Data Driven Modelling of a Chemically Defined Growth Medium for *Cupriavidus necator* H16

Christopher C. Azubuike, Martin G. Edwards, Angharad M. R. Gatehouse and Thomas P. Howard

School of Natural and Environmental Sciences, Faculty of Science, Agriculture and Engineering, Newcastle University, Newcastle-upon-Tyne, United Kingdom

***Cupriavidus necator* is a Gram-negative soil bacterium of major biotechnological interest. It is a producer of the bioplastic 3-polyhydroxybutyrate, has been exploited in bioremediation processes, and it's lithoautotrophic capabilities suggest it may function as a microbial cell factory upgrading renewable resources to fuels and chemicals. It remains necessary however to develop appropriate experimental resources to permit controlled bioengineering and system optimisation of this microbial chassis. A key resource for physiological, biochemical and metabolic studies of any microorganism is a chemically defined growth medium. Here we use 1 mL micro-well cell cultures, automated liquid handling and a statistical engineering approach to develop a model that describes the effect of key media components and their interactions on *C. necator* culture cell density. The model is predictive and was experimentally validated against novel media compositions. Moreover, the model was further validated against larger culture volumes at 100 mL and 1 L volumes and found to correlate well. This approach provides valuable and quantifiable insights into the impact of media components on cell growth as well as providing predictions to guide culture scale-up.**

Cupriavidus necator H16 | Design of Experiments (DOE) | chemically defined media | 3-polyhydroxybutyrate (PHB)
Correspondence: *thomas.howard@newcastle.ac.uk*

Introduction

The Gram-negative, non-spore forming soil bacterium, *Cupriavidus necator* H16, is of biotechnological importance principally due to its ability to accumulate >80% of its dry cell weight as polyhydroxyalkanoate (PHA) - a biodegradable polymer and an alternative to petroleum-based polymers [1, 2]. The accumulation of PHA (specifically 3-polyhydroxybutyrate (3-PHB)) by the bacterium is a carbon conservation mechanism. When nitrogen, oxygen, or phosphorus becomes limiting to cell growth the bacterium diverts excess carbon to 3-PHB. This carbon store supports growth when conditions improve [3–6]. *C. necator* H16 is also chemolithoautotrophic, with the ability to use CO_2, and formate or H_2 as carbon and energy sources to support cellular metabolism [7]. In the absence of oxygen it can use alternative electron acceptors (NO_3^- and NO_2^-) to carry out anaerobic respiration by denitrification [8, 9]. Metabolic engineering of *C. necator* H16 has been demonstrated through the introduction of pathways for the biosynthesis of alcohols [6, 7, 10, 11], fatty acids [12–14], alkanes [15] and enzymes [16] under both heterotrophic and autotrophic growth conditions. The production of branched chain alcohols by *C. necator* H16 through electricity powered cellular synthesis (i.e. electrosynthesis), demonstrates the value of this bacterium

as a chassis capable of exploiting renewable feedstock for the biosynthesis of valuable products [7]. More widely, the genus is known to encode genes facilitating metabolism of environmental pollutants such as aromatics and heavy metals making it a potential microbial remediator [8, 9, 17–19].

As with any bacterium, the development of *C. necator* as an industrial chassis requires appropriate tools for studying and engineering the organism. One of these resources is the availability of a characterised, chemically defined growth medium. Chemically defined media are important to enable experimental reproducibility, to reliably characterise the genetics of the organism, to determine genotype by environment interactions, and to facilitate fundamental research of bacterial physiology that underpins bioengineering efforts. While different chemically defined media have been deployed for the cultivation of *C. necator* H16 [3, 10–12, 20–22] there is no consensus regarding the components that are required, the concentration of each component, or how each component interacts to affect growth of the bacterium (SI Table 1).

Design of Experiments (DOE) is an iterative, empirical approach that systematically explores the relationship between input variables (factors) and output variables (responses). The approach yields a structured set of data that can be used to build statistical models employed in understanding or optimising system performance. These statistical models can be validated against prior knowledge, internal statitical methods or ultimately - by their ability to predict responses from new combinations of factors. Despite its early origins within the biological sciences, DOE is not a common-place method for life scientists. Increasing availability of laboratory automation and high throughput technologies may be changing this. DOE has found use in the optimisation of metabolic pathways [23, 24], cell-free protein synthesis reactions [25] and codon-use algorithms [26]. It has been applied to the study of genotype-by-genotype and genotype-by-environment interactions in yeast [27] and in re-purposing enzyme activities [28].

Here we employ a statistical engineering approach to build a data driven model that can accurately predict *C. necator* H16 growth responses to a range of media formulations. The model highlighted different formulations that allow reproducible and robust growth of *C. necator* H16 with the minimal concentration of each component and allowed us to identify and understand interactions between components of the media. Finally, the model allowed the learning from the small scale (1 mL) to be applied at larger volumes (100 mL and 1

L). Understanding the impact of each component of a chemically defined medium on *C. necator* growth is a fundamental tool for controlled exploration of the biotechnological potential of this important bacterium.

Results

Identifying main ingredients in chemically-defined media that affect the growth of *C. necator*. A comparison of the literature for the use of defined media for *C. necator* growth identified variety in both the nature and concentrations of macroelements (C, N, P, Ca, Mg, and S) and trace elements (Cu, Zn, Fe and Mn) required for robust cell growth (Table SI 1). An initial scoping trial was carried out using fructose, glucose, glycerol or sucrose to identify a principle carbon source to be used for subsequent work, and to determine the range of concentrations to be tested. Four scoping trials were performed one each at low and high concentrations of media components, and two trials at the midpoint values between the two extremes. Cultures were grown in 1 mL volumes in a 48-well plate format for 120 h, at 30°C, 200 rpm. We confirmed that optical density (OD_{600nm}) was an appropriate surrogate for determining *C. necator* cell growth (Fig. SI 1A). At the ranges tested glucose, glycerol and sucrose supported little or no growth of *C. necator* (Fig. SI 1B-D), whilst fructose supported high growth except at the lowest concentrations (Fig. 1). The OD_{600nm} for the two midpoint experiments demonstrated a peak at 72 h followed by a plateau. Growth obtained from optical density measurements at $OD_{600\ nm}$ correlated well with the number of viable cells (CFU/mL). From these scoping trials it was established that fructose would be our principle carbon source, OD_{600nm} was an appropriate measure of cell growth, growth assays in 1 mL volumes in a 48-well plate format were appropriate for subsequent experiments, and recording $OD_{600\ nm}$ at 72 h provides a good balance between measuring growth rate and peak culture density.
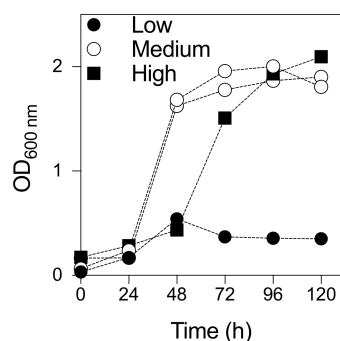

Fig. 1. Scoping trials. Growth of *C. necator* H16 in 48-well format with fructose as the carbon source. Each experiment was carried out at low, medium (*n* = 2) or high concentrations of each media component (details can be found in Table SI 2).

We next identified key factors that might influence OD_{600nm} at 72 h. A definitive screening design (DSD) array was built based on 10 media components (Table 1). DSDs are highly efficient experimental designs in which all main effects can be estimated independently of other main effects and all pos-

sible two-way interactions. The requisite variant media compositions were assembled in a 48-well plate using a liquid-handling robot. Cell growth in these variant media was monitored as above. The resulting data were analysed using both the *Definitive Screening* and the *Two Level Screening* platforms in JMP Pro 13.0. Both analyses indicated that fructose, $CaCl_2$ and amino acids contributed positively to growth, while Na_2HPO_4 and trace elements contribute negatively to growth. Factors such as NaH_2PO_4, K_2SO_4, $MgSO_4$, NH_4Cl and vitamins were not found to have statistically significant effects under the conditions tested (Fig. 2). These analyses also highlighted several two-way interactions that may influence growth responses, however, while DSDs are efficient arrays for de-aliasing main effects from other main effects and from second-order interactions, second-order interactions remain partially aliased with themselves. For this reason attributing effects to specific second-order interactions was not attempted at this stage.

Augmentation of the data set. A definitive screening design can force many of the data-points collected to the edges of the design space. For this reason we ran a second DSD array (Table SI 3) in which the concentration ranges of the components were guided by data from DSD1 and the factors under investigation were restricted to those that were highlighted as significant in DSD1. Examining the combined data for DSD1 and DSD2 did indeed confirm that the extreme concentrations of some of the components were detrimental to cell growth. For example, when the concentrations of fructose were at the highest and lowest values (0.5 and 40 g/L) the $OD_{OD600nm}$ 72 h were both lower and more variable than when fructose was restricted to between 5 and 25 g/L (Fig. 3). This indicates that maintaining fructose between 5 and 25 g/L is key to establishing robust and reliable growth. Likewise, adjustments were made to the concentration ranges of the amino acids (5 and 20 ml/L), $CaCl_2$ (0.1 and 0.459 g/L) and Na_2HPO_4 (0.1 and 3.05 g/L). Other factors not identified as statistically significant were kept at the midpoint from DSD1, with the exception of NH_4Cl which was set at the
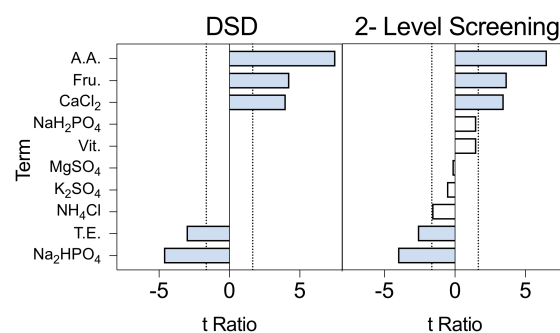

Fig. 2. Definitive Screening Design array analysis. Definitive screening and 2-Level screening of data within DSD1 were performed. The comparative lengths of the t-ratios for each factor and factor interaction are shown. Bars extending to the right have a positive impact on the model, those extending to the left have a negative impact on the model. Terms deemed significant for model projection are highlighted (blue). Analysis is based on two replicated arrays (Table 1). The dashed vertical lines indicate a term that is significant at the 0.10 level.

