

















Computation-aided designs enable developing auxotrophic metabolic sensors for wide-range glyoxylate and glycolate detection

Received: 1 September 2024

Accepted: 18 February 2025

Published online: 04 March 2025

 Check for updates

Enrico Orsi ^{1,7} , Helena Schulz-Mirbach ^{2,7}, Charles A. R. Cotton³, Ari Satanowski ², Henrik M. Petri ², Susanne L. Arnold², Natalia Grabarczyk¹, Rutger Verbakel¹, Karsten S. Jensen¹, Stefano Donati ¹, Nicole Paczia ², Timo Glatter ², Andreas M. Küffner², Tanguy Chotel², Farah Schillmüller², Alberto De Maria ¹, Hai He ², Steffen N. Lindner ^{3,4}, Elad Noor ⁵, Arren Bar-Even ^{3,6}, Tobias J. Erb ² & Pablo I. Nikel ¹ 

Auxotrophic metabolic sensors (AMS) are microbial strains modified so that biomass formation correlates with the availability of specific metabolites. These sensors are essential for bioengineering (e.g., in growth-coupled designs) but creating them is often a time-consuming and low-throughput process that can be streamlined by *in silico* analysis. Here, we present a systematic workflow for designing, implementing, and testing versatile AMS based on *Escherichia coli*. Glyoxylate, a key metabolite in (synthetic) CO₂ fixation and carbon-conserving pathways, served as the test analyte. Through iterative screening of a compact metabolic model, we identify non-trivial growth-coupled designs that result in six AMS with a wide sensitivity range for glyoxylate, spanning three orders of magnitude in the detected analyte concentration. We further adapt these *E. coli* AMS for sensing glycolate and demonstrate their utility in both pathway engineering (testing a key metabolic module for carbon assimilation *via* glyoxylate) and environmental monitoring (quantifying glycolate produced by photosynthetic microalgae). Adapting this workflow to the sensing of different metabolites could facilitate the design and implementation of AMS for diverse biotechnological applications.

Metabolism is the set of chemical reactions sustaining life¹. These reactions can be systematically investigated thanks to standardized approaches in synthetic biology, a field that applies engineering principles to living (micro)organisms². Engineered microbial auxotrophic

strains with defined disruptions of the metabolic network can be designed to have a defined demand for a specific metabolite, which can be supplemented exogenously, and thus are ideal platforms for metabolic studies^{3,4}. Given their dependence on metabolic

¹The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark. ²Max Planck Institute for Terrestrial Microbiology, Marburg, Germany. ³Max Planck Institute of Molecular Plant Physiology, Potsdam, Germany. ⁴Department of Biochemistry, Charité Universitätsmedizin Berlin, Freie Universität Berlin and Humboldt-Universität, Berlin, Germany. ⁵Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel. ⁶Deceased: Arren Bar-Even. ⁷These authors contributed equally: Enrico Orsi, Helena Schulz-Mirbach.

 e-mail: enricoo@biosustain.dtu.dk; pabnik@biosustain.dtu.dk

intermediates, we will refer to these engineered strains as auxotrophic metabolic sensors (AMS). Originally used for biochemistry studies through growth complementation, metabolic sensors are now increasingly used for studying metabolic pathways in vivo^{5–10}.

Construction of AMS requires reliable quantitative predictions of the effect of gene deletions that must be confirmed in vivo. Yet, the implementation of auxotrophic phenotypes still relies on one-at-a-time interventions (e.g., multiple rounds of deletion and testing) with the exception of a handful of cases where systematic analysis of all the possible options was performed through screening of metabolic models^{8,11,12}. In all these cases, computations were performed on a core metabolic model, limited to glycolysis/gluconeogenesis, pentose phosphate pathway, the tricarboxylic acid (TCA) cycle, and oxidative phosphorylation. The output is inherently limited to growth-coupled designs around the 12 universal biomass precursors belonging to the central carbon metabolism¹³. Several additional reactions are involved in the synthesis of key biomass precursors in vivo (e.g., amino acids or lipids biosynthesis), hence larger metabolic models could lead to multiple growth-coupled designs that can be further explored. However, a model encompassing the entirety of an organism's metabolism is highly complex, which could increase computing times and lead to irrelevant or unfeasible solutions based on secondary reactions. A compromise between these two approaches is employing a medium-scale metabolic model, covering core metabolism and essential metabolic pathways (i.e., the synthesis of energy carriers and biosynthetic precursors, such as amino acids), without the redundancies of a genome-scale metabolic reconstruction¹⁴.

