate reduction to molecular nitrogen via nitrite and the subsequent gaseous compounds (nitric oxide and nitrous oxide) is called denitrification. The activity of this pathway has been previously studied in the wild type and *narG* mutant strains of *P. stutzeri* A1501 [68]. The study included a qualitative assay based on production of gaseous compounds whilst growing anaerobically

in a solid medium supplemented either with nitrate or nitrite. Based on the results of this experiment the pathway was found to be active in the wild type strain in both nitrate- and nitrite-containing media. Denitrification was also seen in the mutant defective in the respiratory nitrate reductases growing in the nitrite-supplemented medium, but it was not observed for the mutant strain when nitrate was provided (as the terminal electron acceptor).

We simulated anaerobic growth conditions and performed FVA to see if the model could predict the experimental observations correctly. In the case of the wild type strain, the pathway was active when either nitrate or nitrite was allowed to enter the system. For the simulated mutant strain, the pathway could carry flux only when nitrite was provided (Data are shown in Table S4). Therefore, the computational results were consistent with the experimental data.

3.5.3.3. Nitrate reductase activity

Nitrate reductase activity (NRA) of the wild type and the *narG* mutant strain has been investigated in aerobic, micro-aerobic and anaerobic conditions. It was reported that the mutant strain loses its ability to reduce nitrate in anaerobic and micro-aerobic conditions. The only NRA activity of the *narG* mutant was seen under aerobic conditions, which has been deduced to be related to the periplasmic nitrate reductase. This enzyme is not expressed in anaerobic conditions. For the wild type strain, the NRA was observable in all oxygen concentrations. Anaerobic and micro-aerobic conditions had almost equal NRAs and this activity was lower under aerobic conditions[68].

For evaluating the modeling predictions in this case, we simulated the aforementioned conditions for the wild type and the *narG* mutant strain and ran FVA. The results are shown in Table 3. The experimental observations are presented qualitatively and the computational results are shown in values representing the flux of nitrate uptake reaction. In the cases where the minimum and maximum nitrate uptake flux rates were not identical, both values are shown in the table. In anaerobic conditions, the simulations were consistent with the experimental results. Also in the micro-aerobic conditions, an acceptable consistency is observable. The model predicted a lower NRA for the wild type strain in aerobic conditions compared than in anaerobic and micro-aerobic conditions. However, the "poorly active" state of nitrate reductase activity was not predicted by the model for the *narG* mutant strain (Quantitative values are shown in Table S5).

Table 3. Experimental observations and computational simulations for nitrate reductase activity in the wild type and the *narG* mutant strain in anaerobic, micro-aerobic and aerobic conditions.

	Anaerobic		Micro-aerobic		Aerobic	
	Experimental	Computational	Experimental	Computational	Experimental	Computational
Wild type strain	Active	Active	Active	Active	Inactive	Inactive
NarG mutant strain	Inactive	Inactive	Inactive	Poorly active	Inactive	Inactive

3.5.4. Nitrogen fixation and nitrogenase repression

Pseudomonas stutzeri A1501 is best known as a nitrogen fixing bacterium. In this process, atmospheric nitrogen is converted into ammonium. Nitrogen fixation occurs under micro-aerobic conditions in both endophytic [73] and free-living state in *P. stutzeri* A1501 [38]. An enzyme complex called nitrogenase is responsible for the reduction of N_2 to ammonia. It has been shown that nitrogenase activity is repressed in the presence of ammonium [74].