Table 1. Definitive screening design array. A definitive screening design was developed to assess the impact of 10 ingredients found within the chemically defined media. All concentrations are in g/L except trace elements, amino acids and vitamins which are ml / L. Trace element working concentration contained (g/L): 15 $FeSO_4.7H_2O$, 2.4 $MnSO_4.H_2O$, 2.4 $ZnSO_4.7H_2O$, and 0.48 $CuSO_4.5H_2O$. A 100 × stock amino acid mix contained (g/L): 12.9 arginine, and 10 each of histidine, leucine and methionine. A 1000 × vitamin stock contained (g/L): 0.1 pyridoxine, 0.02 folic acid, 0.05 each of thiamine, riboflavin, niacin, pantothenic acid and nicotinamide. Abbreviations: Fru., fructose; T.E. Trace element mixture; A.A., amino acid mixture; Vit., vitamin mixture. The DSD was performed in replicate.

Media	Fru.	NaH_2PO_4	Na_2HPO_4	K_2SO_4	$MgSO_4$	$CaCl_2$	NH_4Cl	T.E.	A.A.	Vit.	OD1	OD2
1	20.25	0.1	0.1	0.01	0.01	0.01	0.01	0.01	0.05	0.01	1.06	1.37
2	20.25	6	6	4.8	2.8	0.8	4.8	4.8	19	4.8	1.14	1.62
3	0.5	3.05	0.1	4.8	0.01	0.01	0.01	4.8	19	4.8	0.52	0.47
4	40	3.05	6	0.01	2.8	0.8	4.8	0.01	0.05	0.01	0.09	0.78
5	0.5	6	3.05	0.01	2.8	0.01	0.01	0.01	19	4.8	0.77	0.71
6	40	0.1	3.05	4.8	0.01	0.8	4.8	4.8	0.05	0.01	-	-
7	40	6	0.1	2.405	2.8	0.01	4.8	4.8	19	0.01	0.93	0.69
8	0.5	0.1	6	2.405	0.01	0.8	0.01	0.01	0.05	4.8	0.11	0.28
9	0.5	6	0.1	4.8	1.405	0.01	4.8	0.01	0.05	0.01	0.80	0.64
10	40	0.1	6	0.01	1.405	0.8	0.01	4.8	19	4.8	1.62	1.48
11	0.5	6	6	0.01	2.8	0.405	0.01	4.8	0.05	0.01	0.29	0.32
12	40	0.1	0.1	4.8	0.01	0.405	4.8	0.01	19	4.8	1.86	1.41
13	0.5	6	6	4.8	0.01	0.8	2.405	0.01	19	0.01	0.41	0.55
14	40	0.1	0.1	0.01	2.8	0.01	2.405	4.8	0.05	4.8	-	0.01
15	0.5	0.1	6	4.8	2.8	0.01	4.8	2.405	0.05	4.8	0.17	0.43
16	40	6	0.1	0.01	0.01	0.8	0.01	2.405	19	0.01	1.69	1.67
17	0.5	0.1	0.1	4.8	2.8	0.8	0.01	4.8	9.525	0.01	0.17	0.28
18	40	6	6	0.01	0.01	0.01	4.8	0.01	9.525	4.8	-	0.19
19	0.5	0.1	0.1	0.01	2.8	0.8	4.8	0.01	19	2.405	0.48	0.70
20	40	6	6	4.8	0.01	0.01	0.01	4.8	0.05	2.405	-	0.04
21	20.25	3.05	3.05	2.405	1.405	0.405	2.405	2.405	9.525	2.405	1.76	1.70
22	0.5	0.1	6	0.01	0.01	0.01	4.8	4.8	19	0.01	0.86	0.71
23	0.5	6	0.1	0.01	0.01	0.8	4.8	4.8	0.05	4.8	0.10	0.19
24	40	6	0.1	4.8	2.8	0.8	0.01	0.01	0.05	4.8	1.64	1.56
25	40	0.1	6	4.8	2.8	0.01	0.01	0.01	19	0.01	0.20	0.23

lowest value because there was some evidence of a negative impact from the 2-level factor analysis (Fig. 2).

Examination of the distributions for DSD2 revealed that the media compositions that generated the highest cell densities (OD_{600nm} >2.0) were associated with low trace element concentrations and high amino acid content (Fig. SI 2). We therefore probed the putative interaction between amino acids and trace elements using a medium that was formulated to permit robust cell growth over 72 h. It was observed that the absence of trace elements did not adversely affect growth of *C. necator* but that the absence of amino acids did (Fig. 4A). Interestingly, simultaneous exclusion of both amino acids and trace elements resulted in cell densities comparable with the control. We hypothesized that in the absence of one or more of the amino acids (methionine, histidine, leucine and/or arginine), one or more of the components of the trace elements ($CuSO_4$, $FeSO_4$, $MnSO_4$ and/or $ZnSO_4$) inhibits growth of *C. necator*. This was tested first by formulating media without amino acids and withdrawing each trace element in turn (Fig 4B). Under these conditions, media without $CuSO_4$ but containing other trace elements resulted in growth comparable to the control, while all three formulations that contained $CuSO_4$ had reduced cell densities. These observations support the hypothesis that $CuSO_4$ - in the absence of amino

acids - inhibits *C. necator* growth. To determine which amino acid(s) interacts with $CuSO_4$, a series of experiments were performed in which each amino acid was excluded in formulations with and without $CuSO_4$. The first observation was that at high concentrations of amino acids the presence or absence of $CuSO_4$ did not affect growth (Fig 4C). The second observation was that at lower concentrations of amino acids growth was partially suppressed in the presence of $CuSO_4$ but that this was exacerbated in the absence of $CuSO_4$. $CuSO_4$ is therefore an important medium component and cannot simply be excluded from formulations. Similar growth responses were seen in experiments in which either methionine or leucine were excluded (Fig 4D and E). If arginine, or most notably histidine, were removed from the medium, then the presence of $CuSO_4$ impaired growth of *C. necator* (Fig 4F and G). For media lacking histidine this effect was also observable when all other amino acid levels were kept high (Fig. 4G). From this data we conclude two points: first, that histidine protects against the inhibitory effects of copper, but second, $CuSO_4$ is an essential component of the media whose absence retards growth when amino acid content is restricted. Both $CuSO_4$ and histidine concentrations must therefore be balanced for robust growth.

Final data augmentation. With a greater understanding of which components and concentrations are required to construct a medium supporting robust growth we re-investigated the roles that NaH_2PO_4, K_2SO_4, $MgSO_4$ and NH_4Cl have on the system. In the original DSD these factors were not identified as being statistically significant but those experiments were run under conditions in which key components (e.g. fructose and $CaCl_2$ (Fig. 3) were at settings that have since been identified as resulting in poor or unreliable growth responses. We therefore wished to re-evaluated these factors under conditions in which Na_2HPO_4, $CaCl_2$, trace elements and amino acid concentrations were not disruptive to cell growth (Table SI 4). Fructose was set at either 5 g/L or 25 g/L. The results indicated that these components did indeed affect growth rate when primary factors are not restricting growth (Fig. 5). Some of the suggestions from the original DSD were confirmed. Most notable was the observation that increasing NH_4Cl had a detrimental affect on culture cell density. NaH_2PO_4 and K_2SO_4 had some detrimental impact if concentrations were low, whilst $MgSO_4$ concentrations were not significant. These results were true at both high and low concentrations of fructose.

Finally, it is important to draw a distinction between what is statistically significant and what is essential. We observe little growth in formulae lacking $MgSO_4$ yet $MgSO_4$ is not deemed statistically significant. This is because under our experimental conditions, all concentrations of $MgSO_4$ tested are in excess and do not limit to growth. There is therefore

scope to reduce $MgSO_4$ concentrations if this was desirable.

Modelling the media formula-growth response landscape. At this stage, 64 different variant formulations have been experimentally assessed in duplicate across three dif-
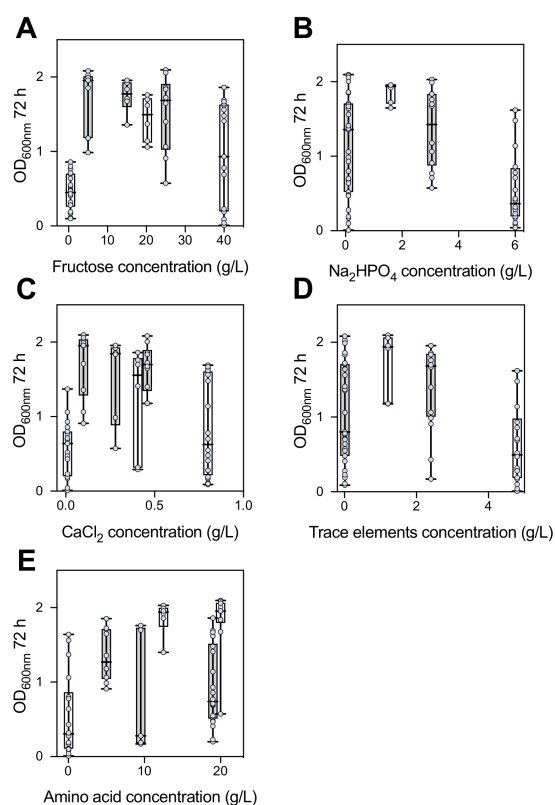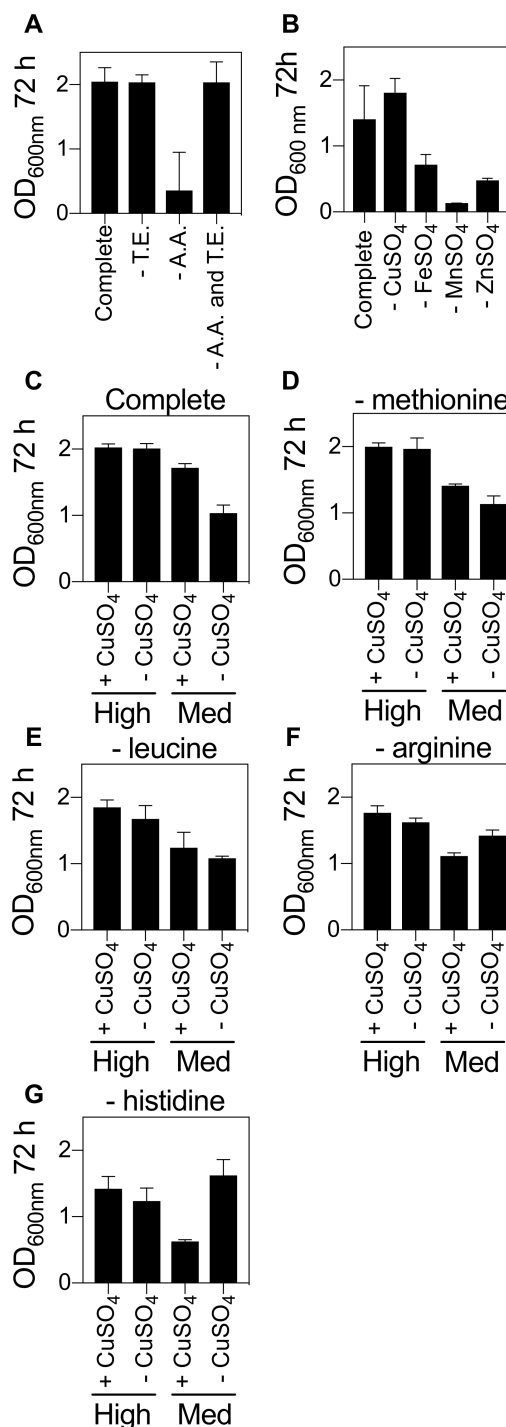