In this work, we designed AMS for glyoxylate, a non-essential metabolite that is not directly involved in the synthesis of biomass precursors in *Escherichia coli*. Creating metabolic sensors for this molecule therefore requires deep metabolic rewiring. We adopted a medium-scale metabolic model to guide the design and engineering of several AMS for this metabolite. The relevance of glyoxylate is evident by a wealth of studies in the context of prebiotic chemistry^{15–18}, synthetic metabolism^{19–22}, and biomanufacturing^{23,24}. We show that complex screening applications can be accessed using our sensor strains. These include in vivo screening of enzyme variants for synthetic one-carbon (C1) assimilation and sensing of excreted glycolate produced during photorespiration as a globally occurring environmental process. This work improves the robustness of the workflow for the computation-guided design of metabolic interventions and yields an array of novel glyoxylate and glycolate sensors supporting applications beyond the traditional uses of metabolic sensors.

Results

Systematic interrogation of a compact metabolic reconstruction for growth-coupled designs around glyoxylate

Since glyoxylate is not an essential metabolite for wild-type *E. coli*, coupling this organism's growth to glyoxylate availability requires complex metabolic rewiring. Therefore, we propose a pipeline that predicts potential designs in silico using a compact metabolic model (Fig. 1A). To systematically investigate all the combinations of knockouts (KOs) that force growth to be dependent on glyoxylate, we used an algorithm we had previously applied to a core *E. coli* model^{12,25}, but replaced the core-model with the recently published iCH360 medium-scale model of *E. coli*'s metabolism¹⁴. This medium-scale model (323 reactions) offers wider coverage of possible KOs compared to a core model (95 reactions) and limits the number of additional experience-based manual interventions required compared to a full genome-scale model (2000 reactions or above). To adapt the medium-scale model to identify growth-coupled designs for glyoxylate, we included four additional reactions: (i) glyoxylate uptake; (ii) an aspartate-glyoxylate aminotransferase (BHC)²⁶ as proxy for promiscuous transaminase activities on glyoxylate; (iii) glyoxylate carboligase (GLXCL); and (iv) tartronate semialdehyde reductase (TRSARr) (Fig. 1B). The procedure

then iterates through a list of KO combinations and uses flux balance analysis to test which combinations render growth dependent on glyoxylate using different carbon substrates^{12,25}.