In order to investigate the modeling predictions in this case, we prevented all nitrogen sources from entering the system, except for N_2, by setting the lower bound of their associated uptake reactions to zero. We also simulated a micro-aerobic environment (by setting the maximum uptake rate of oxygen to 1.00 $mmol.g_{DW}^{-1}.h^{-1}$, deduced from experimental observations [75]) since the maximum nitrogenase activity occurs in this condition [68]. We performed flux variability analysis to see the range of nitrogen fixing reaction flux. In these conditions, both minimum and maximum flux values for this reaction were found to be 1.64 $mmol.g_{DW}^{-1}.h^{-1}$. In the second round of simulation, while keeping the glucose uptake rate at 10 $mmol.g_{DW}^{-1}.h^{-1}$, we allowed ammonium enter the system (by adding the constraint $-1000 \leq v_{ammonium} \leq +1000$) to see how it would affect the nitrogenase activity. By running FVA, the minimum and maximum flux rates for the nitrogen fixation reaction changed to zero and 1.11×10^{-5} $mmol.g_{DW}^{-1}.h^{-1}$, respectively. Therefore, our model correctly predicted ammonium usage instead of nitrogen fixation.

3.5.5. Nitrogenase activity during N_2 fixation under different oxygen concentrations

It has been reported that the nitrogenase activity increases when environmental conditions change from anaerobic to micro-aerobic [76]. By simulating the nitrogen fixing conditions as mentioned in the previous section, we performed 11 rounds of FVA when oxygen uptake reaction flux increased gradually form 0 to 1 with 0.1 steps. In all cases, the flux of nitrogen fixation reaction was fixed, i.e., the minimum and maximum flux values were equal. Figure 6 shows how the oxygen uptake rate influences the nitrogen-fixing rate. As the oxygen uptake rate increases from zero (representing anaerobic conditions) to 1.00 $mmol.g_{DW}^{-1}.h^{-1}$ (representing micro-aerobic conditions), the nitrogen fixing reaction carries higher fluxes. It should be noted that there are 59 reactions in the model that consume or produce ammonia. Consequently, the fixed nitrogen in form of ammonia is potentially able to go through different pathways and a linear relation between the nitrogen fixation and biomass production rate is not necessarily expected. In conclusion, the increasing trend of the nitrogen fixation flux as a result of the O_2 flux agrees with the reported experimental results.

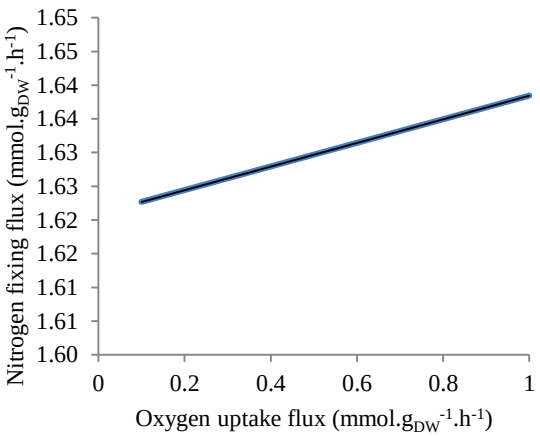


Figure 6. Flux of the nitrogen fixing reaction against oxygen uptake. When the oxygen uptake rate increases from 0 to 1.00 $mmol.g_{DW}^{-1}.h^{-1}$ changing the environmental conditions from anaerobic to micro-aerobic, the nitrogen fixation reaction carries higher flux rates.

3.6. Model evaluation based on literature data: Mutation in *dapB* gene

The *dapB* gene encodes for the L-2,3-dihydrodipicolinate reductase which is essential for meso-diaminopimelate biosynthesis.

Diaminopimelate is an essential component of peptidoglycan and also serves as a major precursor in lysine biosynthesis. Mutation in this gene abolishes microbial growth both on a minimal medium and on a minimal medium containing high concentrations of diaminopimelate. However, the mutant strain can grow in a medium supplemented with high concentrations of lysine and diaminopimelate [77].

We simulated the *dapB* mutant strain by eliminating its associated reaction. By performing FBA, the model predicted no growth in this case. In order to simulate an *in silico* medium containing diaminopimelate, we added a temporary uptake reaction for diaminopimelate so that it would be able to enter the system. It was indicated that diaminopimelate was not readily taken up by the cells [77] therefore, we constrained the uptake rate to have a maximum of 0.100 $mmol.g_{DW}^{-1}.h^{-1}$ and ran FBA. In this case, the biomass producing reaction carried a flux equal to 0.218 h^{-1}. For simulating the growth of the mutant strain on a medium containing high concentrations of diaminopimelate and lysine, we let both of these compounds enter the system and performed FBA. In this case, the model predicted a growth rate equal to 0.565 $mmol.g_{DW}^{-1}.h^{-1}$. Therefore, the simulations could reflect the experimental observations qualitatively. Since the experimental uptake rate was unknown for lysine, we set the lower bound of the uptake reaction to -1000 $mmol.g_{DW}^{-1}.h^{-1}$.