Fig. 4. Interactions between trace elements and amino acids. A. *C. necator* grown in a complete, chemically defined medium with either amino acids, trace elements or both amino acids and trace elements excluded. **B.** *C. necator* grown in a complete medium with each of the four trace elements excluded. **C-G** *C. necator* grown in a complete, chemically defined medium with or without $CuSO_4$ and with each of the four amino acids excluded. Error bars are S.E. Mean, $n = 2$.

Fig. 3. Combined data for DSD1 and DSD2 for key media components. A. Fructose. **B.** Na2HPO4 **C.** CaCl2 **D.** Trace elements **E.** Amino acids.
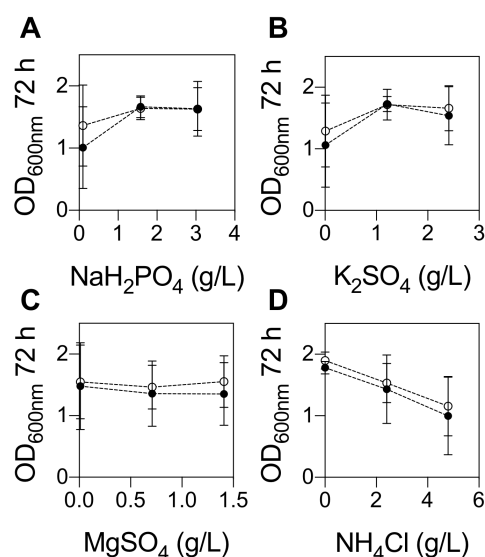
Fig. 5. Re-examination of non-significant media components. The impact of components not deemed statistically significant in DSD1 were re-examined under less extreme conditions. These were **A.** NaH_2PO_4 **B.** K_2SO_4 **C.** $MgSO_4$ and **D.** NH_4Cl. Experiments were conducted at 5 g/L (open circles) or 25 g/L (closed circles) fructose. Error bars are S.E. Mean, $n \geq 6$

ferent experiments (25, 13 and 26 experimental runs respectively). We then built a statistical model, trained against this data set, that could describe our understanding of how the cell cultures respond to changes in media composition and predict performance in novel formulations. We performed two-level screening on all 128 runs. This identified a number of factors and factor interactions deemed significant for model projection. The screening did not highlight fructose or K_2SO_4 (as these had been at held concentrations that did not significantly impact growth during much of DSDS2 and DSD3) but these terms were included manually in the model as we knew they were important factors. These terms were used to construct a standard least squares model. The least squares model was able to describe the relationships within the data with good accuracy (Fig. 6A). Though the model was internally consistent it was important to know if it could predict OD_{600nm} at 72 h in formulations it had not encountered during model training. We assessed 16 new formulations with sampling biased towards media formulations that were predicted to be in the top 25% of media performance. Each of these was assessed in triplicate and the resulting OD_{600nm} compared to predicted OD_{600nm} (Fig. 6B,C). As predicted by the model, all of the new formulations fell within the upper quartile of formula performance.

The model allowed us to visualise phenomenon observed during the data-collection phase of the investigation. For example, it visualises the interaction between amino acids (specifically histidine) and trace elements (specifically copper) that was elucidated in Fig. 4. It shows that low and high concentrations of trace elements are detrimental to growth (low concentrations are to be avoided as these are essential to growth, while high concentrations may be toxic) and that increasing the amino acid concentration can help mitigate the
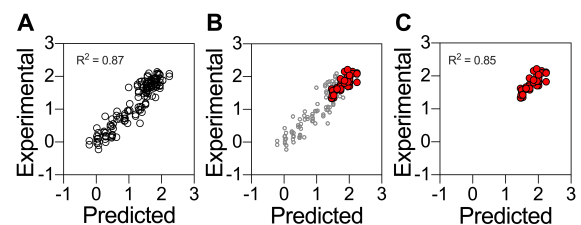

Fig. 6. Experimental validation of model predictions in 48-well plate format. A. Model predicted values plotted against experimental data for least squares model. **B.** New experimental data for 1 mL culture volumes overlaid against the original least squares model. Open circles show the training data set, red circles are new data. **C.** New experimental data for 1 mL cultures plotted on their own against the model predicted values. Three replicates were assessed for each prediction.

inhibitory effects of high concentrations of trace elements (Fig. 7A). The greater the concentration of trace elements included the higher the amino acid concentration needs to be. Nevertheless, increasing the concentration of amino acids impacts in other areas. When fructose is low, increasing the concentration of amino acids increases cell density (Fig 7B). Given that the amino acids present in the media are the only alternative source of carbon for growth, raising the concentration of amino acids too high will impact on interpretation of experiments designed to examine carbon utilisation. A balance therefore needs to be struck between fructose, amino acids and trace element concentrations. The model also visualises interactions between Na_2HPO_4 and fructose (Fig. 7C). Greater concentrations of Na_2HPO_4 results in lower cell densities - an effect that can be partially offset by decreasing fructose concentrations. Finally, we had previously observed a negative effect of increasing NH_4Cl concentrations (Fig. 5D). The model indicates that this can be mitigated by increasing the K_2SO_4 concentration (Fig. 7D). Understanding the interactions between media components is vital for predictions of culture performance and allows the experimenter to alter media formulations for different experimental goals.

Model validation at greater volumes. We conducted two further model validation tests. We transferred these formulations to both 100 mL scale in shake flasks and 1 L batch fermentation in a bioreactor. Shake flask cultivations were carried out under identical conditions (30°C, 200 rpm) over 72 h periods in 250 mL baffled and non-baffled flasks, each containing 100 mL of medium. The formula for each cultivation was selected randomly from an L32 fractional factorial design of experiment (Fig SI 3). Growth (OD_{600nm}) for each formula in both types of flasks were strikingly similar at every interval throughout the cultivation period. Most importantly, the growth rank for baffled and non-baffled flasks were identical. Spearman's correlation showed a significant ($p < 0.05$) relationship between predicted and actual ranks for both flask types, with 0.867 correlation co-efficient ρ (Fig. 8A). Next, three formulae from shake flask cultivation together with two additional formulae were cultivated in 1 L bioreactor (Fig SI 4). Similarly, growth rank was predictable, with significant relationship between predicted rank and actual growth rank ($p = 0.900$) (Fig. 8B). During the culti-
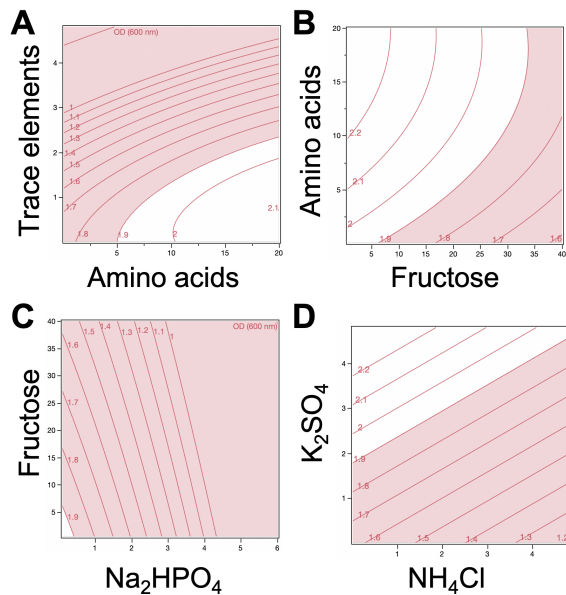
Fig. 7. Interactions of components of the defined media described by the least squares model. A. Amino acid and trace element interactions. **B.** Fructose and amino acid interactions. **C.** Disodium phosphate and fructose interactions. **D.** Ammonium chloride and potassium sulphate interactions. Each panel represents a two way-interaction. Red zones are areas where OD_{600nm} at 72 h fails to reach 1.9, and white zones where it surpasses 1.9. Each contour grid line represents an OD_{600nm} of 0.1. All other media components were kept at concentrations permitting high growth.

Fig. 8. Experimental validation of model predictions at shake-flask and bioreactor scale. A. Spearman's rank correlations between model predicted rank and experimental data for 100 mL culture volumes **B.** Spearman's rank correlations between model predicted rank and experimental data for 1 L bioreactor batch flask cultures.

vations, there were no significant changes in bioprocess parameters. Constant agitation at 200 rpm, with 1 vvm airflow rate was sufficient to maintain dissolved oxygen (dO_2) above 20%. Although no base was added in all cultivations, the pH of the media did not drop below 4.5, the set point. There was no significant relationship ($p = 0.233$; $p = 0.520$) between growth at 72 h and pH of media prior to inoculation or after growth.

