Predicting Strain Engineering Strategies Using iKS1317: A Genome-Scale Metabolic Model of *Streptomyces coelicolor*

Tjaša Kumelj, Snorre Sulheim, Alexander Wentzel, and Eivind Almaas*

Streptomyces coelicolor **is a model organism for the** *Actinobacteria*, **a phylum known to produce an extensive range of different bioactive compounds that include antibiotics currently used in the clinic. Biosynthetic gene clusters discovered in genomes of other** *Actinobacteria* **can be transferred to and expressed in** *S. coelicolor*, **making it a factory for heterologous production of secondary metabolites. Genome-scale metabolic reconstructions have successfully been used in several biotechnology applications to facilitate the over-production of target metabolites. Here, the authors present iKS1317, the most comprehensive and accurate reconstructed genome-scale metabolic model (GEM) for** *S. coelicolor*. **The model reconstruction is based on previous models, publicly available databases, and published literature and includes 1317 genes, 2119 reactions, and 1581 metabolites. It correctly predicts wild-type growth in 96.5% of the evaluated growth environments and gene knockout predictions in 78.4% when comparing with observed mutant growth phenotypes, with a total accuracy of 83.3%. However, using a minimal nutrient environment for the gene knockout predictions, iKS1317 has an accuracy of 87.1% in predicting mutant growth phenotypes. Furthermore, we used iKS1317 and existing strain design algorithms to suggest robust gene-knockout strategies to increase the production of acetyl-CoA. Since acetyl-CoA is the most important precursor for polyketide antibiotics, the suggested strategies may be implemented in vivo to improve the function of** *S. coelicolor* **as a heterologous expression host.**

1. Introduction

The increasing resistance of pathogenic bacteria to antibiotics combined with a continuous low level of discovery and development of new antibacterial drugs is a major threat to global health. One reason for the lack of novel drugs is that the traditional method of bioprospecting, involving cultivation and high-throughput screening, is no longer efficient, partly because of a high rate of rediscovery.[1–3] A promising approach for discovery and production of novel bioactive metabolites is based on the heterologous expression of biosynthetic gene clusters in specialized expression host strains. One organism in which this already has been achieved is *Streptomyces coelicolor*.[4–8]

The genus *Streptomyces* is one of the most important sources of bioactive, microbial metabolites. *S. coelicolor* is a model organism for this genus[9] with a capacity to produce 31 different secondary metabolites,[10] including four antibiotics: actinorhodin, calcium-dependent antibiotic, undecylprodigiosin and methylenomycin.[11,12] Note that none of these four antibiotics are of medical relevance. Thus, it has a metabolic machinery capable of providing precursors for a large range of different classes of bioactive metabolites, which is a necessary feature of a host for heterologous expression of biosynthetic gene clusters and production of the encoded compounds. An improved *S. coelicolor* strain for heterologous expression has already been developed by removing two plasmids naturally present in *S. coelicolor* A3(2) and four major biosynthetic gene clusters from the chromosome,[13] resulting in a reduced metabolic and bioactive background. To further improve this

T. Kumelj, S. Sulheim, Prof. E. Almaas
Department of Biotechnology and Food Science
NTNU - Norwegian University of Science and Technology
Trondheim, Norway
E-mail: eivind.almaas@ntnu.no

S. Sulheim, Dr. A. Wentzel
SINTEF Industry
Department of Biotechnology and Nanomedicine
Trondheim, Norway

Prof. E. Almaas
K.G. Jebsen Center for Genetic Epidemiology
NTNU – Norwegian University of Science and Technology
Trondheim, Norway

organism as an expression host, it is necessary to develop a more comprehensive understanding of the metabolism.

A genome-scale metabolic model (GEM) is a network representation of the metabolic capabilities of an organism, constructed from an annotated genome by using inferred or proven gene-protein-reaction relations, in addition to transport reactions, and an estimated biomass composition. A detailed reconstruction protocol is described in ref. [14]. The network of reactions and metabolites in a GEM can be mathematically represented by a stoichiometric matrix. Using a variety of constraint-based modeling approaches, this stoichiometric matrix serves as a core input to predict phenotypes of the organism subject to perturbations or its behavior in different growth environments. GEMs have been used successfully to direct strain engineering of organisms (e.g., see ref. [15] for a review).

Currently, there exist three genome-scale metabolic models for *S. coelicolor*; the iIB711,[16] the iMA789,[17] and the iMK1208,[18] with the most recent being published in 2014. The iMA789 is an improved version of iIB711, and it includes a more comprehensive reconstruction of pathways for the production of antibiotics. The iMK1208 was reconstructed de novo based on annotations in StrepDB,[19] Kyoto Encyclopedia of Genes and Genomes (KEGG),[20] BioCyc,[21] and TransportDB,[22] and with updated biomass and ATP-maintenance reactions [18].

Here, we present iKS1317, a more validated and comprehensive GEM of *S. coelicolor* based on the previous model iMK1208,[18] appended and corrected with knowledge obtained from iMA789,[17] KEGG,[20] and BioCyc.[21] Both reactions and metabolites are annotated with KEGG-identifiers when possible, and they are named according to guidelines and existing names in BiGG.[23] The recent transposon mutagenesis study by Xu et al.[24] enabled a thorough evaluation of the model's accuracy in predicting single gene knockout growth phenotypes. iKS1317 is written in the SBML format (level 3) and is compatible with both the COBRA Toolbox for Matlab and COBRApy.[25,26]

Using iKS1317 as the basis for constraint-based optimization analyses, we suggest engineering strategies that may increase the heterologous production of polyketide antibiotics because of increased availability of the primary precursor acetyl-CoA. We predict the optimal yield of acetyl-CoA for *S. coelicolor* in response to single, double and triple reaction deletions in three different growth environments. Furthermore, we compare the results from OptKnock[27] and Genetic Design through Local Search (GDLS),[28] the two strain engineering methods used in this study.

2. Experimental Section

2.1. iKS1317 Model Reconstruction

The GEM presented here, iKS1317, is based on the iMK1208 model by Kim et al.[18] An overview of the origin of reactions and metabolites in iKS1317 is given in **Table 1**. Of the 1859 reactions in iMK1208, 1852 were included in iKS1317 with no or minor changes. Four out of the seven removed reactions consisted of lumped reactions that we replaced by detailed, multi-reaction steps. One reaction was a duplicate, and two reactions in the actinorhodin pathway were mapped to reactions R09312 and R09313 in KEGG.[20] A complete list of the removed reactions is provided in S1, Supporting information.

Based on the original biomass reaction in iMK1208, a second biomass reaction were constructed where the amino acids have been replaced by their respective tRNA-charged versions. The corresponding released tRNA molecules were added as products to balance the equation, and we adjusted the stochiometric coefficients of ATP, ADP, water, protons, and phosphate to account for the energy consumed by the reactions charging the amino acids with tRNA molecules. The content of both biomass reactions is given in S2, Supporting information.