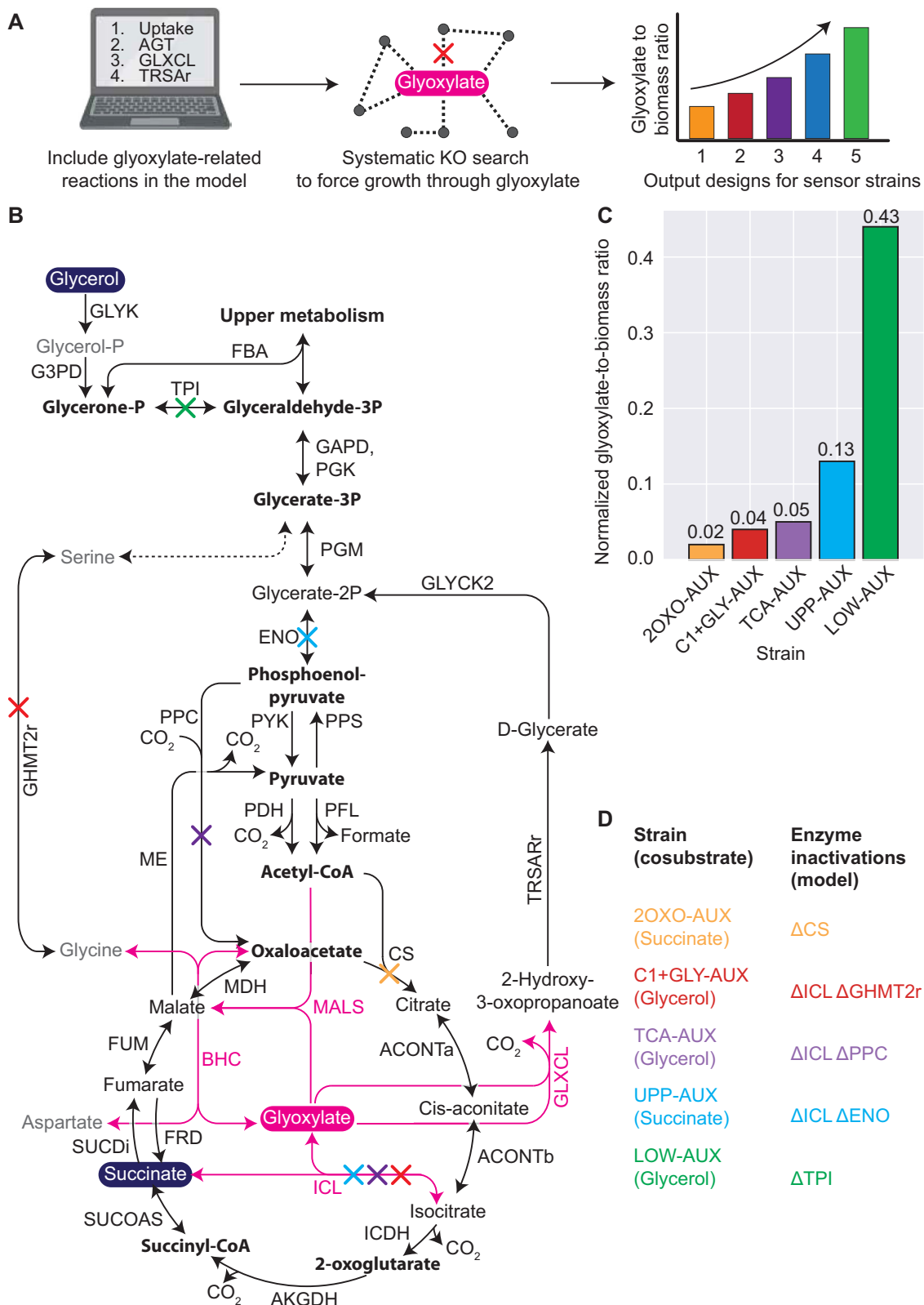
With this extended medium-scale model, we iteratively explored how different combinations of gene deletions affect the dependency of *E. coli* growth on glyoxylate. We set a maximum of two KOs to force glyoxylate consumption for biomass buildup. To broaden the spectrum of identified designs, we performed this search for growth on two different main carbon sources (glycerol and succinate) not affected by catabolite repression (Fig. 1B). Although it would be computationally possible to expand the search to more than two KOs and two alternative carbon sources (see “Methods”), we found that the solutions provided by this smaller search space were sufficient. This feature highlights the advantage of using smaller models (here, iCH360) compared to genome-scale metabolic reconstructions, where typically many more KOs are necessary to achieve growth coupling. Very often, the essential KOs are the ones found by the small model, while the others tackle obscure bypasses that are unlikely to be relevant in vivo. In our case, the output of the algorithm was a set of different KO combinations for each of the two carbon sources selected, spanning different levels of dependency on glyoxylate (Supplementary Fig. 1).