Discussion

We developed a model trained against a structured data-set for the cultivation of *C. necator* H16 in a chemically defined medium with fructose as the primary carbon source. Our approach identified significant growth factors and their effects on culture density at 72 h. Fructose, glucose, glycerol and sucrose were used in the preliminary phase with fructose considered as the best carbon source supporting robust growth under heterotrophic conditions. While *C. necator* H16 has been reported to have broad substrate range its ability to utilize carbohydrates as a carbon source during heterotrophic growth appears to be limited to fructose and *N*-acetylglucosamine (8, 17, 29). Utilisation of fructose by *C. necator* H16 is most likely to occur via substrate import via an ATP-binding cassette (ABC-type) transporter, followed by catabolism via 2-keto-3-deoxy-6-phosphogluconate (KDPG) and the Entner-Doudoroff pathway. The responsible genes, notably a putative regulator (*frcR*), ribosome transporters (i.e. *frcA*, *frcC* and *frcB* orthologs in *Escherichia coli* and *Ralstonia solanacearum*) and other essential genes facilitating such metabolism are located on chromosome 2 in-
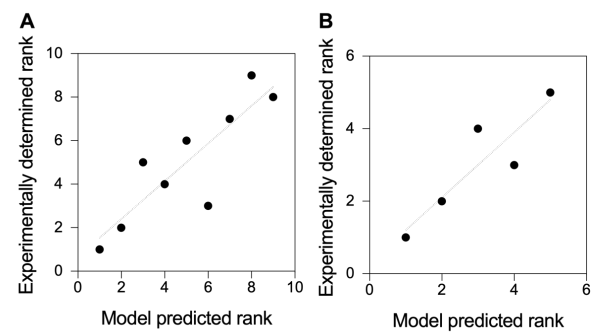
side gene clusters for glucose, 2-ketogluconate, and glucosamine catabolism (17). In contrast, phoshphofructokinase and 6-phosphogluconate dehydrogenase, key enzymes of the Embden-Meyerhoff-Parnas and oxidative pentose phosphate pathways, respectively, appear to be absent from the *C. necator* H16 genome. It is perhaps therefore not surprising that glucose supported little or no growth yet it has been reported that prolonged cultivation (> 70 h) with glucose as the sole carbon source resulted in a mutant that was able to utilize glucose (30). The mutant acquired such ability by mutating the *N*-acetylglucoasmine phosphotransferase system, which when deleted led to inability to utilize glucose (30, 31). Glycerol supported low growth of *C. necator* H16 especially in the medium forcing during the scoping trial. Such low growth was attributed to oxidative stress resulting from high levels of reactive oxygen species (ROS) formed as a result of elevated level of hydrogenases, leading to cell destruction by damaging DNA and other cell components (32).

Formulations of chemically defined media previously described for the cultivation of *C. necator* H16 are typically prepared without amino acids (3, 7, 11, 12, 22, 33). In this study, bacterial growth is improved when a small number of amino acids (arginine, histidine, leucine and methionine) are included in the medium. The data suggest that they serve as preferred sources of nitrogen, compared to NH_4Cl. Their inclusion resulted in a shorter lag phase compared to media lacking amino acids. Although the effect of histidine is greater than methionine, arginine and leucine, their action was synergistic. Interestingly, the model indicated that there was an interaction between fructose and amino acid content that indicates the carbon : nitrogen ratio balance must be maintained, irrespective of the actual values. Growth reductions are less under low amino acid and low fructose concentrations (5 ml/L and 5 g/L, respectively) compared to that under low amino acid and high fructose concentrations (5 ml/L and > 20 g/L, respectively).

Metal ions, especially divalent cations of d-block transition metals, are important for bacterial growth where they act as metalloenzyme cofactors in living cells. Their presence in high concentrations however, tends to be detrimental to cells (34). Copper is known to play diverse structural and catalytic

functions owing to its ability to exist either in a reduced (Cu^+) state with affinity for thiol and thioether groups, or an oxidized (Cu^{2+}) state with more likely coordination for oxygen or imidazole nitrogen group of amino acids including histidine [35]. Bacteria have been reported to respond to high levels of copper by one of three families of metalloregulatory repressors: CopY, CsoR, and CueR. Under high concentration, an integral membrane Cu^+ transporter (P1B-type ATPase) - characterised by histidine-rich domains, and conserved cytosine/histidine motifs within specific transmembrane domains - exports copper from the cytoplasm into the periplasm where in Gram-negative bacteria further detoxification and exportation are carried out by other related enzymes [35]. Further, studies have reported that copper toxicity affected sugar (glucose) utilization in different microorganisms, and its growth inhibition on fungi (*Neurospora crassa* and *Saccharomyces cerevisiae*) was due to impaired/suppression of histidine biosynthesis [36, 37]. In general, the toxic effect of copper appears to be enhanced under histidine limitation or impaired histidine biosynthesis, and is neutralized by addition of histidine.

This study provides insight into the impact of media formulations on growth and cell density of *C. necator* H16. Our growth model shows that the bacterium is able to grow to high cell density with minimal concentration of each growth component, provided the concentration of amino acids and trace elements are balanced. When cultivated under our formulae, growth peaked at 72 h and remained stationary after a further 20 h. Fructose, amino acids, and $CaCl_2$, and interactions between amino acids and fructose, fructose and $CaCl_2$, and amino acids and trace elements have significant positive effects on cell density, whilst Na_2HPO_4, and trace elements, and interactions between fructose and Na_2HPO_4, have negative effects. Addition of amino acids shortens the lag phase of growth and results in reproducible high cell density over repeated experiments. Besides fructose, only magnesium is considered essential for growth, whilst amino acids and K_2SO_4 are considered important. Thus, a defined medium supporting high cell density of *C. necator* H16 lacking some components (i.e. Na_2HPO_4, $CaCl_2$, NH_4Cl, and trace elements) can still result in high cell density comparable to formulae with all components. Both the information provided and the statistical approaches taken in this study will inform further efforts aimed at optimising other *C. necator* responses. This includes the biosynthesis of polyhydroxyalkanoate, platform chemicals, proteins and other products as well as future optimisation of growth under lithoautotrophic conditions with CO_2 and molecular hydrogen serving as carbon and energy sources, respectively.

Materials and Methods

Bacterial strains. *Cupriavidus necator* H16 (DSM 428) was obtained from Deutsche Sammlung von Mikrooganismen und Zellkulturen GmbH, German Collection of Microorganisms and Cell Cultures (DSMZ). The bacterial strain was resuscitated on nutrient agar (peptone 5 g/L, and meat extract 3 g/L) according to the supplier's instructions, and incubated at 30°C for 48 h.

Chemicals. Carbon sources (glucose, fructose, glycerol and sucrose), salts (with exception of $MgSO_4.H_2O$ and NH_4Cl), trace metals, amino acids (histidine, leucine and arginine) and some vitamins (thiamine, niacin, pantothenic acid) were obtained from Sigma-Aldrich. The remainder of the vitamins, $MgSO_4.H_2O$, NH_4Cl and methionine were obtained from Duchefa Biochemie B.V., BDH chemicals and Formedium, respectively.

Preparation of media. Solutions of glucose, fructose, sucrose, vitamins, amino acid and each trace metal solution were filter sterilized through 0.22 μm filter. Glycerol, NaH_2PO_4, Na_2HPO_4, $MgSO_4.H_2O$, NH_4Cl, K_2SO_4, $CaCl_2.2H_2O$ solutions were autoclaved. The trace solution was made of $CuSO_4.5H_2O$ (dissolved in 0.1M HCl), $FeSO_4.7H_2O$ (freshly prepared during each trace reconstitution from individual stock), $MnSO_4.H_2O$, and $ZnSO_4.7H_2O$. Amino acid stock solution contained: arginine, histidine, leucine, and methionine, while vitamin stock solutions contained: folic acid, niacin, nicotinamide, pantothenic acid, pyridoxine, riboflavin and thiamine. Subsequently, medium components were added from individual stock solutions except for trace elements which were added from reconstituted working solution. Forty-eight well plates were used in all trials, and medium reconstitution in each well was carried out using an automated liquid handling system (Eppendorf epMotion M5073). All stock solutions were prepared using water as a solvent, and were further diluted in sterile distilled water unless stated otherwise.

Inoculate preparation and growth measurement. For each experiment, 48 h colonies from LB agar were washed twice in sterile distilled water and diluted to a working inoculate concentration in the range of 108 cfu/ml. The inoculate was further diluted 1:100 in wells containing medium. Plates were incubated at 30°C, 170 rpm. Optical density (OD) at 600 nm was measured every 24 h using a Varioskan LUX[TM] Multimode Microplate reader (Thermo Scientific).

Batch cultivation in a bioreactor system. Large-scale cultivations were carried out in a batch mode with 1 L chemically defined media contained in 2 L capacity fermentors (Applikon ADI fermentation system). During the cultivations, pH was maintained above 4.5, temperature at 30°C, agitation at 200 rpm. Dissolved oxygen (dO_2) was maintained above 20 % [1, 6] with airflow at 1 vvm (volume of air per volume of medium). No anti-foam agents or base were added during cultivation. Starter culture media were of the same formulations with that used in fermenters, and were prepared by growing 100 mL culture in 250 mL non-baffled flasks to late exponential growth at 30°C, 200 rpm. Following polarization and calibration of the dO_2 probe, fermenters were inoculated with 10 mL (48 h) starter culture and cultivations were monitored on-line and off-line over 72 h. Samples were taken every 24 h for off-line OD_{600nm}.

Data analyses. Experimental designs were created using JMP Pro statistical software (version 13.0) and data from each experiment was analysed using the same software. Graphics were generated using GraphPad Prism 7.0.

Supporting Information

- **Table SI 1. Previously described chemically defined media for the growth of *C. necator*.** All concentrations are g/L unless otherwise indicated.

- **Table SI 2. Scoping Experiment.** A scoping trial was developed to assess the impact of 10 basic ingredients found within the chemically defined media. All concentrations are in g/L except trace elements, amino acids and vitamins which are ml / L. Abbreviations: T.E. Trace element mixture; A.A., amino acid mixture; Vit., vitamin mixture. Carbon: fructose, glucose, glycerol or sucrose. The medium trial was performed in duplicate..