Metabolites in iMA789 and reactions in both iMA789 and iMK1208 were mapped to KEGG-identifiers,[20] making it possible to directly compare the two models.[17,18] The iMK1208 reconstruction is not based upon the iMA789, and we found 167 reactions in iMA789 not present in iMK1208 that we chose to include in iKS1317. The metabolites were mapped to KEGG-identifiers based on their name and chemical formula, and the reactions were mapped based on name, reactants, products, and co-factors. With the reactions annotated with KEGG-identifiers we could compare the content of iKS1317 with the list of reactions in KEGG associated to genes in the genome of *S. coelicolor* A3(2). This allowed us to find reactions in KEGG not already present in the model, and this investigation resulted in another 87 reactions appended to iKS1317. We assumed that the annotations in KEGG were correct if the reactions fitted well with the existing content of the model. If the KEGG-annotations were in contradiction to our existing model or involved a new pathway, they were further evaluated by using BioCyc, published literature, BRENDA, or Uniprot-SwissProt.[21,29,30]

The metabolite formulas in iMA789 and KEGG are given in neutral (non-charged) form, while charged formulas are used in iMK1208. Most metabolites are charged in the cellular environment and this is also recommended in the 96-step protocol for model reconstruction by Thiele and Palsson.[14] We calculated the charged chemical formula of the metabolites added from KEGG and iMA789 at pH 7 using eQuilibrator.[31] Since not all chemical formulas could be calculated in this fashion, the formula for some metabolites were inferred by comparing neutral and charged formulas for similar metabolites.

eQuilibrator was also used to calculate the change in Gibbs free energy at standard conditions to infer reaction directionality in reactions from KEGG and iMA789. Because the concentration of each metabolite in the cell is unknown, we cannot accurately predict the change in Gibbs free energy of a reaction. Hence, we assumed most reactions to be reversible unless eQuilibrator predicted a large ($>30\,\text{kJ mol}^{-1}$)[32,33] change in Gibbs free energy of the reaction.

Table 1. Origin of reactions and metabolites in the reconstructed metabolic network iKS1317.

	Reactions	Metabolites
iMK1208	1853	1435
iMA789	167	68
KEGG	87	69
BioCyc	12	9

A detailed overview of all reactions and their origin is found in S2, Supporting information.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**Biotechnology
Journal**

www.biotechnology-journal.com

By using the excellent review of the biosynthetic pathways in *S. coelicolor* by Challis,[11] we added pathways for three secondary metabolites (geosmin, albaflavenon, methylenomycin) and extended the undecylprodigiosin pathway to include streptorubin. Most of these reactions were also described in BioCyc.[21]

Fifty new transport reactions were added from iMA789, one of them transporting sucrose into the cytoplasm. The latter reaction enabled growth with sucrose as the sole carbon source, but according to Hodgson[34] (as referred to by Borodina et al.[16]) *S. coelicolor* is unable to catabolize sucrose. However, *S. coelicolor* is supposed to easily take up sucrose to balance the osmotic pressure (Bibb, 1985, as cited by Elibol, 1998).[35,36] To avoid this erroneous in silico prediction, the uptake reaction rate for the sucrose-transport reaction was set to zero.

2.2. Validation of iKS1317

A listing of growth phenotyping data is available in the paper describing iMA711.[16] Through additional literature review, we were able to identify growth data for 63 conditions for wild-type or mutant strains.[16,34,37–44] The transposon mutagenesis data published by Xu et al.[24] provide a valuable resource for model validation and improvement. This data set includes the growth phenotyping for 497 different gene knockout mutants, from which 365 are considered to be non-essential and 132 considered to be essential, unconditionally of the growth environment. The non-essential genes are confirmed by the ability of their knockout mutants to grow. The 132 unconditionally essential genes are identified by their lack of presence in any of the cultivated knockout mutants. Because the probability of having a gap in the genome in the transposon mutagenesis data increases with decreasing gap length, only gaps, and thus only genes, longer than 1.9 kb were included in the list of unconditionally essential genes.[24] One hundred thirty-seven of these 497 different genes are present in iKS1317, of which 77 are non-essential and 60 are considered unconditionally essential genes. The transposon mutagenesis study was carried out on the *S. coelicolor* X737 mutant, from which all genes in the actinorhodin gene cluster are removed. We therefore removed these genes from our in silico model before computing the gene knockout growth phenotypes of iKS1317. In this study, we have assumed that the observations from the transposon mutagenesis study are absolutely correct, however this kind of large scale knockout experiments is difficult and can contain errors. A detailed overview of all tested growth conditions is found in S3, Supporting information.

We performed the comparison of in silico and in vivo growth in a binary fashion, classifying each condition as either *growth* or *no growth*. To compare growth rates in different environmental conditions we constrained the total uptake of carbon and nitrogen to 12.6 and 1.85 mmol g dry weight^{-1} h^{-1}, respectively. This corresponds to the maximal experimentally observed uptake of glucose (2.1 mmol g dry weight^{-1} h^{-1})[45] and the simulated corresponding ammonium uptake (1.85 mmol g dry weight^{-1} h^{-1}). The specific uptake rates for each of the evaluated environments are given in S3, Supporting information. When we compared in silico predictions with experimental data, we needed to impose a lower threshold for the experimentally measurable growth rate. The measured growth rate for the maximal uptake rate of glucose is

0.128 h^{-1},[45] corresponding to a doubling time of about 5.4 h. Consequently, we considered growth with a doubling time of more than 1 day as a reasonable lower threshold. This choice of threshold value only affected the growth on L-phenylalanine as carbon source, with a computationally predicted doubling time of 70 h.

For comparing our in silico predictions with growth data from the transposon mutagenesis study,[24] we used a threshold of 50% of the in silico wild-type growth rate in the complex cultivation medium to decide the binary growth versus no growth test.[46] This relatively large threshold was considered to be appropriate because of the methods used in a transposon mutagenesis study. We used the second biomass reaction where the amino acids in the primary biomass function are replaced by their respective tRNA charged versions to predict the knockout phenotypes. This enabled correct phenotype prediction for mutants where the knocked out genes are related to tRNA charging of amino acids.

The transposon mutants were sporulated and grown in complex media (SFM[47] and YBP,[48] respectively). This has direct implications for the modeling, since it is difficult to determine the carbon and nitrogen sources that actually were available and utilized by the organism. We therefore assumed that all carbon and nitrogen sources with an exchange reaction present in iKS1317 were available, a detailed list is given in S3, Supporting information. However, to shed light on some of the limitations of the use of the transposon mutagenesis data, we also included a comparison where the in silico growth of mutants was predicted with only glucose and ammonium available as the carbon and nitrogen source, respectively.

2.3. Suggesting Optimal Knockout Strategies with iKS1317

Two different strain design algorithms, OptKnock[27] and GDLS,[28] were used to predict genetic manipulations for target overproduction. The methods use constraint-based optimization to suggest reaction knockout (constraining the metabolic flux of a reaction to zero) strategies to gain targeted overproduction while optimizing internal (biomass yield) and external (product yield) cellular objectives. A direct consequence of these algorithms design is that overproduction of the target becomes an obligatory by-product of growth.[27,28] These methods are accessible through the COBRA Toolbox v3.0[25] in Matlab. The suggested reactions can be removed in vivo by knocking out one or more of the genes encoding the enzymes catalyzing the reaction.