From these results, we decided to pursue five designs covering different glyoxylate-derived biomass precursors and glyoxylate demands, namely the amount required for the synthesis of biomass (Fig. 1C and Supplementary Methodological Note). The first group included disruptions of glycolysis/gluconeogenesis, causing reliance on glyoxylate either for the synthesis of intermediates in upper glycolysis (e.g., sugar phosphates, “upper metabolism”) or for the synthesis of “lower metabolism” (UPP-AUX and LOW-AUX, respectively). The second group included deletions requiring glyoxylate for anaplerosis of the TCA cycle (TCA-AUX and 2OXO-AUX). The third group included a design that requires glyoxylate for glycine biosynthesis through transamination (called C1 + GLY-AUX, since glycine is also involved in the synthesis of C1 mediators). Notably, this last group of designs would not have been accessible by using a core model of *E. coli* metabolism because key metabolic routes of amino acid metabolism were not included, e.g., serine and glycine interconversion.

The algorithm provided the minimum set of KOs required for generating each auxotrophic phenotype in the metabolic sensors. However, since this in silico prediction does not consider additional factors, such as metabolic regulation and enzyme promiscuities, it needs to be complemented by literature knowledge of the host's metabolic network and physiology. Thus, we manually expanded the knockout selection with additional targets known from the literature to support the strain engineering step (Fig. 2A). In the next section, we describe how we moved from the *in-silico* predictions to successfully engineering the respective AMS.

Metabolic sensor engineering and corresponding glyoxylate contribution to biomass buildup

We started by engineering a group of AMS that rely on disruptions of glycolysis/gluconeogenesis. The selection predicted to have the highest glyoxylate demand is based on the inactivation of triose phosphate isomerase (encoded by the gene *tpiA*; Fig. 2B). Yet this deletion alone is known to be bypassed during growth on glycerol by activation of the methylglyoxal pathway^{27,28}. To avoid this bypass, we additionally deleted *mgsA* (encoding methylglyoxal synthase), thus generating a double mutant $\Delta tpiA \Delta mgsA$ ²⁸. We refer to this strain as LOW-AUX because glyoxylate supplements all metabolism downstream of glyceraldehyde 3-phosphate (“lower metabolism”). After validating the sensor's dependency on glyoxylate, we confirmed the expected metabolic fluxes by tracing ¹³C incorporation into proteino-genic amino acids during growth of the LOW-AUX strain on uniformly labeled ¹³C₂-glyoxylate, with unlabeled glycerol serving as the main



carbon source. The resulting data confirmed labeling incorporation in the majority of the proteinogenic amino acids, thereby demonstrating the contribution of glyoxylate to the buildup of a large fraction of biomass precursors (Fig. 2B and Supplementary Figs. 2 and 3A). The other glycolytic disruption requires the combined inactivation of enolase (ENO, encoded by *eno*) and isocitrate lyase (ICL, encoded by *aceA*). This strain relies on succinate supplementation to feed “lower

metabolism”, while supplied glyoxylate should be converted to glycerate 2-phosphate, feeding “upper metabolism” (therefore named UPP-AUX). To prevent futile consumption of glyoxylate via the glyoxylate shunt, we decided to delete the relevant reactions. These additional deletions included *aceB* and *glsB* (malate synthases), as well as *ghrA* and *ghrB* (glyoxylate/hydroxypropyruvate reductases) (Fig. 2C). In this strain, glyoxylate followed a different fate, which was traced using

Fig. 1 | In silico modeling for exploring glyoxylate-dependent growth-coupling schemes. **A** Our workflow started by adapting a medium-scale metabolic model of *E. coli* by including reactions involved in glyoxylate production or consumption: glyoxylate uptake, BHC (aspartate-glyoxylate aminotransferase), GLXCL (glyoxylate carboligase), TRSAr (tartronate semialdehyde reductase). Next, the updated model was used to run the algorithm for systematically determining combinations of gene deletions that require glyoxylate for growth. As output, the algorithm proposes different combinations of such gene inactivations and ranks them quantitatively in terms of glyoxylate demand (represented by a “glyoxylate-to-biomass ratio”, GBR). Created in BioRender. Orsi (2025) <https://BioRender.com/x91g989>. **B** Schematic overview of the central carbon metabolism of *E. coli* covered by the adapted model. Glyoxylate is shown in magenta together with the reactions