3.7. Model evaluation based on literature data: Indole acetate production

Indole-3-acetate is a phytohormone derived from tryptophan. Production of this compound has been observed in diverse groups of bacteria including *Pseudomonas stutzeri* A15. The pathway begins with the conversion of tryptophan to indolepyruvate, continues to the decarboxylation of indolepyruvate to indole-3-acetyldehyde, and ends with the oxidation of indole-3-acetyldehyde to indole-3-acetate. In an experimental analysis, it was shown that when the cell-free crude extracts of the bacteria is supplemented with tryptophan, indole-3-acetate production increases [78].

We simulated the M9 medium and allowed tryptophan to enter the system by setting the lower bound of its exchange reaction to -10 $mmol.g_{DW}^{-1}.h^{-1}$. We ran FVA to investigate the activity of this pathway. The minimum and maximum of indole-3-acetate production reaction were both 9.67 $mmol.g_{DW}^{-1}.h^{-1}$. Therefore, our model could simulate the activity of this pathway and production of indole-3-acetate when tryptophan was provided.

4. Discussion

The main advantage of genome-scale metabolic networks is that they can provide us with a holistic view of metabolism. Such networks enable us to better understand the quantitative dependencies between metabolic elements which may not be recognizable if a metabolic subnetwork is studied in isolation [79-81]. Consequently, these models can serve as tools for intelligent experimental designs like metabolic engineering and drug targeting [82, 83]. Furthermore, since these networks represent metabolism in a genome scale, they can be used to predict cellular responses to environmental or genetic perturbations such as nutrient deficiency or mutation in one gene or a set of genes.

Here, the first genome-scale metabolic network model for *Pseudomonas stutzeri* A1501 is presented. The model accounts for 890 genes (which cover 21.5% of the genome). These genes are responsible for 995 metabolic reactions. The gene-associated reactions account for 88% of all the reactions in the network. Most of the remaining non-gene-associated reactions are exchange reactions, which are added to the model in order to simulate environmental conditions such as the available carbon source,

environmental oxygen concentration and growth medium formulation. There are 1033 non-exchange reactions in our model and 995 of them are gene-associated. Therefore, 96% of the non-exchange reactions are gene-associated. The percentage of gene-associated reactions in *P. aeruginosa* and *P. putida* models are 95% and 93%, respectively.

Obviously, the number of the genes and the metabolic reactions in our model can increase as more biochemical data become available. For instance, in a recent study a novel form of L-lactate dehydrogenase has been identified in *P. stutzeri* A1501, which uses iron-sulfur clusters as cofactor [84]. The catalyzed reaction by this enzyme can be incorporated into the model if the exact stoichiometry of the reaction is determined in the future.

For network reconstruction, a combination of automated and manual approaches was used in this work. The draft network was constructed using two *Pseudomonas* models using the RAVEN toolbox based on bidirectional BLAST results. Then in an iterative process of gap filling the metabolic reactions specific to *Pseudomonas stutzeri* were added to the draft network from KEGG until the model was able to simulate *in silico* growth and correctly predict other cellular phenotypes.