The reaction knockout strategy was constrained with upper and lower bound on the ATP-maintenance reaction set to 2.65 mmol g dry weight^{-1} h^{-1}[18] and with a lower bound on the biomass growth rate of 0.05/h. The uptake rate of available carbon and nitrogen sources were set to 0.8 mmol g dry weight^{-1} h^{-1}, except for the uptake of ammonium which was unlimited. We constrained OptKnock and GDLS to only allow knockout of enzymatic reactions with one or more associated genes, that is, all exchange reactions, 90 transport reactions, the biomass reactions, and the ATP-maintenance reaction were restricted from being removed. Additionally, 18 reactions related to oxidative phosphorylation were neither allowed to be knocked out (see S4, Supporting information, for details). The triple knockouts with OptKnock were computed on one node of an HPC platform with access to two Intel Xeon E5-2660 v3 CPUs.

We implemented the strain design algorithms with iKS1317 in three different growth environments: 1) a basic environment with glucose and ammonium; 2) a glucose-based environment enriched with nitrogen sources, that is, glutamate, nitrate, and ammonium; and 3) an environment enriched with carbon sources, that is, galactose, glycerol, mannitol, and with ammonium as the nitrogen source. An overview is given in **Table 3**A. By comparing the predicted strain engineering strategies in these three environments, we could evaluate the consistency of the suggested genetic modifications.

3. Results

The genome-scale metabolic model iKS1317 of *S. coelicolor*, contains 1317 genes (16% of the protein-coding genes in the genome),[49–51] 2119 reactions and 1581 metabolites. An overview of the iKS1317 composition is given in **Figure 1**B. Both the model reactions and metabolites are now consistently annotated with KEGG IDs whenever possible. Reactions are annotated to 96 different pathways within 10 different subsystems, as defined by KEGG (Figure 1B).[20,53] The additional reactions and genes relative to iMK1208 are mostly located in primary metabolism. However, we have also added pathways allowing the production of geosmin, albaflavenon, coenzyme F420, and methylenomycin, none of which are present in the previous model, iMK1208.[18]

3.1. 83.3% Accuracy in Predicting Growth and Knockout Phenotypes

When we test the ability of our model to correctly predict experimental growth phenotypes, we find correct predictions in 96.5% (55/57) of the tested growth environments for the wild-type *S. coelicolor* (**Table 2**). Another approach for assessing the quality of a genome-scale reconstructed metabolic network, is to compare in vivo with in silico predictions of growth phenotypes for single-gene knockout mutants. This test is assessing the quality of the reconstruction of alternative metabolic pathways, and thus, is a complementary test to that of wild-type growth phenotyping. We found that iKS1317 predicts the correct knockout phenotype in 78.4% (120/153) of the compared conditions (Table 2). Eight of the correctly predicted knockout phenotypes are related to tRNA charging of the amino acids and would be false if the primary biomass reaction was used in the validation. In sum, iKS1317 has been evaluated in 210 different conditions and has an accuracy of 83.3% (175/210) in predicting growth and knockout phenotypes. The iMK1208 provides similar accuracy for predicting growth phenotypes (96.5%) and a 71.4% (105/147) accuracy for knockout mutants, resulting in an overall accuracy of 78.4% (160/204) (Figure 1A). Six of the 153 knockout phenotypes could not be evaluated by iMK1208 because the genes were not present in the model. A spreadsheet describing all growth comparisons is provided in S3, Supporting information.

3.2. OptKnock and GDLS Predicts Approximately 2-Fold Increase of Acetyl-CoA Production for Double-Knockout Mutants

Since the pyruvate dehydrogenase (PDH) reaction produces acetyl-CoA from pyruvate and coenzyme A (CoA), we selected it as the target reaction for the strain engineering analyses. The two strain engineering strategies, OptKnock[27] and GDLS,[28] predicted identical reaction knockouts as the optimal solution
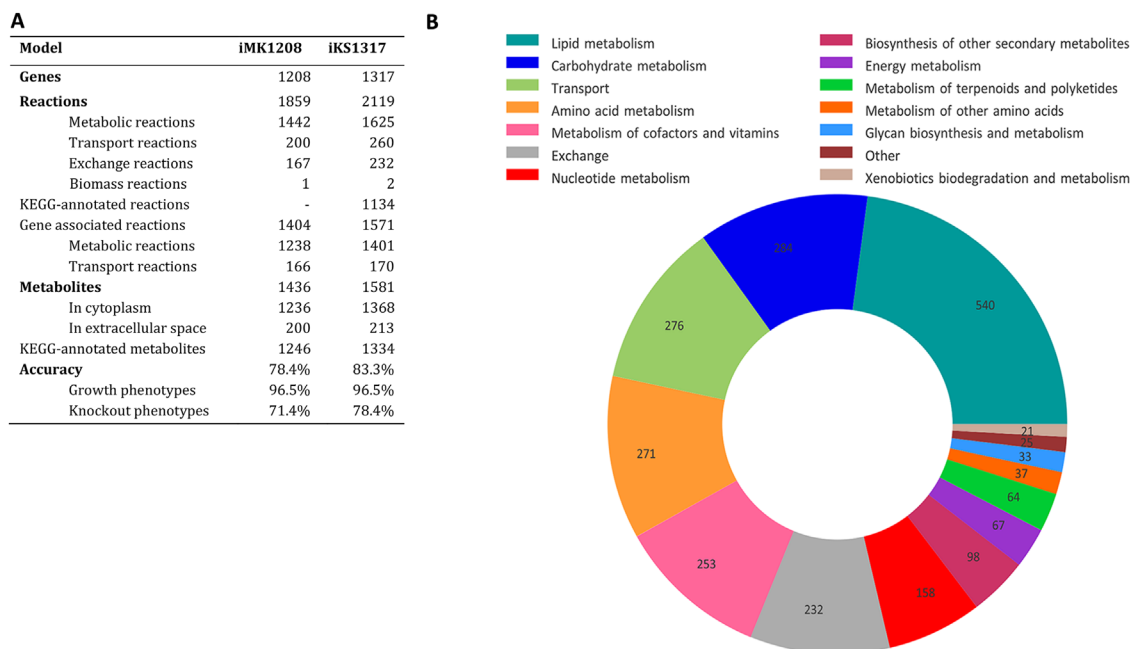
Model	iMK1208	iKS1317
Genes	1208	1317
Reactions	1859	2119
Metabolic reactions	1442	1625
Transport reactions	200	260
Exchange reactions	167	232
Biomass reactions	1	2
KEGG-annotated reactions	-	1134
Gene associated reactions	1404	1571
Metabolic reactions	1238	1401
Transport reactions	166	170
Metabolites	1436	1581
In cytoplasm	1236	1368
In extracellular space	200	213
KEGG-annotated metabolites	1246	1334
Accuracy	78.4%	83.3%
Growth phenotypes	96.5%	96.5%
Knockout phenotypes	71.4%	78.4%

Figure 1. **A**) The table displays a side-by-side comparison of the two most recent genome-scale metabolic models for *S. coelicolor*. When assuming minimal cultivation medium with glucose and ammonium as the only carbon and nitrogen source, iKS1317 has a total accuracy of 90%. See Section 4 for details. **B**) The pie chart shows how reactions in iKS1317 are distributed over the main KEGG pathways, as well as exchange and transport reactions.