- **Table SI 3. Definitive Screening Design 2.** A second definitive screening design was developed to assess the impact of the five ingredients identified as important to the chemically defined media. All concentrations are in g/L. Abbreviations with exception of T.E. and A.A, which are in ml/L: Fru., fructose, T.E., Trace element mixture; A.A., amino acid mixture. The DSD array was performed in duplicate.

- **Table SI 4. Definitive Screening Design 3.** A final set of experiments were performed to re-assess the impact of components not deemed significant in the first DSD. All concentrations are in g/L. Abbreviations: Fru., fructose, T.E. Trace element mixture; A.A., amino acid mixture. The DSD array was performed in duplicate.

- **Fig. SI 1. Growth characteristics of *C. necator* from initial scoping experiments. A.** Colony forming units against OD_{600nm}. Growth of *C. necator* on **B.** Glucose, **C.** Glycerol, and **D.** Sucrose as sole carbon source.

- **Fig. SI 2. Distribution of data for Definitive Screening Design 2.** Highlighted in blue cross-hatch are the settings that resulted in the greatest OD_{600nm} at 72 h for DSD2. Fructose, Na_2HPO_4 and $CaCl_2$ may be at either the highest or lowest concentrations, whereas trace elements and amino acid concentrations are found at the lowest or highest concentrations respectively.

- **Fig. SI 3. 100 mL validation experiment.** Nine different formulations were assessed at 100 mL culture volumes. For each row, the model predicted rank is shown on the left hand side, media settings for each component are shown in red underneath, the measured $OD_{(600nm)}$ 72 h for two replicate experiments and experimentally determined rank on the right hand side.

- **Fig. SI 4. 1 L validation experiment.** Five different formulations were assessed at 1 L culture volumes. For each row, the model predicted rank is shown on the left hand side, media settings for each component are shown in red underneath, the measured $OD_{(600nm)}$ 72 h for two replicate experiments and experimentally determined rank on the right hand side.

Bibliography

1. E. Grousseau, J. Lu, N. Gorret, S. E. Guillouet, and A. J. Sinskey. Isopropanol production with engineered cupriavidus necator as bioproduction platform. *Appl Microbiol Biotechnol*, 98(9):4277–90, 2014. ISSN 1432-0614 (Electronic) 0175-7598 (Linking). doi: 10.1007/s00253-014-5591-0.

2. M. Schmidt, J. L. Ienczak, L. K. Quines, K. Zanfonato, W. Schmidell, and G. M. F. de Aragao. Poly(3-hydroxybutyrate-co-3-hydroxyvalerate) production in a system with external cell recycle and limited nitrogen feeding during the production phase. *Biochemical Engineering Journal*, 112:130–135, 2016. ISSN 1369-703x. doi: 10.1016/j.bej.2016.04.013.

3. K. K. Aneja, R. D. Ashby, and D. K. Solaiman. Altered composition of *Ralstonia eutropha* poly(hydroxyalkanoate) through expression of pha synthase from *Allochromatium vinosum* atcc 35206. *Biotechnol Lett*, 31(10):1601–12, 2009. ISSN 1573-6776 (Electronic) 0141-5492 (Linking). doi: 10.1007/s10529-009-0052-z.

4. C. J. Brigham, C. F. Budde, J. W. Holder, Q. Zeng, A. E. Mahan, C. Rha, and A. J. Sinskey. Elucidation of beta-oxidation pathways in *Ralstonia eutropha* h16 by examination of global gene expression. *J Bacteriol*, 192(20):5454–64, 2010. ISSN 1098-5530 (Electronic) 0021-9193 (Linking). doi: 10.1128/JB.00493-10.

5. C. H. Bi, P. Su, J. Muller, Y. C. Yeh, S. R. Chhabra, H. R. Beller, S. W. Singer, and N. J. Hillson. Development of a broad-host synthetic biology toolbox for *Ralstonia eutropha* and its application to engineering hydrocarbon biofuel production. *Microbial Cell Factories*, 12, 2013. ISSN 1475-2859. doi: Artn10710.1186/1475-2859-12-107.

6. J. Marc, E. Grousseau, E. Lombard, A. J. Sinskey, N. Gorret, and S. E. Guillouet. Over expression of groesl in cupriavidus necator for heterotrophic and autotrophic isopropanol production. *Metab Eng*, 42:74–84, 2017. ISSN 1096-7184 (Electronic) 1096-7176 (Linking). doi: 10.1016/j.ymben.2017.05.007.

7. H Li, P H Opgenorth, D G Wernick, S Rogers, T Y Wu, W Higashide, P Malati, Y X Huo, K M Cho, and J C Liao. Integrated electromicrobial conversion of co2 to higher alcohols. *Science*, 335(6076):1596–1596, 2012.

8. R. Cramm. Genomic view of energy metabolism in ralstonia eutropha h16. *J Mol Microbiol Biotechnol*, 16(1-2):38–52, 2009. ISSN 1660-2412 (Electronic) 1464-1801 (Linking). doi: 10.1159/000142893.

9. S. Gruber, J. Hagen, H. Schwab, and P. Koefinger. Versatile and stable vectors for efficient gene expression in ralstonia eutropha h16. *J Biotechnol*, 186:74–82, 2014. ISSN 1873-4863 (Electronic) 0168-1656 (Linking). doi: 10.1016/j.jbiotec.2014.06.030.

10. J. Lu, C. J. Brigham, C. S. Gai, and A. J. Sinskey. Studies on the production of branched-chain alcohols in engineered ralstonia eutropha. *Appl Microbiol Biotechnol*, 96(1):283–97, 2012. ISSN 1432-0614 (Electronic) 0175-7598 (Linking). doi: 10.1007/s00253-012-4320-9.

11. T Oda, K Oda, H Yamamoto, A Matsuyama, M Ishii, Y Igarashi, and H Nishihara. Hydrogen-driven asymmetric reduction of hydroxyacetone to (r)-1,2-propanediol by ralstonia eutropha transformant expressing alcohol dehydrogenase from from kluyveromyces lactis. *Microbial Cell Factories*, 12, 2013.

12. J. Muller, D. MacEachran, H. Burd, N. Sathitsuksanoh, C. Bi, Y. C. Yeh, T. S. Lee, N. J. Hillson, S. R. Chhabra, S. W. Singer, and H. R. Beller. Engineering of ralstonia eutropha h16 for autotrophic and heterotrophic production of methyl ketones. *Appl Environ Microbiol*, 79(14):4433–9, 2013. ISSN 1098-5336 (Electronic) 0099-2240 (Linking). doi: 10.1128/AEM.00973-13.

13. J. S. Chen, B. Colon, B. Dusel, M. Ziesack, J. C. Way, and J. P. Torella. Production of fatty acids in ralstonia eutropha h16 by engineering beta-oxidation and carbon storage. *PeerJ*, 3:e1468, 2015. ISSN 2167-8359 (Print) 2167-8359 (Linking). doi: 10.7717/peerj.1468.

14. D. Przybylski, T. Rohwerder, C. Dilssner, T. Maskow, H. Harms, and R. H. Muller. Exploiting mixtures of h2, co2, and o2 for improved production of methacrylate precursor 2-hydroxyisobutyric acid by engineered cupriavidus necator strains. *Appl Microbiol Biotechnol*, 99(5):2131–45, 2015. ISSN 1432-0614 (Electronic) 0175-7598 (Linking). doi: 10.1007/s00253-014-6266-6.

15. L. Crepin, E. Lombard, and S. E. Guillouet. Metabolic engineering of cupriavidus necator for heterotrophic and autotrophic alka(e)ne production. *Metab Eng*, 37:92–101, 2016. ISSN 1096-7184 (Electronic) 1096-7176 (Linking). doi: 10.1016/j.ymben.2016.05.002.

16. G. C. Barnard, G. E. Henderson, S. Srinivasan, and T. U. Gerngross. High level recombinant protein expression in *Ralstonia eutropha* using t7 rna polymerase based amplification. *Protein Expr Purif*, 38(2):264–71, 2004. ISSN 1046-5928 (Print) 1046-5928 (Linking). doi: 10.1016/j.pep.2004.09.001.

17. A. Pohlmann, W. F. Fricke, F. Reinecke, B. Kusian, H. Liesegang, R. Cramm, T. Eitinger, C. Ewering, M. Potter, E. Schwartz, A. Strittmatter, I. Voss, G. Gottschalk, A. Steinbuchel, B. Friedrich, and B. Bowien. Genome sequence of the bioplastic-producing "knallgas" bacterium ralstonia eutropha h16. *Nat Biotechnol*, 24(10):1257–62, 2006. ISSN 1087-0156 (Print) 1087-0156 (Linking). doi: 10.1038/nbt1244.

18. A. Lykidis, D. Perez-Pantoja, T. Ledger, K. Mavromatis, I. J. Anderson, N. N. Ivanova, S. D. Hooper, A. Lapidus, S. Lucas, B. Gonzalez, and N. C. Kyrpides. The complete multipartite genome sequence of cupriavidus necator jmp134, a versatile pollutant degrader. *PLoS*

One, 5(3):e9729, 2010. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0009729.

19. N. Berezina, B. Yada, and R. Lefebvre. From organic pollutants to bioplastics: insights into the bioremediation of aromatic compounds by *Cupriavidus necator*. *N Biotechnol*, 32(1):47–53, 2015. ISSN 1876-4347 (Electronic) 1871-6784 (Linking). doi: 10.1016/j.nbt.2014.09.003.

20. O. P. Peoples and A. J. Sinskey. Poly-beta-hydroxybutyrate (phb) biosynthesis in alcaligenes eutrophus h16. identification and characterization of the phb polymerase gene (phbc). *J Biol Chem*, 264(26):15298–303, 1989. ISSN 0021-9258 (Print) 0021-9258 (Linking).

21. S. Obruca, I. Marova, M. Stankova, L. Mravcova, and Z. Svoboda. Effect of ethanol and hydrogen peroxide on poly(3-hydroxybutyrate) biosynthetic pathway in cupriavidus necator h16. *World J Microbiol Biotechnol*, 26(7):1261–7, 2010. ISSN 0959-3993 (Print) 0959-3993 (Linking). doi: 10.1007/s11274-009-0296-8.