In this study, a set of experimental data was used to validate the model, including specific growth rate, high-throughput substrate utilization test, and experimental data in the literature. We measured the aerobic growth rate in the minimal M9 medium and found it to be very close to the growth rate predicted by the model in the same environmental condition. The value for the *in silico* growth rate fell within 6% of the maximum growth rate determined experimentally for *P. stutzeri* A1501. For anaerobic conditions, we estimated the specific growth rate using a growth curve of *P. stutzeri* represented in a previous study and compared it to the predicted anaerobic growth rate. Only a slight difference (12%) between the experimental and computational growth rates was observed. The carbon source consumption profile of *P. stutzeri* A1501 was determined using a Biolog phenotyping assay. There was an 87% consistency in experimental and computational results regarding the growth phenotype on different carbon sources. However, for many of the carbon sources in Biolog MicroPlate, we could not simulate the computational growth since the catabolic pathways of these compounds are not entirely known and need to be studied more thoroughly. We also performed a comprehensive literature search looking for more experimental data to evaluate our model with an emphasis on nitrogen metabolism because the modeled strain is best known for its ability to fix atmospheric nitrogen. In most cases, the model could correctly predict the experimental data qualitatively.

Pseudomonas stutzeri A1501 has been originally isolated from rice rhizosphere, and it has been shown that it forms an associative lifestyle with this plant. In this associative relationship, this bacterium can provide the plant with fixed nitrogen [85] and phytohormones, promoting the plant growth[78]. The represented genome-scale metabolic network model for *P. stutzeri* A1501 can provide us with proper strategies to enhance these capabilities so that this bacterium could be used as a natural growth promoter for rice. Since a metabolic network model for rice has been recently reconstructed [86], the interactions between these two organisms can also be modeled and analyzed and to help us better understand the aforementioned association and find potential strategies to optimize the growth of this plant.

Acknowledgements

The financial support for this work by the Iran National Science Foundation (INSF) is gratefully acknowledged. We

thank Ali Muhammadzadeh for editorial comments in connection with writing this paper.

Notes and references

[a] Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran

[b] Corresponding authors. E-mail addresses: marashi@ut.ac.ir (S.-A. Marashi); asad@ut.ac.ir (S. Asad).

† Electronic Supplementary Information (ESI) available:

Supplementary file S1: Metabolic subsystems of unique reactions of *P. stutzeri* format.

Supplementary file S2: The metabolic network of *Pseudomonas stutzeri* A1501 in the standard SBML.

Supplementary file S3: Details of the *in silico* simulations.

Supplementary file S4: Results of Biolog assay for those carbon sources which are not present in *i*PB890.

See DOI: 10.1039/b000000x/

References:

1. N. Beckloff, S. Starkenburg, T. Freitas and P. Chain, in *Microbial Systems Biology*, Springer2012, pp. 471-503.
2. X. Mao, T. Cai, J. G. Olyarchuk and L. Wei, *Bioinformatics*, 2005, **21**, 3787-3793.
3. J. J. Hamilton and J. L. Reed, *Environmental microbiology*, 2014, **16**, 49-59.
4. C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Linsay and R. L. Stevens, *Nature biotechnology*, 2010, **28**, 977-982.
5. M. Kanehisa and S. Goto, *Nucleic acids research*, 2000, **28**, 27-30.
6. R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer and C. Tissier, *Nucleic acids research*, 2008, **36**, D623-D631.
7. J. W. Whitaker, I. Letunic, G. A. McConkey and D. R. Westhead, *Nucleic acids research*, 2009, **37**, D531-D538.
8. J. D. Orth, T. M. Conrad, J. Na, J. A. Lerman, H. Nam, A. M. Feist and B. Ø. Palsson, *Molecular Systems Biology*, 2011, **7:535**, 535.
9. M. Heinemann, A. Kümmel, R. Ruinatscha and S. Panke, *Biotechnology and bioengineering*, 2005, **92**, 850-864.
10. Y.-K. Oh, B. O. Palsson, S. M. Park, C. H. Schilling and R. Mahadevan, *Journal of Biological Chemistry*, 2007, **282**, 28791-28799.
11. M. A. Oberhardt, J. Puchałka, K. E. Fryer, V. A. M. Dos Santos and J. A. Papin, *Journal of bacteriology*, 2008, **190**, 2790-2803.
12. J. Puchałka, M. A. Oberhardt, M. Godinho, A. Bielecka, D. Regenhardt, K. N. Timmis, J. A. Papin and V. A. M. dos Santos, *PLoS computational biology*, 2008, **4**, e1000210.
13. H. Knoop, M. Gründel, Y. Zilliges, R. Lehmann, S. Hoffmann, W. Lockau and R. Steuer, *PLoS computational biology*, 2013, **9**, e1003081.
14. J. Nogales, B. Ø. Palsson and I. Thiele, *BMC systems biology*, 2008, **2**, 79.
15. A. Navid and E. Almaas, *Molecular BioSystems*, 2009, **5**, 368-375.
16. S. Aggarwal, I. A. Karimi and D. Y. Lee, *Molecular BioSystems*, 2011, **7**, 3122-3131.
17. O. Gonzalez, S. Gronau, M. Falb, F. Pfeiffer, E. Mendoza, R. Zimmer and D. Oesterhelt, *Molecular BioSystems*, 2008, **4**, 148-159.
18. A. M. Feist, J. Scholten, B. Ø. Palsson, F. J. Brockman and T. Ideker, *Molecular systems biology*, 2006, **2:0004**.
19. N. Goyal, H. Widiastuti, I. A. Karimi and Z. Zhou, *Molecular BioSystems*, 2014, **10**, 1043-1054.
20. S. Mintz-Oron, S. Meir, S. Malitsky, E. Ruppin, A. Aharoni and T. Shlomi, *Proceedings of the National Academy of Sciences*, 2012, **109**, 339-344.
21. J. Förster, I. Famili, P. Fu, B. Ø. Palsson and J. Nielsen, *Genome research*, 2003, **13**, 244-253.
22. J. M. Dreyfuss, J. D. Zucker, H. M. Hood, L. R. Ocasio, M. S. Sachs and J. E. Galagan, *PLoS computational biology*, 2013, **9**, e1003126.
23. R. L. Chang, L. Ghamsari, A. Manichaikul, E. F. Hom, S. Balaji, W. Fu, Y. Shen, T. Hao, B. Ø. Palsson and K. Salehi-Ashtiani, *Molecular systems biology*, 2011, **7**.