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**Biotechnology
Journal**
www.biotechnology-journal.com

for single, double and triple knockouts in environments 1 and 3. For environment 2, both the single and double knockouts were identical, but GDLS provided a sub-optimal solution for the triple-knockout.

OptKnock also provides alternative solutions when several strategies are optimal (Strategy 7 and 8, Table 3B) and a pre-determined number of sub-optimal solutions (S4, Supporting information). The predicted production rates of acetyl-CoA through the PDH reaction are reported as the ratio of the rate in the knockout mutant to the rate in the wild-type in the same growth environment (**Figure 2**A, Table 3B). Absolute production rates, growth rates, associated genes, and sub-optimal solutions are provided in S4, Supporting information.

A numbering of the suggested strain engineering strategies and the reactions suggested for knockout are given in Table 3B. The name, KEGG ID and related pathway of these reactions are given in Table 3C. In the following section we refer to the strategies by the numbering and the reactions by their name.

The triple knockout in Environment 2 (Strategy 1) disrupt the synthesis of glutamate from alpha-ketoglutarate and provide the largest maximal relative production of acetyl-CoA (2.84). However, this strategy is not recommended for in vivo experiments because the minimal possible relative flux through PDH in this strategy is zero (Figure 2A). The double knockout in this environment (Strategy 2) provide a better solution: a large maximal relative rate of acetyl-CoA production (2.62) and a minimal relative equal the wild-type strain. In this strategy succinyl-CoA synthetase and glycine hydroxymethyltransferase is removed, and the effect on the flux distribution in the TCA cycle and central metabolism is displayed in Figure 2B. The reactions with major increase and decrease in flux is highlighted in red and blue, respectively. The removed succinyl-CoA synthetase is marked by a red X, and it is obvious that this has major impact on the flux distribution in the TCA cycle. We observe that the flux is rerouted out of the TCA cycle, into glyoxylate cycle to produce succinate and malate through isocitrate lyase and malate synthase. This increases the flux through the PDH reaction because malate is converted to pyruvate by malate dehydrogenase. Glycine hydroxymethyl-transferase converts serine to glycine and is not shown in this figure, but the removal reroutes the synthesis of glycine through threonine. It is not obvious how this connects to the flux change in the TCA cycle, and this illustrates well why a computational approach is necessary to identify the optimal engineering strategies.

Both the double and triple reaction knockouts in Environments 1 and 3 (Strategy 3–5) give similar changes in the flux distribution as Strategy 2 (Figure 2B). The overall pattern is that succinyl-CoA synthetase is knocked out along with either glutamate dehydrogenase or glycine hydroxymethyltransferase. The triple knockouts provide the largest predicted maximal production of acetyl-CoA in these environments (1.97 and 1.79 for Environments 1 and 3, respectively), but the small gain compared to Strategy 4 (1.80 and 1.71 for Environments 1 and 3, respectively) will probably not be worth the extra effort required to perform the additional knockout in vivo. The single reaction knockout strategies provide almost no increase in the production of acetyl-CoA (Figure 2A).

4. Discussion

This work presents iKS1317, an updated genome-scale metabolic network for *S. coelicolor*, providing more accurate predictions than any previous genome-scale reconstruction for this organism (Table 1). The iKS1317 network is also more comprehensive, more thoroughly annotated, and better validated than previous models.

Many of the discrepancies between in vivo growth and in silico predictions present in previous metabolic reconstructions of *S. coelicolor* are related to the degradation and biosynthesis of branched-chain amino-acids, and they have been resolved in iKS1317. It is known that the *S. coelicolor* Δvdh (SCO4089) knockout mutant is unable to grow with L-valine, L-leucine, or L-isoleucine as the sole carbon source,[43] an observation not supported by iMK1208.[18] However, by only changing the L-valine (R01214), L-leucine (R01090), and L-isoleucine (R02199) transaminase reactions from reversible to irreversible, these reactions are prevented from participating in the degradation of branched-chain amino acids, and these experimental results are recovered in iKS1317. In contradiction, the estimated changes in Gibbs free energy at 1 mM concentration and standard conditions are 3.2 ± 6.9, -1.2 ± 3.2, and $-4.7 \pm 6.9 \, \text{kJ mol}^{-1}$ for R01214, R01090, and R02199, respectively,[31] not indicating that the reactions are irreversible in any direction. However, different metabolite concentrations and the efficiency of upstream and downstream reactions have a major impact on these values.

Second, the $\Delta msdA$ (SCO2726) knockout mutant is incapable of growth in vivo, with L-valine as the sole carbon source. The gene *msdA* encodes for the enzyme methylmalonate-semi-aldehyde dehydrogenase which catalyze the reactions methyl-malonate-semialdehyde: NAD+ oxidoreductase (R00935) and 3-oxopropanoate:NAD+ oxidoreductase (R00705). Removing the reaction methylmalonate semialdehyde: NAD+ oxidoreductase disrupts the primary degradation pathway of L-valine, but according to the metabolic reconstruction L-valine can also be degraded through the pathways for biosynthesis and degradation of L-leucine. This connection is possible because of the enzyme catalysing the reaction 2-isopropylmalate synthase (R01213). We have in iKS1317 introduced a redox coupling (acetyl-CoA/CoA) between 3-oxopropanoate: NAD+ oxidoreductase (R00705) and

Table 2. This table details the result of the comparison between in silico predictions and in vivo observations.

		Growth environments		Knockout mutants	
	In silico	Growth	No growth	Growth	No growth
In vivo	Growth	TP: 51	FN: 0	TP: 78	FN: 6
	No growth	FP: 2	TN: 4	FP: 27	TN: 42