(and associated enzymes) originating from this molecule. Succinate and glycerol are shown (navy blue) because these metabolites are chosen as added carbon sources for the design of individual glyoxylate sensor strains. Essential biomass precursors are indicated in bold. Enzymes are abbreviated following BiGG nomenclature. Colored arrows indicate reactions deleted in certain sensor strain designs as proposed by the algorithm to force glyoxylate dependence. Each color represents a different sensor strain design. **C** Graphic output highlighting the different combinations of enzyme inactivations proposed by the algorithm. The normalized GBR value is a unitless value obtained by dividing the GBR of the sensor by the GBR of the wild-type strain growing on glyoxylate as a unique carbon source. **D** Legend of strains and associated enzyme inactivations selected from the algorithm.

^{13}C labeling, confirming its contribution to the synthesis of amino acids derived from upper metabolism (Fig. 2C and Supplementary Figs. 2 and 3B).

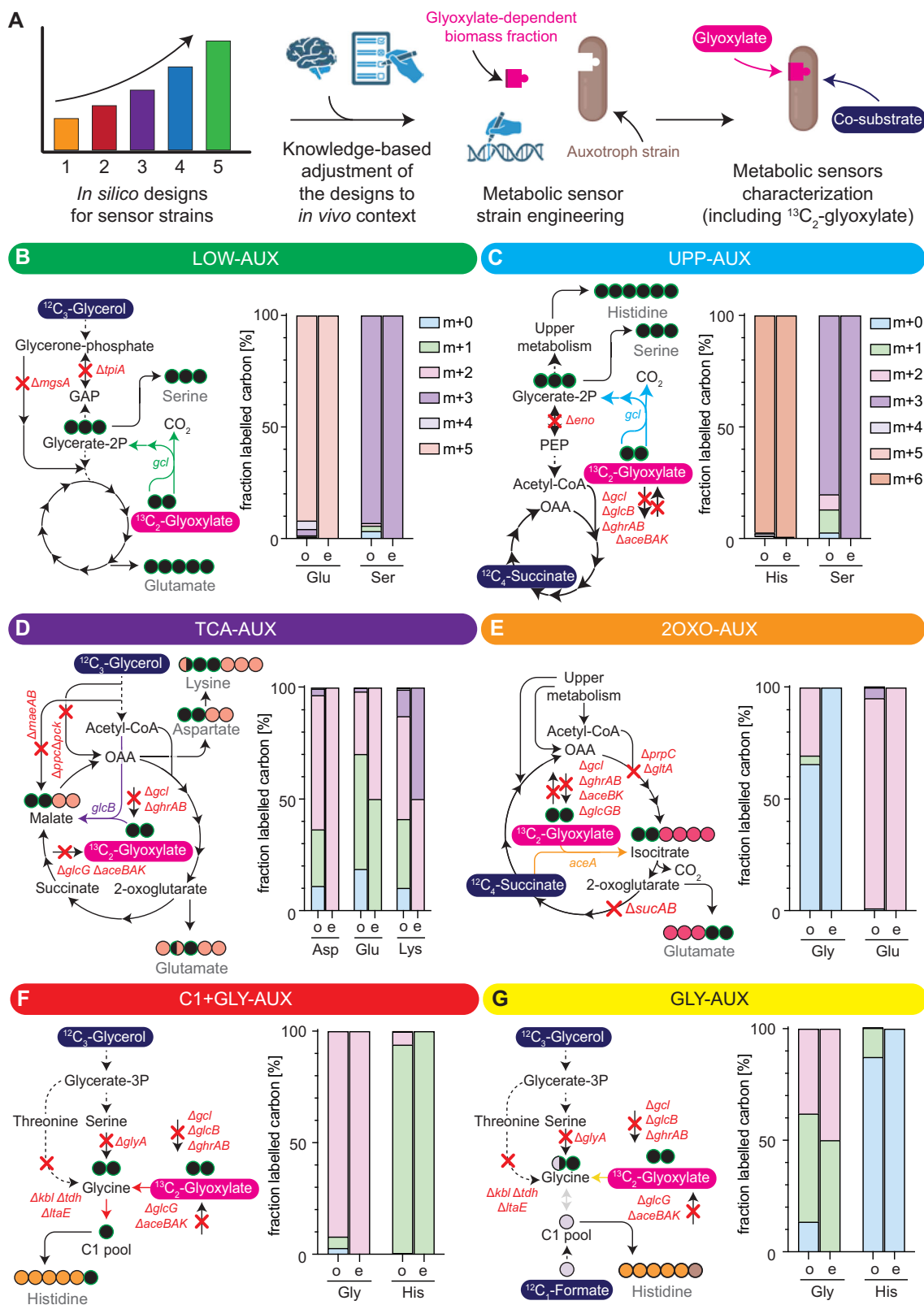
Then, we implemented the designs interrupting TCA cycle anaplerosis. We started with a strain including inactivations of PEP carboxylase (*ppc*), PEP carboxykinase (*pckA*), and isocitrate lyase (*aceA*)²⁹. In addition, the strain included adaptations to streamline glyoxylate utilization through constitutive, chromosomal overexpression of malate synthase (*glcB*, Fig. 2D). Moreover, we deleted *maeA* and *maeB* (encoding two malic enzymes, both canonically involved in gluconeogenesis but thermodynamically reversible under certain conditions), *ΔaceB* (additional copy of malate synthase) and *ΔaceK* (regulator of isocitrate dehydrogenase), *Δgcl* (glyoxylate carboligase, deleted to prevent wasteful conversion of glyoxylate through tartronate semialdehyde), and *ΔghrA* and *ΔghrB* (glyoxylate/hydroxypyruvate reductase). Therefore, this corresponding TCA-AUX strain relies on malate synthase activity for the condensation of glyoxylate and acetyl-CoA for the biosynthesis of all TCA cycle intermediates (Fig. 2D). Labeling patterns in the TCA-AUX grown with $^{13}\text{C}_2$ -glyoxylate and unlabeled glycerol reflect the expected fluxes, with amino acids derived from TCA cycle intermediates predominantly labeled (Fig. 2D and Supplementary Figs. 2 and 3C).