24. I. Thiele, N. Swainston, R. M. Fleming, A. Hoppe, S. Sahoo, M. K. Aurich, H. Haraldsdottir, M. L. Mo, O. Rolfsson and M. D. Stobbe, *Nature biotechnology*, 2013, **31**, 419-425.
25. S. Selvarasu, I. A. Karimi, G.-H. Ghim and D.-Y. Lee, *Molecular BioSystems*, 2009, **6**, 152-161.
26. M. Hadi and S.-A. Marashi, *Molecular BioSystems*, 2014, **10**, 3014-3021.
27. A. R. Zomorrodi, P. F. Suthers, S. Ranganathan and C. D. Maranas, *Metabolic engineering*, 2012, **14**, 672-686.
28. S. Y. Lee, D.-Y. Lee and T. Y. Kim, *Trends in biotechnology*, 2005, **23**, 349-358.
29. N. E. Lewis, H. Nagarajan and B. O. Palsson, *Nature Reviews Microbiology*, 2012, **10**, 291-305.
30. J. Lalucat, A. Bennasar, R. Bosch, E. García-Valdés and N. J. Palleroni, *Microbiology and Molecular Biology Reviews*, 2006, **70**, 510-547.
31. S. W. Park, J. H. Back, S. W. Lee, J. H. Song, C. H. Shin, G. E. Kim and M.-J. Kim, *Kidney Research and Clinical Practice*, 2013, **32**, 81-83.
32. G. L. Gilardi and H. J. Mankin, *N. Y. State J. Med.*, 1973, **73**, 2789-2791.
33. P. Ahamad and A. Kunhi, *Letters in applied microbiology*, 1996, **22**, 26-29.
34. E. Garcia-Valdes, E. Cozar, R. Rotger, J. Lalucat and J. Ursing, *Applied and environmental microbiology*, 1988, **54**, 2478-2485.
35. E. García-Valdés, M. M. Castillo, A. Bennasar, C. Guasp, A. M. Cladera, R. Bosch, K. H. Engesser and J. Lalucat, *Systematic and applied microbiology*, 2003, **26**, 390-403.
36. W. T. Stringfellow and M. D. Aitken, *Applied and Environmental Microbiology*, 1995, **61**, 357-362.
37. R. Rosselló-Mora, J. Lalucat and E. García-Valdés, *Applied and environmental microbiology*, 1994, **60**, 966-972.
38. N. Desnoues, M. Lin, X. Guo, L. Ma, R. Carreño-Lopez and C. Elmerich, *Microbiology*, 2003, **149**, 2251-2262.
39. C. Gao, J. Qiu, C. Ma and P. Xu, *PloS one*, 2012, **7**, e40755.
40. W. Zhang, C. Gao, B. Che, C. Ma, Z. Zheng, T. Qin and P. Xu, *Bioresource technology*, 2012, **110**, 719-722.
41. C. Gao, W. Zhang, C. Lv, L. Li, C. Ma, C. Hu and P. Xu, *Applied and environmental microbiology*, 2010, **76**, 1679-1682.
42. B. Deng, L. Fu, X. Zhang, J. Zheng, L. Peng, J. Sun, H. Zhu, Y. Wang, W. Li and X. Wu, *PloS one*, 2014, **9**, e114886.
43. O. Nwokoro and M. E. U. Dibua, *Archives of Industrial Hygiene and Toxicology*, 2014, **65**, 113-119.
44. Y. Yan, J. Yang, Y. Dou, M. Chen, S. Ping, J. Peng, W. Lu, W. Zhang, Z. Yao and H. Li, *Proceedings of the National Academy of Sciences*, 2008, **105**, 7564-7569.
45. H. Vermeiren, A. Willems, G. Schoofs, R. De Mot, V. Keijers, W. Hai and J. Vanderleyden, *Systematic and applied microbiology*, 1999, **22**, 215-224.
46. J. Hallmann, A. Quadt-Hallmann, W. Mahaffee and J. Kloepper, *Canadian Journal of Microbiology*, 1997, **43**, 895-914.
47. S. A. Becker, A. M. Feist, M. L. Mo, G. Hannum, B. Ø. Palsson and M. J. Herrgard, *Nature protocols*, 2007, **2**, 727-738.