The left part displays the growth phenotypes for 57 different growth environments and the right part display growth phenotypes for 153 different knockout mutants. The two false positive predictions for the growth environments are with glutamine and aspartate as the sole carbon source. The 27 false predictions for the knockout mutants are examined in the Discussion section. The following abbreviations are used: true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN).
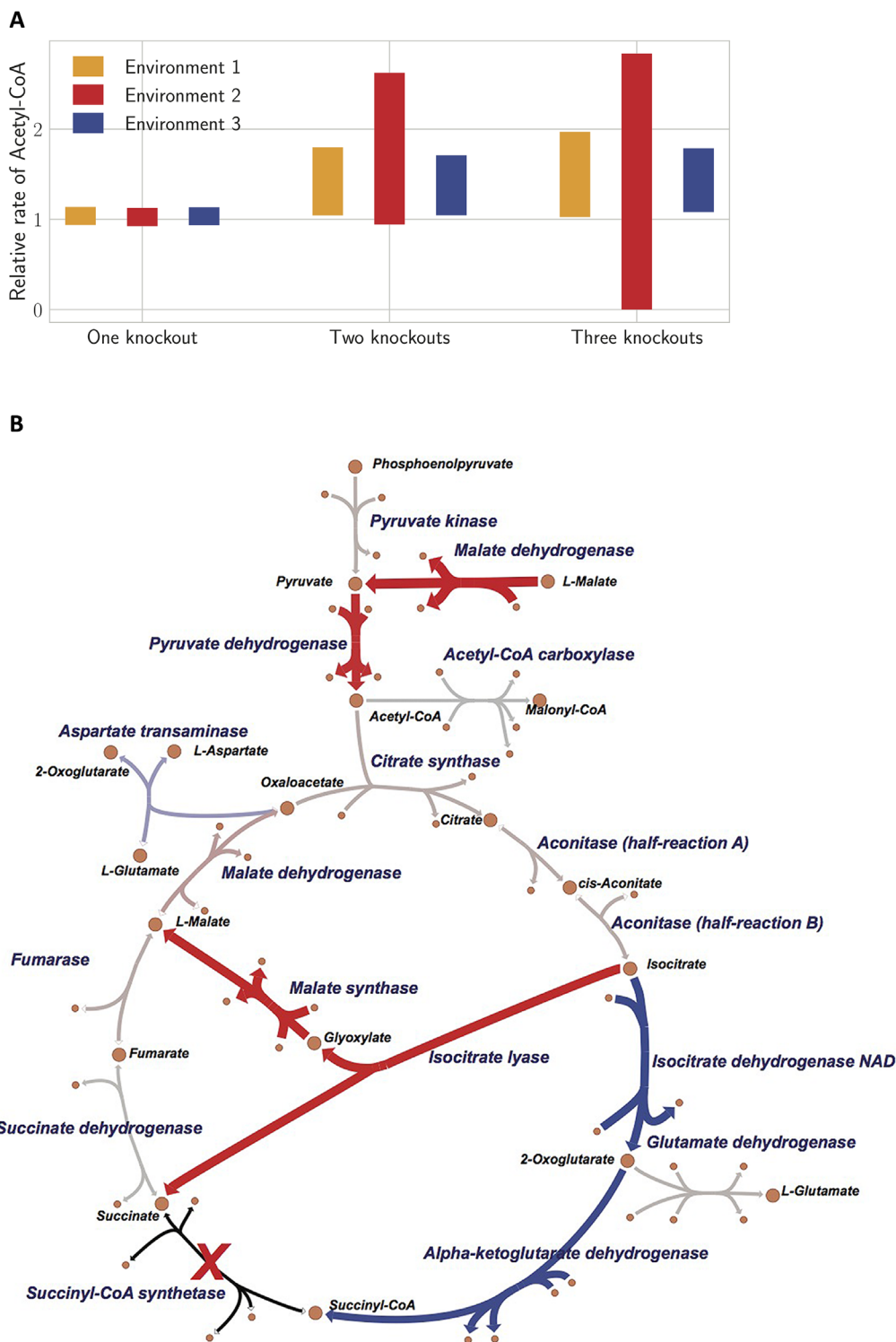
ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

Biotechnology
Journal
www.biotechnology-journal.com

Figure 2. A) Predicted relative production rate ranges of the primary antibiotic precursor acetyl-CoA through the pyruvate dehydrogenase (PDH) reaction for single, double, and triple reaction knockout mutants in the three different growth environments (Table 3A). Each bar spans between the minimum and maximum computed rates, relative to the rate of the PDH reaction for the wild-type in the same growth environment. OptKnock[27] and GDLS[28] provided identical predictions except for the triple knockout in environment 3, where GDLS returned sub-optimal solutions. **B)** The predicted change in flux distribution when succinyl-CoA synthetase (marked by red X) and glycine hydroxymethyltransferase (not in figure) are knocked out (Strategy 2, Table 3B). The flux is rerouted out of the TCA cycle and into the glyoxylate cycle to produce succinate and malate. This increases the flux through malate dehydrogenase and pyruvate dehydrogenase. The map is drawn using Escher and display only the TCA cycle and related reactions in iKS1317.[52] The reactions with major increase and decrease in flux are highlighted in red and blue, respectively.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**Biotechnology
Journal**

www.biotechnology-journal.com

Table 3. An overview of the strain engineering environments and results.

(A) Environments

Environment no.	Carbon sources	Nitrogen sources
1	Glucose	Ammonium
2	Glucose	Glutamate[a], nitrate, ammonium
3	Galactose, glycerol, mannitol	Ammonium

(B) Strain engineering predictions

Strategy no.	Environment no.	Reactions	Max production
1	2	R00114, R00093, R00248	2.84
2	2	R00405, R00945	2.62
3	1	R00405, R00658, R00945	1.97
4	1	R00405, R00248	1.80
5	3	R00405, R00248, R00955	1.79
4	3	R00405, R00248	1.71
6	2	R00844	1.05
7	1	R04780	1.04
8	1	R01070	1.04
9	3	R00248	1.03

(C) KEGG ID, name and pathway of suggested reactions

KEGG ID	Reaction name	Pathway
R00093	Glutamate synthase	Glutamate metabolism
R00114	Glutamate synthase	Glutamate metabolism
R00248	Glutamate dehydrogenase	Glutamate metabolism
R00405	Succinyl-CoA synthetase	TCA-cycle
R00658	Enolase	Glycolysis/gluconeogenesis
R00844	Glycerol-3-phosphate dehydrogenase	Glycerophospholipid metabolism
R00945	Glycine hydroxymethyltransferase	Glycine, serine and threonine metabolism
R00955	Uridyl transferase	Galactose metabolism
R01070	Fructose-bisphosphate aldolase	Glycolysis
R04780	Fructose 1,6-bisphosphatase	Gluconeogenesis

Table 3A display the carbon and nitrogen sources of the three environments. Table 3B display the predicted optimal strain engineering strategies: The maximal production is the relative rate of the pyruvate dehydrogenase reaction (PDH) with respect to the maximal rate of this reaction for the wild-type in the same growth environment. Table 3C display the KEGG ID, name and pathway of each of the reactions in the suggested strategies. The absolute minimal and maximal production rates, growth rates, associated genes, and sub-optimal solutions are provided in S4, Supporting information. [a]Glutamate is a source of both carbon and nitrogen.

2-isopropylmalate synthase (R01213) which blocks the latter reaction for the Δ*msdA* knockout mutant and solves this discrepancy between in vivo observation and in silico prediction. A detailed description is given in S5, Supporting information.

While our analysis and manual curation of the metabolic network reconstruction has removed the two before mentioned discrepancies, we are still left with an incorrect growth prediction of the Δ*msdA* (SCO2726) mutant with propionate as the sole carbon source. In contrast to in vivo experiments,[37] the iKS1317 predicts no growth, since the introduced coupling also disrupts the synthesis of L-leucine, which is an essential amino acid. Further research is required to fully understand the regulatory mechanisms involved in the synthesis and degradation of the branched-chain amino acids in *S. coelicolor*.

In iKS1317, we propose a new, possible pathway for the biosynthesis of L-isoleucine: four intermediate steps catabolizing pyruvate to 2-oxobutanoate. According to KEGG, *S. coelicolor* is missing the initial reaction of this pathway, R-citramalate synthase. However, upon conducting a protein BLAST (BLASTP) search,[54] we uncovered a sequence similarity of 47% (E-value: 2e-168) between SCO5529 and the *cimA* gene in *Geobacter sulfurreducens*, where this pathway has been experimentally validated.[55] The same reaction has also been identified in vivo in *Cyanobacteria*, and it has been suggested that this pathway may be present in other organisms.[56] The output from our BLASTP search is available in S6, Supporting information.