The other TCA cycle-related design, which involved inactivation of citrate synthase, was realized by deleting *glcA* (citrate synthase) and *prpC* [2-methylcitrate synthase, reported to have promiscuous citrate synthase activity^{30,31}] (Fig. 2E). In addition to these deletions, we aimed to reduce the demand for glyoxylate by deleting *sucAB* (2-oxoglutarate dehydrogenase), thereby limiting the glyoxylate requirement for the synthesis of the biomass precursor 2-oxoglutarate (2OXO-AUX strain). This strain required additional supplementation of succinate to prevent succinyl-CoA auxotrophy (Supplementary Fig. 4). To prevent undesired glyoxylate consumption through other routes we abolished suspected glyoxylate sinks by adding further deletions as in the TCA-AUX, i.e., *Δgcl*, *ΔghrAB*, *ΔglcB*, and *ΔaceBK*. Isotopic ^{13}C labeling patterns of the 2OXO-AUX strain growing with $^{13}\text{C}_2$ -glyoxylate and unlabeled succinate confirmed the incorporation of glyoxylate into 2-oxoglutarate with 100% of glutamate being labeled twice ($m+2$), while the other TCA cycle-derived amino acids aspartate, lysine, and threonine remained unlabeled ($m+0$) (Fig. 2E, Supplementary Figs. 2 and 3D).