22. P. K. Sharma, J. Fu, V. Spicer, O. V. Krokhin, N. Cicek, R. Sparling, and D. B. Levin. Global changes in the proteome of cupriavidus necator h16 during poly-(3-hydroxybutyrate) synthesis from various biodiesel by-product substrates. *AMB Express*, 6(1):36, 2016. ISSN 2191-0855 (Print) 2191-0855 (Linking). doi: 10.1186/s13568-016-0206-z.

23. Kedar G. Patel, Mark Welch, and Claes Gustafsson. *Leveraging gene synthesis, advanced cloning techniques, and machine learning for metabolic pathway engineering*, book section Chapter 4, pages 53–71. Springer International Publishing, 2016. ISBN 978-3-319-41964-0 978-3-319-41966-4. doi: 10.1007/978-3-319-41966-4_4.

24. P. Xu, E. A. Rizzoni, S. Y. Sul, and G. Stephanopoulos. Improving metabolic pathway efficiency by statistical model-based multivariate regulatory metabolic engineering. *ACS Synth Biol*, 6(1):148–158, 2017. ISSN 2161-5063 (Electronic) 2161-5063 (Linking). doi: 10.1021/acssynbio.6b00187.

25. F. Caschera, M. A. Bedau, A. Buchanan, D. Gazzola G. Hanczyc M. Cawse, J. de Lucrezia, and N. H. Packard. Coping with complexity: machine learning optimization of cell-free protein synthesis. *Biotechnol. Bioeng.*, 108(9):2218–2228, 2011.

26. C. Gustafsson, J. Minshull, S. Govindarajan, J. Ness, A. Villalobos, and M. Welch. Engineering genes for predictable protein expression. *Protein Expression and Purification*, 83 (1):37–46, 2012. ISSN 1046-5928.

27. S. R. Brown, M. Staff, R. Lee, J. Love, D. A. Parker, S. J. Aves, and T. P. Howard. Design of experiments methodology to build a multifactorial statistical model describing the metabolic interactions of alcohol dehydrogenase isozymes in the ethanol biosynthetic pathway of the yeast saccharomyces cerevisiae. *ACS Synth Biol*, 2018. ISSN 2161-5063 (Electronic) 2161-5063 (Linking). doi: 10.1021/acssynbio.8b00112.

28. S. Govindarajan, B. Mannervik, J. A. Silverman, K. Wright, D. Regitsky, U. Hegazy, T. J. Purcell, M. Welch, J. Minshull, and C. Gustafsson. Mapping of amino acid substitutions conferring herbicide resistance in wheat glutathione transferase. *ACS Synth Biol*, 4(3):221–7, 2015. ISSN 2161-5063 (Electronic) 2161-5063 (Linking). doi: 10.1021/sb500242x.

29. S. Sichwart, S. Hetzler, D. Broker, and A. Steinbuchel. Extension of the substrate utilization range of ralstonia eutropha strain h16 by metabolic engineering to include mannose and glucose. *Appl Environ Microbiol*, 77(4):1325–34, 2011. ISSN 1098-5336 (Electronic) 0099-2240 (Linking). doi: 10.1128/AEM.01977-10.

30. A. Franz, R. Rehner, A. Kienle, and H. Grammel. Rapid selection of glucose-utilizing variants of the polyhydroxyalkanoate producer ralstonia eutropha h16 by incubation with high substrate levels. *Lett Appl Microbiol*, 54(1):45–51, 2012. ISSN 1472-765X (Electronic) 0266-8254 (Linking). doi: 10.1111/j.1472-765X.2011.03171.x.

31. M. Raberg, J. Bechmann, U. Brandt, J. Schluter, B. Uischner, B. Voigt, M. Hecker, and A. Steinbuchel. Versatile metabolic adaptations of ralstonia eutropha h16 to a loss of pdhl, the e3 component of the pyruvate dehydrogenase complex. *Appl Environ Microbiol*, 77 (7):2254–63, 2011. ISSN 1098-5336 (Electronic) 0099-2240 (Linking). doi: 10.1128/AEM.02360-10.

32. E. Schwartz, B. Voigt, D. Zuhlke, A. Pohlmann, O. Lenz, D. Albrecht, A. Schwarze, Y. Kohlmann, C. Krause, M. Hecker, and B. Friedrich. A proteomic view of the facultatively chemolithoautotrophic lifestyle of ralstonia eutropha h16. *Proteomics*, 9(22):5132–42, 2009. ISSN 1615-9861 (Electronic) 1615-9853 (Linking). doi: 10.1002/pmic.200900333.

33. J. Lu, C. J. Brigham, C. Rha, and A. J. Sinskey. Characterization of an extracellular lipase and its chaperone from ralstonia eutropha h16. *Appl Microbiol Biotechnol*, 97(6):2443–54, 2013. ISSN 1432-0614 (Electronic) 0175-7598 (Linking). doi: 10.1007/s00253-012-4115-z.

34. P. Chandrangsu, C. Rensing, and J. D. Helmann. Metal homeostasis and resistance in bacteria. *Nat Rev Microbiol*, 15(6):338–350, 2017. ISSN 1740-1534 (Electronic) 1740-1526 (Linking). doi: 10.1038/nrmicro.2017.15.

35. J. M. Arguello, D. Raimunda, and T. Padilla-Benavides. Mechanisms of copper homeostasis in bacteria. *Front Cell Infect Microbiol*, 3:73, 2013. ISSN 2235-2988 (Print) 2235-2988 (Linking). doi: 10.3389/fcimb.2013.00073.

36. C. Subramanyam and G. Venkateswerlu. The effect of copper on histidine biosynthesis in neurospora crassa. *Journal of Biosciences*, 1(2):143–149, 1979. ISSN 0250-5991 0973-7138. doi: 10.1007/bf02706326.

37. D. A. Pearce and F. Sherman. Toxicity of copper, cobalt, and nickel salts is dependent on histidine metabolism in the yeast saccharomyces cerevisiae. *J Bacteriol*, 181(16):4774–9, 1999. ISSN 0021-9193 (Print) 0021-9193 (Linking).

Supporting Information

Table SI 1. Previously described chemically defined media for the growth of *C. necator*. All concentrations are g/l unless otherwise indicated. References: 1. Grousseau et al. Metab Eng 42:74–84. 2. Müller et al. 2013. Appl Environ Microbiol. 3. Obruca et al. 1261–1267. 4. Sharma et al. AMB Express 6:36. 5. Aneja et al. Biotechnol Lett 31:1601–1612. 6. Lu et al. Appl Microbiol Biotechnol 96:283–297. 7. Lu et al. Appl Microbiol Biotechnol 97:2443–2454. 8. Oda et al 2013 Microb Cell Fact 12:2.

Carbon source	Phosphate buffer	Potassium	Magnesium	Calcium	Nitrogen	Trace elements	Reference
50 fructose	1.5 Na2HPO4.12H2O 0.25 KH2PO4		0.75 MgSO4.7H2O	0.015 CaCl2	2.39 (NH4)2SO4 0.285 nitrlotriacetic acid 28% 0.9 Fe(NH4)2(citrate)	0.00045 H3BO3 0.0003 CoCl2 0.00015 ZnSO4.7H2O 0.000045 MnCl2.4H2O 0.000045 NaMoO4.2H2O 0.00003 NiCl2.6H2O 0.000015CuSO4.5H2O	1
0.5-1% fructose	12 ml 1.1 M H3PO4	0.45			0.1 or 0.01% NH4Cl	24 ml/trace: 2.6 CaCl2.2H2O 0.1 MnSO4.H2O 0.1 ZnSO4.7H2O 0.02 CuSO4.5H2O 9.7 FeCl3	2
10-20 Fructose	1 KH2PO4 11.1 Na2HPO4.12H2O		0.2 MgSO4		3 (NH4)2SO4	7.8 CaCl2 0.156 CuSO4.5H2O 0.119 CoCl2 0.118 NiCl2 0.062 CrCl2 0.3 H3BO3	3
2% v/v biodiesel-driven glycerol, 2 % w/v glycerol bottom 1% v/v free fatty acids	6.7 Na2HPO4.2H2O 1.5 KH2PO4		0.2 MgSO4.7H2O	0.01 CaCl2.2H2O	1 (NH4)2SO4, 0.06 Fe(NH4)2(citrate)	0.2 CoCl2 0.1 ZnSO4.7H2O 0.03 MnCl2.4H2O 0.02 NaMoO4.2H2O 0.02 NiCl2.6H2O 0.01CuSO4.5H2O 1.97 FeCl3	4
20 fructose	1.51 Na2HPO4 2.65 KH2PO4		0.3 MgSO4.7H2O			9.7 CaCl2.2H2O 0.156 CuSO4.5H2O 0.184 CoCl2.6H2O 0.118 NiCl2.6H2O 0.062 CrCl3	5
2% fructose, 2% sodium gluconate, 1% palm oil or 0.5% Tween-60	4 NaH2PO4 4.6 Na2HPO4	0.45 K2SO4	0.39 MgSO4	0.062 CaCl2	0.5 NH4Cl	15 FeSO4.7H2O 2.4 MnSO4.H2O 2.4 ZnSO4.7H2O 0.48 CuSO4.5H2O	6, 7
2-4 Fructose, 2 ml glycerol	1.5 KH2PO4 9 Na2HPO4		0.2 MgSO4.7H2O	0.01 CaCl2.2H2O	2 NH4Cl	0.05 FeCl3.6H2O 0.019-0.190 NiCl2.6H2O	8

Table SI 2. Scoping experiment. A scoping trial was developed to assess the impact of 10 basic ingredients found within the chemically defined media. All concentrations are in g/L except trace elements, amino acids and vitamins which are ml / L. Abbreviations: T.E. Trace element mixture; A.A., amino acid mixture; Vit., vitamin mixture. Carbon: fructose, glucose, glycerol or sucrose. The medium trial was performed in duplicate.