We have used published transposon mutagenesis data[24] to validate and correct our model predictions for essential genes. In contradiction to the iMK1208 model predictions,[18] the

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**Biotechnology
Journal**
www.biotechnology-journal.com

transposon study indicates that SCO5626 is an essential gene encoding the ATP/UMP phosphotransferase. According to the gene annotations in iMK1208, the *cmk* gene (SCO1760) also encodes for this enzyme. However, the enzyme encoded by *cmk* is highly specific for the ATP/CMP phosphotransferase in prokaryotes.[57] Consequently, we removed it from this gene-reaction rule in iKS1317. Another apparently essential gene is SCO3894, encoding a transmembrane protein involved in the murein biosynthesis. This differs in KEGG and iMK1208, which suggest that the enzyme encoded by SCO2709 is an isozyme of the enzyme encoded by SCO3894. A BLASTP search of the *murJ* gene in *E. coli*, where the function of the encoded protein is shown,[58] show high similarity with both SCO2709 (E-value 3e-13) and SCO3894 (E-value 1e-20), and we have therefore kept SCO2709 and SCO3894 as isogenes in iMK1208. The output from our BLASTP search is available in S7, Supporting information.

The disagreements between our model predictions and in vivo observed wild-type growth phenotypes are associated with the utilization of glutamine and aspartate as carbon sources.[34] These erroneous phenotypes are also predicted in iMA789 and iMK1208, and it has previously been suggested that this lack of in vivo growth is caused by regulatory effects.[16] It is surprising that *S. coelicolor* is unable to grow on L-aspartate, because it is observed in vivo that it can utilize L-asparagine as a carbon source,[34] which is then degraded to L-aspartate. Thus, by only taking the topology of the metabolic network into account, the organism should be able to grow on L-aspartate as well.

Contrary to our predictions, *S. coelicolor* is unable to grow with glutamine as the carbon source ([34] as cited in ref. [16]). However, glutamine can be used as a nitrogen source, providing ammonium through conversion to glutamate by glutaminase intracellularly. Glutamate is further decarboxylated into γ-aminobutanoate upon uptake.[59] Borodina et al.[16] suggested that the intracellular glutamate provided by glutaminase cannot be further degraded, explaining why glutamine can function as a nitrogen source but not a carbon source. On the other hand, a $\Delta glnA$ (SCO2198) mutant, lacking glutamine synthase, can utilize glutamine as a carbon source.[39] According to optimization theory it is not possible to increase the metabolic repertoire of a GEM through a gene knockout, since a gene knockout reduces the solution space.[60] This observation is in support of the suggestion that a regulatory effect is the cause for the observed lack of utilization of glutamine as a carbon source in vivo.

The growth predictions for knockout mutants may seem less accurate (78.4%) than the growth phenotype predictions (96.5%). The erroneous predictions are mostly false positives, that is, iKS1317 predicts growth for knockout mutants that do not grow in vivo. This indicates that the model is too flexible and contains optional pathways or isogenes when an in vivo essential gene is knocked out. As is customary for constraint-based modeling, unless explicit knowledge is present, one does not account for possible differences in reaction rates between alternative enzymes or pathways.[60] Thus, a pathway providing a perfect replacement in silico may actually be too slow to support detectable growth in vivo. While it is possible to use enzyme kinetics to limit the reaction rates in a GEM, it is difficult to acquire reliable values.[61]

We determined 27 false positive predictions for the different knockout mutants, that is, 27 genes that are observed as essential in vivo but not in silico. Fourteen of the false positive predictions can be traced back to uncertainty of the nutrient environment, more specifically the available carbon and nitrogen sources present in the complex growth medium in which the transposon mutagenesis mutants were cultivated.[24] One example is the knockout of the gene *ddl* (SCO5560), encoding the D-alanine-D-alanine ligase, which is observed to be an essential gene in the transposon mutagenesis study, in contradiction to the iKS1317 predictions. However, by simply removing D-alanyl-D-alanine from the growth medium in silico, iKS1317 correctly predicts no growth for the Δddl mutant. These 14 false positives are changed to the true negative category if we assume a minimal medium with glucose and ammonium as the sole carbon and nitrogen sources, respectively. With this assumption, iKS1317 has an accuracy of 87.1% (134/153) for predicting knockout phenotypes, resulting in a total accuracy of 90% (189/210). This demonstrates a challenge with the typical use of transposon mutagenesis data for model curation and validation.

Another limiting factor is the lower bound of 1.9 kb on the length of the genes identified as essential reactions. Of the 1317 genes present in iKS1317, only 129 (9.8%) are longer than 1.9 kb. A more extensive transposon mutagenesis study with a lower threshold would increase value of the data because it could enable the evaluation of a larger number of essential genes.

We observe that many of the genes annotated to their respective enzymes, and thus reactions, are inferred from homology with similar genes in different organisms. It is in most cases not certain that these genes actually encode for the same enzymes, potentially leading to erroneous model predictions. Another possibility which will provide false-positive results is that a gene that is present in the iKS1317 model may not be expressed in vivo. There exist several methods for using transcriptomics (see Ref. [62] for a comparison) to restrict the model solution space to only include genes expressed in the chosen conditions, and such data may improve model predictions.

By using OptKnock[27] and GDLS[28] we have predicted optimal single, double, and triple reaction-knockout strategies to increase the production of acetyl-CoA through PDH. The GDLS algorithm use heuristics to perform a local search, and it is not guaranteed to find the optimal solution.[28] With the GDLS algorithm the optimal solution was found for all single and double reaction knockouts, and in two of the three environments for the triple knockouts. While OptKnock always find the optimal solution, the CPU-time for a triple reaction knockout using OptKnock and iKS1317 is about 160 h on a HPC-platform, compared to a few minutes with GDLS on an average laptop. Thus, for more than three knockouts with iKS1317, OptKnock becomes practically infeasible when all possible reactions are considered.

Production of acetyl-CoA was selected as a target for the strain engineering algorithms because it is the most important precursor for biosynthesis of polyketides. Increasing the precursor pool has previously provided increased secondary metabolite production in *S. coelicolor*[63,64] and similar strains.[65–67] Increasing the precursor pool can be combined with overexpression of the biosynthetic gene cluster encoding the

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**Biotechnology
Journal**

www.biotechnology-journal.com

pathway producing the target compound to further improve the likelihood of increased production.[68]

Several precise regulatory mechanisms are involved in the biosynthesis of antibiotics in vivo. Among these mechanisms, carbon-source interplay appears to be one of the main factors controlling secondary metabolism.[69,70] Therefore, reaction deletion strategies for overproduction of polyketide antibiotics in silico were suggested in various growth environments (Table 3A), that included different nitrogen and especially carbon sources. Both Strategy 2 and 4 seem like good suggestions for in vivo strain optimization: Strategy 2 provides the largest increase in acetyl-CoA production of these two, but Strategy 4 is more robust to different growth environments.