In the third and last selection strategy, we pursued transamination of glyoxylate for the synthesis of the essential biomass precursor glycine. Glycine is also involved in the synthesis of the essential C1 metabolites [5,10-methylene-tetrahydrofolate (methylene-THF) and 10-formyl-THF]. Thus, we envisioned two setups for glyoxylate sensor strains depending on the inclusion or not of the C1 pool, named C1+GLY-AUX and GLY-AUX strains, respectively (Fig. 2F, G). First, we implemented deletions that force glyoxylate conversion to glycine through transamination (*Δgcl* *ΔghrAB* *ΔglcB*, and *ΔaceBAK*). Second, we prevented glycine synthesis from glycerol by deleting *glyA* (serine hydroxymethyltransferase), *kbl* (2-amino-3-ketobutyrate CoA ligase), and *ltaE* [low-specificity L-threonine

aldolase³²]. We demonstrated that this C1+GLY-AUX strain grew in a glyoxylate-dependent manner (described in detail in the next section), thus suggesting that one of the endogenous transaminases is employed for glyoxylate conversion to glycine. Indeed, isotopic labeling patterns of C1+GLY-AUX grown with $^{13}\text{C}_2$ -glyoxylate and unlabeled glycerol confirmed the ability to support glycine synthesis from glyoxylate by showing 100% of the glycine pool being labeled twice ($m+2$) (Fig. 2F and Supplementary Figs. 2 and 3E). Moreover, we proved the glyoxylate contribution to the synthesis of the C1 pool by measuring 100% of the histidine pool with one label, since C1 moieties are involved in the synthesis of this amino acid³³. To gain insights into native (promiscuous) transaminases involved in the conversion of glyoxylate to glycine, we searched for differences in protein abundances between the C1+GLY-AUX and a wild-type strain grown either with 20 mM glycerol + 1 mM glyoxylate or 20 mM glycerol + 5 mM glycine (Supplementary Fig. 5). Here, we found that of all known > 20 *E. coli* transaminases, *hisC* (encoding histidinol-phosphate transaminase) was the only one that was significantly upregulated in both comparisons of C1+GLY-AUX vs the wild-type strain on 20 mM glycerol + 1 mM glyoxylate and of C1+GLY-AUX grown with glyoxylate vs with glycine. Additionally, we observed an upregulation of allantoin metabolism genes (*allB* and *allC*) in C1+GLY-AUX grown with glyoxylate in all comparisons, which is in line with previous reports of glyoxylate-dependent expression of these genes^{34,35}. While the degradation of the purine synthesis intermediate allantoin proceeds via a ureidoglycine:glyoxylate transaminase in other organisms (PucG in *Bacillus subtilis* or HpxJ in *Klebsiella pneumoniae*), *E. coli* MG1655 is not known to harbor a homolog. However, when we used the DELTA-BLAST domain search and PucG from *B. subtilis* (Uniprot ID A0A6M4JLB6) as a query, histidinol-phosphate transaminase (HisC), which was upregulated in our proteomics data, was amongst the hits with the highest identity score (18.32% identity, e -value $1e^{-17}$). While these findings hint towards the relevance of HisC for glyoxylate use in the C1+GLY-AUX, it remains to be investigated whether HisC can promiscuously interconvert glyoxylate and glycine using ureidoglycine as an amine donor.

To further enhance the selection sensitivity through the same rescue mechanism, we engineered an additional strain GLY-AUX (derived from C1+GLY-AUX) with modifications including heterologous gene expressions not represented in the model. This strain expressed a specific enzyme from *Paracoccus denitrificans* (*bhcA*) that helps convert glyoxylate to glycine thanks to its aspartate:glyoxylate transaminase activity²⁶. Additionally, we modified the strain to produce methylene-THF from formate by integrating the first module of the reductive glycine pathway directly into its chromosome³⁶. These modifications are expected to increase the strain's sensitivity to glyoxylate. Remarkably, the strain showed a reduced pool of fully labeled glycine ($m+2$ pool reduced of 60% compared to C1+GLY-AUX strain; Fig. 2G), thereby suggesting a partial contribution of formate to the synthesis of glycine through the reverse glycine cleavage system. Moreover, the release of glyoxylate contribution to histidine was confirmed by observing a



drastic change in the histidine labeling pattern, with the majority of this amino acid pool remaining unlabeled ($m+0$) (Fig. 2G and Supplementary Figs. 2 and 3F). Once we confirmed that all strains required glyoxylate for the synthesis of one or more essential biomass components, we progressed further into their quantitative characterization.