Media	Carbon	NaH2PO4	Na2HPO4	K2SO4	Mg2SO4	CaCl2	NH4Cl	T.E.	A.A.	Vit.
Low	0.5	0.1	0.1	0.01	0.01	0.01	0.01	0.01	0.05	0.01
Medium	20.25	3.05	3.05	2.405	1.405	0.405	2.405	2.405	9.525	2.405
High	40	6	6	4.8	2.8	0.8	4.8	4.8	19	4.8

Table SI 3. Definitive Screening Design 2. A second definitive screening design was developed to assess the impact of the five ingredients identified as important to the chemically defined media. All concentrations are in g/L. Abbreviations with exception of T.E. and A.A, which are in ml/L: Fru., fructose, T.E., Trace element mixture; A.A., amino acid mixture. The DSD array was performed in duplicate.

Media	Fru.	NaH2PO4	Na2HPO4	K2SO4	Mg2SO4	CaCl2	NH4Cl	T.E.	A.A.	OD1	OD2
1	15	3.05	0.100	2.405	1.405	0.10	0.01	0.010	5	1.71	1.36
2	15	3.05	3.050	2.405	1.405	0.46	0.01	2.405	20	1.84	1.67
3	25	3.05	1.575	2.405	1.405	0.46	0.01	0.010	5	1.72	1.65
4	5	3.05	1.575	2.405	1.405	0.10	0.01	2.405	20	1.95	1.95
5	5	3.05	0.100	2.405	1.405	0.28	0.01	2.405	5	1.85	0.99
6	25	3.05	3.050	2.405	1.405	0.28	0.01	0.010	20	0.57	1.84
7	5	3.05	3.050	2.405	1.405	0.46	0.01	1.280	5	1.18	1.18
8	25	3.05	0.100	2.405	1.405	0.10	0.01	1.280	20	2.10	2.06
9	5	3.05	3.050	2.405	1.405	0.10	0.01	0.010	13	2.03	1.99
10	25	3.05	0.100	2.405	1.405	0.46	0.01	2.405	13	1.86	1.40
11	15	3.05	1.575	2.405	1.405	0.28	0.01	1.208	13	1.92	1.96
12	5	3.05	0.100	2.405	1.405	0.46	0.01	0.010	20	2.01	2.08
13	25	3.05	3.050	2.405	1.405	0.10	0.01	2.405	5	1.06	0.91

Table SI 4. Definitive Screening Design 3. A final set of experiments were performed to re-assess the impact of components not deemed significant in the first DSD. All concentrations are in g/L. Abbreviations: Fru., fructose, T.E. Trace element mixture; A.A., amino acid mixture. The DSD array was performed in duplicate.