When comparing the suggested strain engineering strategies with experimental data, we find good agreement with the results from Huang et al.[71] that found increased production of FK506 in a Δ*gdhA Streptomyces tsukubaensis* knockout mutant. FK506 is a combined polyketide synthase and non-ribosomal peptide with acetyl-CoA as one of the main precursors, and Δ*gdhA* encodes for glutamate dehydrogenase (R00248), one of the suggested knockouts in Strategy 1, 4, 5, and 9 (Table 3B).

However, when taking the transposon mutagenesis data into account, some of the suggested strategies are in contradiction to observed knockout mutants: the *sucC* (SCO4808) and *sucD* (SCO4809) genes encoding the succinyl-CoA synthetase complex are classified as essential genes.[24] Succinyl-CoA synthetase (R00405) is one of the two knocked out reactions in both Strategy 2 and 4 (Table 3B). According to iKS1317 and protein BLAST SCO4808 and SCO4809 are not essential because SCO6585 and SCO6586 are isogenes encoding the same enzyme complex with E-values of 2e-145 and 1e-156, respectively (S8, Supporting information). Additionally, iKS1317 predicts less than 2% reduction in growth rate if succinyl-CoA synthetase is knocked out. Possible reasons for this discrepancy include: 1) The genes SCO6585 and SCO6586 may not be expressed in the cultivation media used in the transposon mutagenesis experiment[24] and 2) iKS1317 is too flexible and predicts an efficient flux rerouting when succinyl-CoA synthetase is removed which does not occur in vivo. Consequently, our computational in-depth analysis of iKS1317 serves as an example of the systems-biology science iteration paradigm, by producing further hypothesis that need experimental follow up.

Abbreviations

GDLS, genetic design through local search; GEM, genome-scale metabolic model; KEGG, kyoto encyclopedia of genes and genomes.

Supporting information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgement

Conflict of Interest

The authors declare no commercial or financial conflict of interest.

Keywords

[1] C. L. Ventola, *Pharm. Ther.* **2015**, *40*, 277.
[2] C. L. Ventola, *Pharm. Ther.* **2015**, *40*, 344.
[3] D. J. Payne, M. N. Gwynn, D. J. Holmes, D. L. Pompliano, *Nat. Rev. Drug Discov.* **2007**, *6*, 29.
[4] H. Ikeda, K. Shin-ya, S. Omura, *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 233.
[5] J. Yin, M. Hoffmann, X. Bian, Q. Tu, F. Yan, L. Xia, X. Ding, A. F. Stewart, R. Müller, J. Fu, *Sci. Rep.* **2015**, *5*, 15081.
[6] B. Bonet, R. Teufel, M. Crüsemann, N. Ziemert, B. S. Moore, *J. Nat. Prod.* **2014**, *78*, 539.
[7] K. Yamanaka, K. A. Reynolds, R. D. Kersten, K. S. Ryan, D. J. Gonzalez, V. Nizet, P. C. Dorrestein, B. S. Moore, *Proc. Natl. Acad. Sci.* **2014**, *111*, 1957.
[8] J. P. Gomez-Escribano, M. J. Bibb, *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 425.
[9] J. Berdy, *J. Antibiot.* **2005**, *58*, 1.
[10] M. Nett, H. Ikeda, B. S. Moore, *Nat. Prod. Rep.* **2009**, *26*, 1362.
[11] G. L. Challis, *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 219.
[12] D. A. Hopwood, *Streptomyces in Nature and Medicine: The Antibiotic Makers.* Oxford University Press, Oxford, UK **2007**.
[13] J. P. Gomez-Escribano, M. J. Bibb, *Microb. Biotechnol.* **2011**, *4*, 207.
[14] I. Thiele, B. Ø. Palsson, *Nat. Protoc.* **2010**, *5*, 93.
[15] E. Simeonidis, N. D. Price, *J. Ind. Microb. Biotechnol.* **2015**, *42*, 327.
[16] I. Borodina, P. Krabben, J. Nielsen, *Genome Res.* **2005**, *15*, 820.
[17] M. T. Alam, M. E. Merlo, D. A. Hodgson, E. M. Wellington, E. Takano, R. Breitling, T. S. C. (stream), *BMC Genomics* **2010**, *11*, 202.
[18] M. Kim, J. Sang Yi, J. Kim, J.-N. Kim, M. W. Kim, B.-G. Kim, *Biotechnol. J.* **2014**, *9*, 1185.
[19] http://strepdb.streptomyces.org.uk
[20] M. Kanehisa, S. Goto, *Nucleic Acids Res.* **2000**, *28*, 27.
[21] R. Caspi, R. Billington, L. Ferrer, H. Foerster, C. A. Fulcher, I. M. Keseler, A. Kothari, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, P. Subhraveti, D. S. Weaver, P. D. Karp, *Nucleic Acids Res.* **2015**, *44*, D471.
[22] L. D. Elbourne, S. G. Tetu, K. A. Hassan, I. T. Paulsen, *Nucleic Acids Res.* **2016**, *45*, D320.
[23] Z. A. King, J. Lu, A. Dräger, P. Miller, S. Federowicz, J. A. Lerman, A. Ebrahim, B. Ø. Palsson, N. E. Lewis, *Nucleic Acids Res.* **2016**, *44*, D515.
[24] Z. Xu, Y. Wang, K. F. Chater, H.-Y. Ou, H. H. Xu, Z. Deng, M. Tao, *Appl. Environ. Microbiol.* **2017**, *83*, e02889.
[25] L. Heirendt, S. Arreckx, T. Pfau, S. N. Mendoza, A. Richelle, A. Heinken, H. S. Haraldsdottir, S. M. Keating, V. Vlasov, J. Wachowiak, S. Magnusdottir, C. Yu Ng, G. Preciat, A. Zagare, S. H. Chan, M. K. Aurich, C. M. Clancy, J. Modamio, J. T. Sauls, A. Noronha, A. Bordbar, B. Cousins, A. M. Assal, CEL Diana, M. B. Guenila, Apaolaza, A. Kostromins, H. M. Le, S. Y. Ma, Ding, L. V. Valcarcel, L. Wang, J. T. Yurkovich, P. T. Young, H. S. Assal, P Lemmer El, W. A. Bryant, F. J. Aragioen Artacho, F. J. Planes,

E. Stalidzans, V. S. Maass, Alejandro, M. Hucka, M. A. Saunders, C. D. Maranas, N. E. Lewis, T. Sauter, B. Ø. Palsson, I. Thiele, R. M. Fleming, *Nat. Protoc.* **2011**, *6*, 1290.

[26] A. Ebrahim, J. A. Lerman, B. Ø. Palsson, D. R. Hyduke, *BMC Syst. Biol.* **2013**, *7*, 74.

[27] A. P. Buargard, P. Pharkya, C. D. Maranas, *Biotechnol. Bioeng.* **2003**, *84*, 647.

[28] D. S. Lun, G. Rockwell, N. J. Guido, M. Baym, J. A. Kelner, B. Berger, Galagan, G. M. J. E. Church, *Mol. Syst. Biol.* **2009**, *5*, 1.

[29] M. Scheer, A. Grote, A. Chang, I. Schomburg, C. Munaretto, M. Rother, C. Söhngen, M. Stelzer, J. Thiele, D. Schomburg, *Nucleic Acids Res.* **2010**, *39*, D670.