48. O. Resendis-Antonio, J. L. Reed, S. Encarnación, J. Collado-Vides and B. Ø. Palsson, *PLoS Comput Biol*, 2007, **3**, e192.
49. H. Zhao, M. Li, K. Fang, W. Chen and J. Wang, *PLoS ONE*, 2012, **7**, e31287.
50. C. Wang, Z.-L. Deng, Z.-M. Xie, X.-Y. Chu, J.-W. Chang, D.-X. Kong, B.-J. Li, H.-Y. Zhang and L.-L. Chen, *FEBS Letters*, **589**, 285-294.
51. I. Thiele and B. Ø. Palsson, *Nature protocols*, 2010, **5**, 93-121.
52. R. Agren, L. Liu, S. Shoaie, W. Vongsangnak, I. Nookaew and J. Nielsen, *PLoS computational biology*, 2013, **9**, e1002980.
53. S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *Journal of molecular biology*, 1990, **215**, 403-410.
54. M. A. Oberhardt, J. Puchałka, V. A. M. dos Santos and J. A. Papin, *PLoS Computational Biology*, 2011, **7**, e1001116.
55. P. Babaei, T. Ghasemi-Kahrizsangi and S.-A. Marashi, *The Scientific World Journal*, 2014, **2014**, 11.
56. A. M. Feist and B. O. Palsson, *Current Opinion in Microbiology*, 2010, **13**, 344-349.
57. J. Schellenberger, R. Que, R. M. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis and S. Rahmanian, *Nature protocols*, 2011, **6**, 1290-1307.
58. E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch, *Nucleic acids research*, 2003, **31**, 3784-3788.
59. J. D. Orth, I. Thiele and B. Ø. Palsson, *Nature biotechnology*, 2010, **28**, 245-248.
60. R. Mahadevan and C. H. Schilling, *Metabolic engineering*, 2003, **5**, 264-276.
61. T. Chen, Z. W. Xie and Q. Ouyang, *Chinese Science Bulletin*, 2009, **54**, 2610-2619.
62. J. Pramanik and J. D. Keasling, *Biotechnology and Bioengineering*, 1998, **60**, 230-238.
63. J. B. van Duuren, J. Puchałka, A. E. Mars, R. Bücker, G. Eggink, C. Wittmann and V. A. M. dos Santos, *BMC biotechnology*, 2013, **13**, 93.
64. P. Chomczynski and N. Sacchi, *Analytical Biochemistry*, 1987, **162**, 156-159.
65. Y.-C. Chen and J. Jaczynski, *Journal of agricultural and food chemistry*, 2007, **55**, 1814-1822.
66. Y.-C. Chen and J. Jaczynski, *Journal of agricultural and food chemistry*, 2007, **55**, 9079-9088.
67. B. A. Bidlingmeyer, *Preparative liquid chromatography*, Elsevier1987.
68. H. Rediers, J. Vanderleyden and R. De Mot, *Microbiological research*, 2009, **164**, 461-468.
69. D. J. Richardson, *Microbiology*, 2000, **146**, 551-571.
70. L. Philippot, *Biochimica et biophysica acta (BBA)-Gene structure and expression*, 2002, **1577**, 355-376.
71. J. T. Lin and V. Stewart, *Advances in microbial physiology*, 1997, **39**, 1-30.
72. L. Reitzer, *Annual Reviews in Microbiology*, 2003, **57**, 155-176.
73. E. James, *Field Crops Research*, 2000, **65**, 197-209.
74. T. Zhang, Y. Yan, S. He, S. Ping, K. M. Alam, Y. Han, X. Liu, W. Lu, W. Zhang and M. Chen, *Research in microbiology*, 2012, **163**, 332-339.
75. F. Bergersen, *Soil Biology and Biochemistry*, 1997, **29**, 875-880.
76. Y. Yan, S. Ping, J. Peng, Y. Han, L. Li, J. Yang, Y. Dou, Y. Li, H. Fan and Y. Fan, *BMC genomics*, 2010, **11**, 11.
77. H. Rediers, V. Bonnecarrere, P. B. Rainey, K. Hamonts, J. Vanderleyden and R. De Mot, *Applied and environmental microbiology*, 2003, **69**, 6864-6874.
78. R. O. Pedraza, A. Ramírez-Mata, M. Xiqui and B. E. Baca, *FEMS microbiology letters*, 2004, **233**, 15-21.
79. S.-A. Marashi, L. David and A. Bockmayr, *Algorithms for Molecular Biology*, 2012, **7**, 17.
80. A. Rezvan, S.-A. Marashi and C. Eslahchi, *Journal of Bioinformatics and Computational Biology*, 2014, **12**, 1450028.
81. B. P. Ingalls and E. Bembenek, *In Silico Biology.*, 2015, In press. DOI: 10.3233/ISB-140464.
82. Z. A. King, C. J. Lloyd, A. M. Feist and B. O. Palsson, *Current Opinion in Biotechnology*, 2015, **35**, 23-29.
83. H. U. Kim, S. B. Sohn and S. Y. Lee, *Biotechnology Journal*, 2012, **7**, 330-342.
84. C. Gao, Y. Wang, Y. Zhang, M. Lv, P. Dou, P. Xu and C. Ma, *Journal of Bacteriology*, 2015, **197**, 2239-2247.
85. C. You and F. Zhou, *Canadian journal of microbiology*, 1989, **35**, 403-408.
86. P. Dharmawardhana, L. Ren, V. Amarasinghe, M. Monaco, J. Thomason, D. Ravenscroft, S. McCouch, D. Ware and P. Jaiswal, *Rice*, 2013, **6**, 1-15.