| Fru.	NaH$_2$PO$_2$	Na$_2$HPO$_2$	K$_2$SO$_2$	MgSO$_2$	CaCl$_2$	NH$_4$Cl	T.E.	A.A.	OD1	OD2
5	0.1	0.1	0.01	0.01	0.459	4.8	1	10	0.13	0.17
5	0.1	0.1	1.2075	1.405	0.459	0.01	1	10	1.79	1.74
5	0.1	0.1	2.405	0.01	0.459	0.01	1	10	1.64	1.71
5	0.1	0.1	0.01	1.405	0.459	2.405	1	10	0.95	0.57
5	3.05	0.1	1.2075	0.01	0.459	4.8	1	10	1.85	1.76
5	3.05	0.1	0.01	0.7075	0.459	0.01	1	10	1.86	1.65
5	3.05	0.1	2.405	1.405	0.459	0.01	1	10	1.89	1.88
5	1.575	0.1	0.01	0.01	0.459	0.01	1	10	1.76	1.92
5	3.05	0.1	0.01	1.405	0.459	4.8	1	10	0.84	0.79
5	0.1	0.1	2.405	0.7075	0.459	4.8	1	10	0.66	0.73
5	1.575	0.1	2.405	1.405	0.459	4.8	1	10	1.51	1.55
5	3.05	0.1	2.405	0.01	0.459	2.405	1	10	1.95	1.89
5	1.575	0.1	1.2075	0.7075	0.459	2.405	1	10	1.49	1.76
20	0.1	0.1	0.01	0.01	0.459	4.8	1	10	0.15	0.81
20	0.1	0.1	1.2075	1.405	0.459	0.01	1	10	1.86	2.09
20	0.1	0.1	2.405	0.01	0.459	0.01	1	10	2.09	1.92
20	0.1	0.1	0.01	1.405	0.459	2.405	1	10	1.66	0.95
20	3.05	0.1	1.2075	0.01	0.459	4.8	1	10	1.79	1.52
20	3.05	0.1	0.01	0.7075	0.459	0.01	1	10	1.71	1.90
20	3.05	0.1	2.405	1.405	0.459	0.01	1	10	2.05	1.73
20	1.575	0.1	0.01	0.01	0.459	0.01	1	10	1.81	1.85
20	3.05	0.1	0.01	1.405	0.459	4.8	1	10	1.08	0.98
20	0.1	0.1	2.405	0.7075	0.459	4.8	1	10	1.17	0.96
20	1.575	0.1	2.405	1.405	0.459	4.8	1	10	1.65	1.48
20	3.05	0.1	2.405	0.01	0.459	2.405	1	10	1.74	1.80
20	1.575	0.1	1.2075	0.7075	0.459	2.405	1	10	1.65	1.39
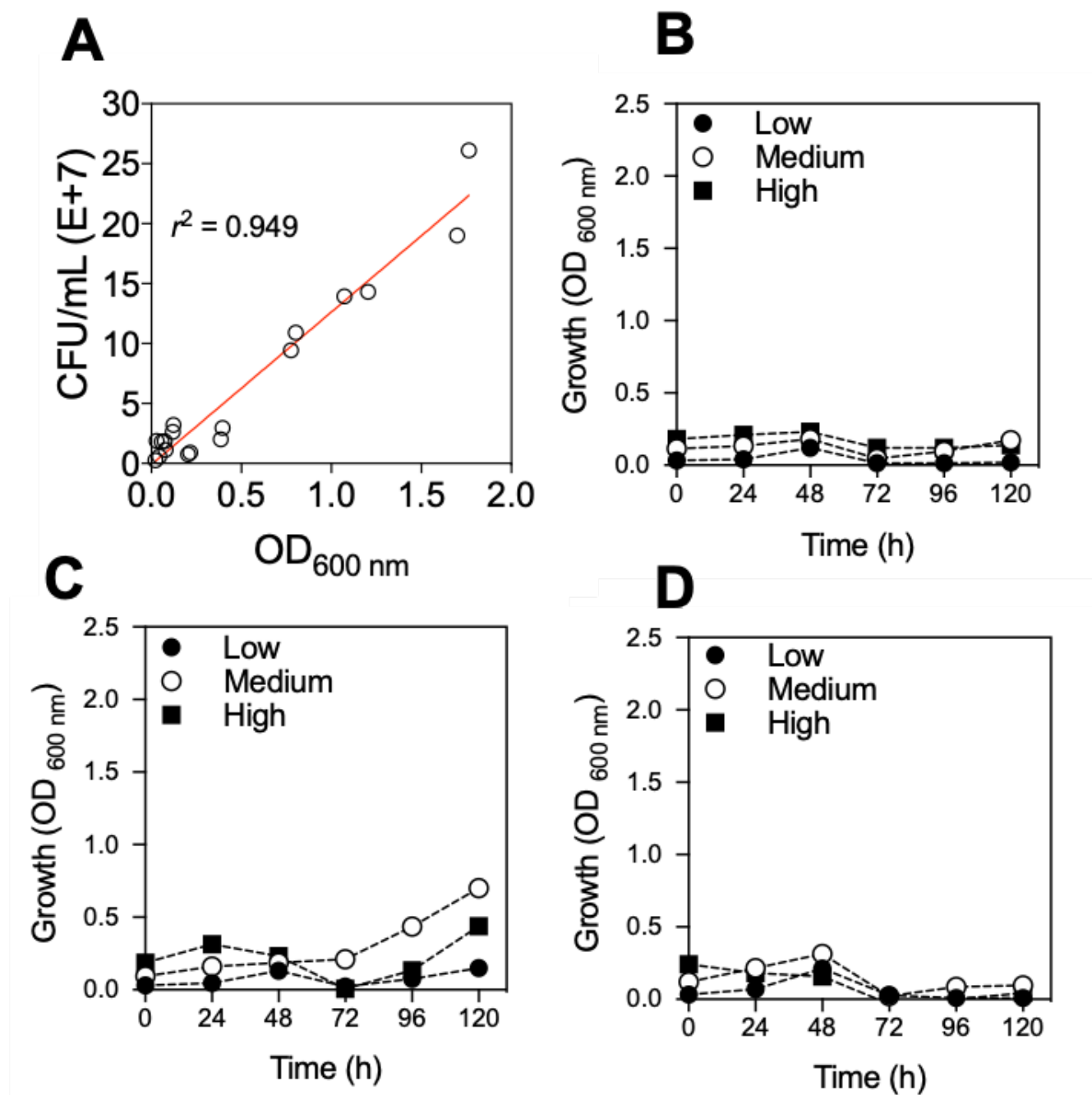
Fig. SI 1. Growth characteristics of *C. necator* from initial scoping experiments. A. Colony forming units against OD_{600nm}. Growth of *C. necator* on **B.** Glucose, **C.** Glycerol, and **D.** Sucrose as sole carbon source.
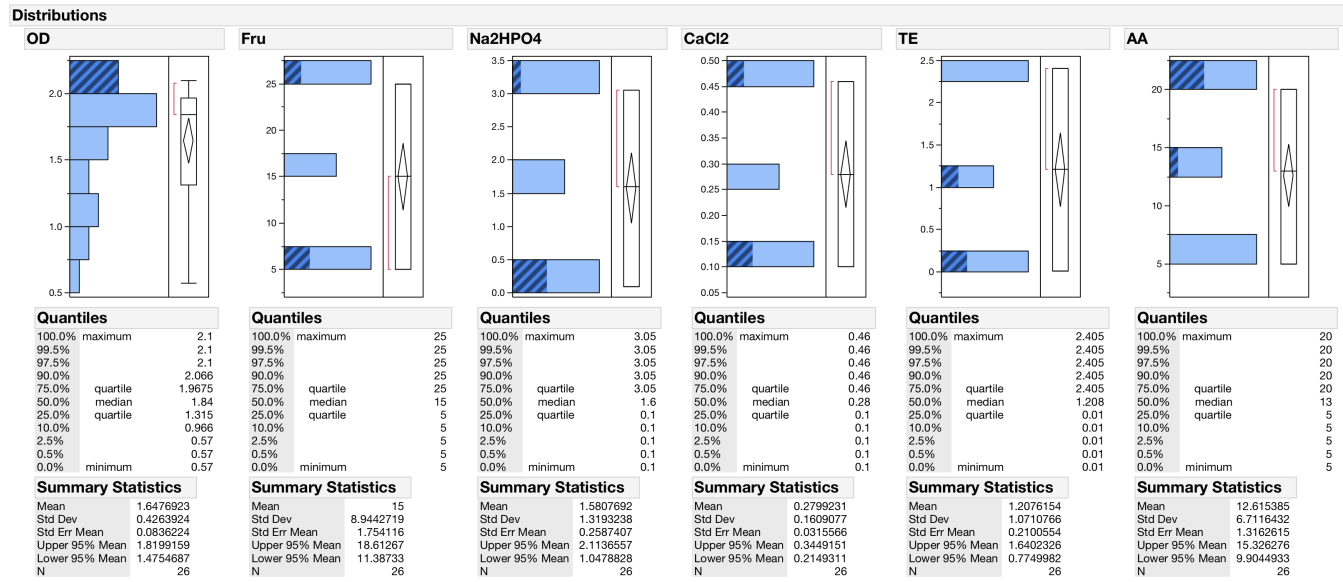
Fig. SI 2. Distributions of data for Definitive Screening Design 2. Highlighted in blue cross-hatch are the settings that resulted in the greatest OD_{600nm} at 72 h for DSD2. Fructose, Na_2HPO_4 and $CaCl_2$ may be at either the highest or lowest concentrations, whereas trace elements and amino acid concentrations are found at the lowest or highest concentrations respectively.
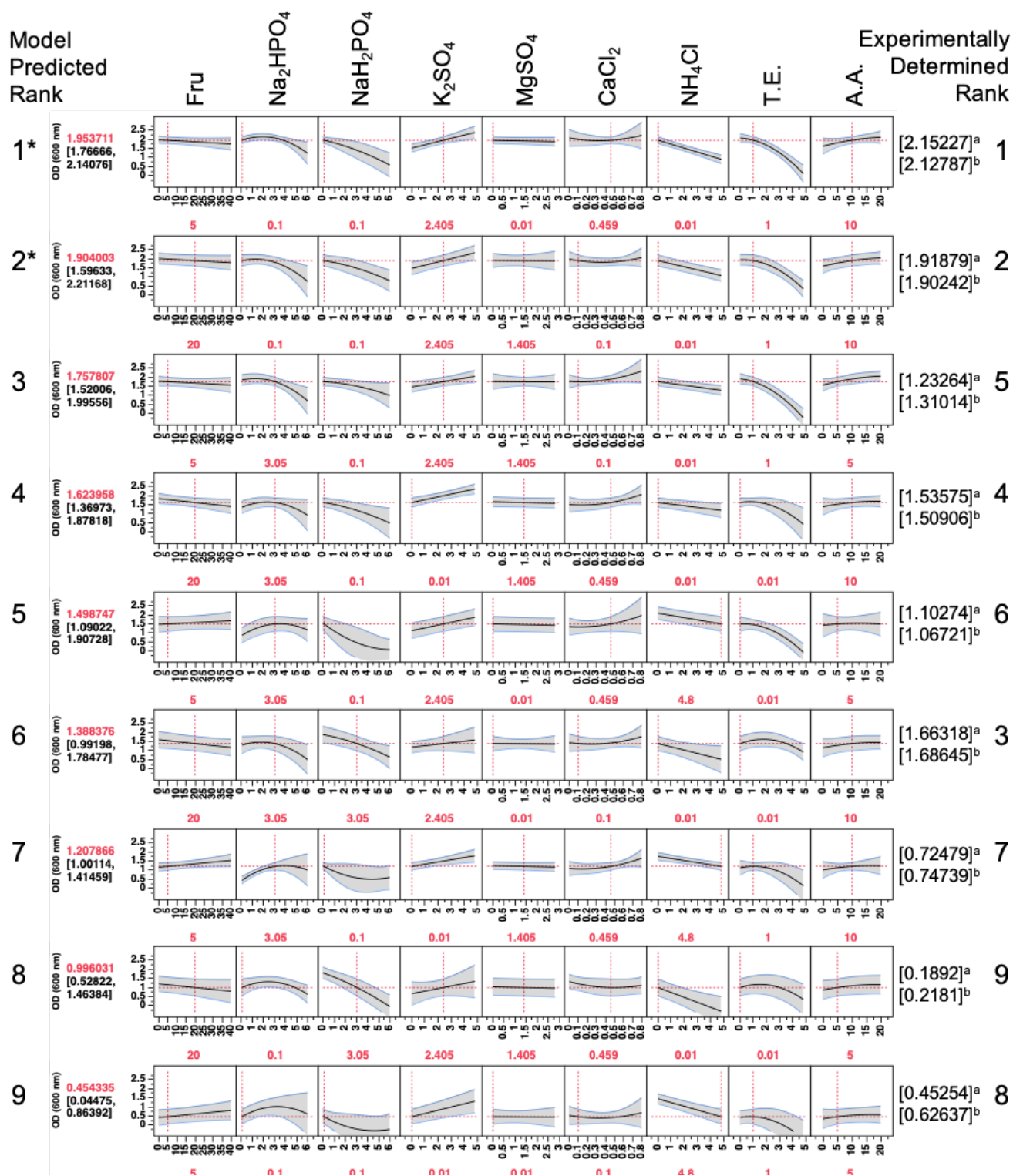
Fig. SI 3. 100mL validation. Nine different formulations were assessed at 100 mL culture volumes. For each row, the model predicted rank is shown on the left hand side, media settings for each component are shown in red underneath, the measured $OD_{(600nm)}$ 72 h for two replicate experiments and experimentally determined rank on the right hand side.
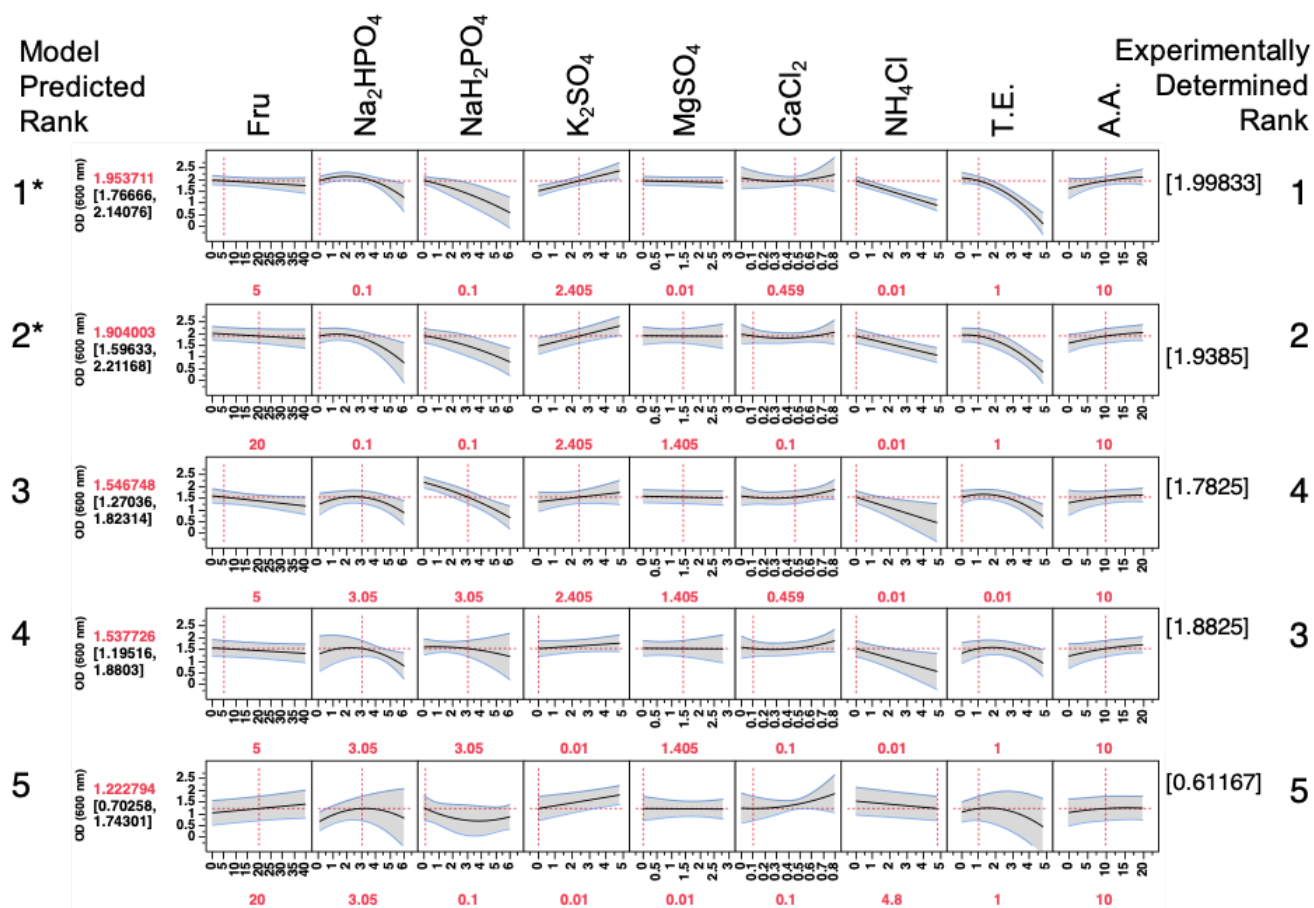
Fig. SI 4. 1 L validation. Five different formulations were assessed at 1 L culture volumes. For each row, the model predicted rank is shown on the left hand side, media settings for each component are shown in red underneath, the measured $OD_{(600nm)}$ 72 h for two replicate experiments and experimentally determined rank on the right hand side.