[30] U. Consortium, *Nucleic Acids Res.* **2018**, *46*, 2699.

[31] A. Flamholz, E. Noor, A. Bar-Even, R. Milo, *Nucleic Acids Res.* **2011**, *40*, D770.

[32] A. Bar-Even, A. Flamholz, E. Noor, R. Milo, *Biochim. Biophys. Acta (BBA)-Bioenerg.* **2012**, *1817*, 1646.

[33] N. Elad, A. Bar-Even, A. Flamholz, Y. Lubling, R. Milo, *Bioinformatics* **2012**, *2012*, 2037.

[34] D. Hodgson, Carbohydrate utilization in *Streptomyces coelicolor* A3 (2), *Ph.D. Thesis*, University of East Anglia **1980**.

[35] M. Elibol, F. Mavituna, *Process Biochem.* **1998**, *33*, 307.

[36] M. Bibb, D. Hopwood, K. Chater, T. Kieser, C. Bruton, H. Kieser, D. Lydiate, C. Smith, J. Ward, H. Schrempf, *Genetic Manipulation of Streptomyces: A Laboratory Manual*. John Innes Foundation, Norwich, UK **1986**.

[37] Y.-X. Zhang, L. Tang, C. R. Hutchinson, *J. Bacteriol.* **1996**, *178*, 490.

[38] G. Van Wezel, J. White, M. Bibb, P. Postma, *Mol. Gen. Genet. MGG* **1997**, *254*, 604.

[39] D. Fink, D. Falke, W. Wohlleben, A. Engels, *Microbiology* **1999**, *145*, 2313.

[40] F. Barona-Gómez, D. A. Hodgson, *EMBO Rep.* **2003**, *4*, 296.

[41] F. Flett, J. Platt, J. Cullum, *J. Basic Microbiol.* **1987**, *27*, 1.

[42] D.-J. Hu, D. Hood, R. Heidstra, D. Hodgson, *Mol. Microbiol.* **1999**, *32*, 869.

[43] L. Tang, C. Hutchinson, *J. Bacteriol.* **1993**, *175*, 4176.

[44] M. Fischer, C. Schmidt, D. Falke, R. G. Sawers, *Res. Microbiol.* **2012**, *163*, 340.

[45] K. Melzoch, M. J. T. de Mattos, O. M. Neijssel, *Biotechnol. Bioeng.* **1997**, *54*, 577.

[46] B. Papp, C. Pal, L. D. Hurst, *Nature* **2004**, *429*, 661.

[47] G. Hobbs, C. M. Frazer, D. C. Gardner, J. A. Cullum, S. G. Oliver, *Appl. Microbiol. Biotechnol.* **1989**, *31*, 272.

[48] X. Ou, B. Zhang, L. Zhang, G. Zhao, X. Ding, *Appl. Environ. Microbiol.* **2009**, *75*, 2158.

[49] S. D. Bentley, K. F. Chater, A.-M. Cerdeño-Tárraga, G. L. Challis, N. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.-H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M.-A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, D. A. Hopwood, *Nature* **2002**, *417*, 141.

[50] S. D. Bentley, S. Brown, L. D. Murphy, D. E. Harris, M. A. Quail, J. Parkhill, B. G. Barrell, J. R. McCormick, R. I. Santamaria, R. Losick, M. Yamasaki, H. Kinashi, C. W. Chen, G. Chandra, D. Jakimowicz, H. M. Kieser, T. Kieser, K. F. Chater, *Mol. Microbiol.* **2004**, *51*, 1615.

[51] I. Haug, A. Weissenborn, D. Brolle, S. Bentley, T. Kieser, J. Altenbuchner, *Microbiology* **2003**, *149*, 505.

[52] Z. A. King, A. Dräger, A. Ebrahim, N. Sonnenschein, N. E. Lewis, B. Ø. Palsson, *PLOS Comput. Biol.* **2015**, *11*, e1004321.

[53] This use of subsystems comes from Pathway maps for Metabolism in KEGG (http://www.genome.jp/kegg/pathway.html).

[54] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403.

[55] C. Risso, S. J. Van Dien, A. Orloff, D. R. Lovley, M. V. Coppi, *J. Bacteriol.* **2008**, *190*, 2266.

[56] B. Wu, B. Zhang, X. Feng, J. R. Rubens, R. Huang, L. M. Hicks, H. B. Pakrasi, Y. J. Tang, *Microbiology* **2010**, *156*, 596.

[57] T. Bertrand, P. Briozzo, L. Assairi, A. Ofiteru, N. Bucurenci, H. Munier-Lehmann, B. Golinelli Pimpaneau, O. Bârzu, A.-M. Gilles, *J. Mol. Biol.* **2002**, *315*, 1099.

[58] L.-T. Sham, E. K. Butler, M. D. Lebar, D. Kahne, T. G. Bernhardt, N. Ruiz, *Science* **2014**, *345*, 220.

[59] L. Inbar, A. Lapidot, *J. Bacteriol.* **1991**, *173*, 7790.

[60] B. Ø. Palsson, *Systems Biology: Constraint-Based Reconstruction and Analysis*. Cold Spring Harbor Laboratory Press (CSH Press), Cold Spring Harbor, New York, USA **2015**.

[61] R. Adadi, B. Volkmer, R. Milo, M. Heinemann, T. Shlomi, *PLOS Comput. Biol.* **2012**, *8*, e1002575.

[62] D. Machado, M. Herrgard, *PLOS Comput. Biol.* **2014**, *10*, e1003580.

[63] Y.-G. Ryu, M. J. Butler, K. F. Chater, K. J. Lee, *Appl. Environ. Microbiol.* **2006**, *72*, 7132.

[64] Y. J. Kim, J. Y. Song, M. H. Moon, C. P. Smith, S.-K. Hong, Y. K. Chang, *Appl. Microbiol. Biotechnol.* **2007**, *76*, 1119.

[65] S. Mo, Y.-H. Ban, J. W. Park, Y. J. Yoo, Y. J. Yoon, *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 1473.

[66] D. Zabala, A. F. Braña, A. B. Flórez, J. A. Salas, C. Méndez, *Metab. Eng.* **2013**, *20*, 187.

[67] H. L. Robertsen, T. Weber, H. U. Kim, S. Y. Lee, *Biotechnol. J.* **2018**, *13*, 1700465.

[68] W. S. Jung, E. Kim, Y. J. Yoo, Y. H. Ban, E. J. Kim, Y. J. Yoon, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 3701.

[69] A. Wentzel, P. Bruheim, A. Øverby, M. Ø. Jakobsen, H. Sletta, W. A. Omara, D. A. Hodgson, T. E. Ellingsen, *BMC Syst. Biol.* **2012**, *6*, 59.

[70] S. Sanchez, A. Chávez, A. Forero, Y. García-Huante, A. Romero, M. Sánchez, D. Rocha, B. Sánchez, M. Avalos, S. Guzmán-Trampe, R. Rodríguez-Sanoja, E. Langley, B. Ruiz, *J. Antibiot.* **2010**, *63*, 442.

[71] D. Huang, S. Li, M. Xia, J. Wen, X. Jia, *Microb. Cell Fact.* **2013**, *12*, 52.