Metabolic sensors can detect glyoxylate over a wide concentration range spanning three orders of magnitude

We further analyzed the metabolic sensors by monitoring their growth with varying supplemented concentrations of glyoxylate (Fig. 3A). This approach is routinely used in the characterization of metabolic sensors^{5,25,37} and allows identifying the range of direct correlation

Fig. 2 | Auxotrophic metabolic sensor strains engineering and characterization. **A** The workflow for creating the metabolic sensors starts with the output from the algorithm, which is complemented by curated literature knowledge on the host's metabolic network. The strains are engineered accordingly, and the auxotrophic phenotype for the target molecule (glyoxylate) is characterized and confirmed. Created in BioRender. Orsi (2025) <https://BioRender.com/x91g989>.

B–G Overview of the metabolic sensors and the corresponding ^{13}C -labeled amino acids pattern upon cultivation with $^{13}\text{C}_2$ -glyoxylate (magenta). For each scheme, the unlabeled co-substrate is highlighted in navy blue. For each strain, we proposed the labeling pattern observed ('o') and expected ('e') in terms of label incorporation. Labeling patterns are likewise identified with colored circles to trace the fate of the ^{13}C atoms within the metabolic network. The color code for the mass isotopomers ($m + 0$, $m + 1$, $m + 2$, etc.) is the same for panels B through G. Abbreviations for the

genes *mgsA* methylglyoxal synthase, *tpiA* triose-phosphate isomerase, *eno* enolase, *gcl* glyoxylate carboligase, *gclB* malate synthase G, *ghrA* glyoxylate/hydroxypyruvate reductase A, *ghrB* glyoxylate reductase, *aceA* isocitrate lyase, *aceB* malate synthase A, *aceK* isocitrate dehydrogenase kinase/isocitrate dehydrogenase phosphatase, *maeA* malate dehydrogenase (oxaloacetate-decarboxylating), *maeB* malate dehydrogenase (oxaloacetate-decarboxylating) (NADP⁺), *ppc* phosphoenolpyruvate carboxylase, *pck* phosphoenolpyruvate carboxykinase (ATP), *glcA* citrate synthase, *prpC* 2-methylcitrate synthase, *sucA* 2-oxoglutarate decarboxylase, thiamine-requiring, *sucB* 2-oxoglutarate dehydrogenase E2 subunit, *kbl* 2-amino-3-ketobutyrate CoA ligase, *tdh* threonine dehydrogenase, *ltaE* low-specificity L-threonine aldolase, *ghyA* serine hydroxymethyltransferase. Source data are provided as a Source Data file.

between the concentration of the limiting substrate (glyoxylate) and the final biomass concentration (Fig. 3B). This is determined by plotting the maximal optical density (maxOD_{600}) as a function of the concentration of the target metabolite. For our sensitivity analysis, we focused on an operational range giving linear response between substrate concentration and maxOD_{600} , i.e., excluding those concentrations where glyoxylate is no longer the limiting metabolite (upper limit) and those where we observed lack of growth because the concentration of glyoxylate is too low (lower limit). We confirmed that the experimentally determined glyoxylate demand was qualitatively consistent with the sensitivity ranking based on predicted “glyoxylate-to-biomass ratios” (GBR) calculated beforehand in flux balance analysis (Fig. 3B). In combination, the different constructed metabolite sensor strains enabled a glyoxylate sensitivity range covering three orders of magnitude from 10 μM to 20 mM of glyoxylate (Fig. 3B and 1). We confirmed robustness in auxotroph phenotypes for all strains by ruling out any unexpected growth behavior over several days of incubation (Supplementary Figs. 6 and 7). The upper limit of detection was determined by growing a wild-type strain with glyoxylate serving as the only carbon source, i.e., representing the highest achievable glyoxylate demand (Supplementary Fig. 8).

Notably, in some strains, the growth rate varied depending on supplied glyoxylate concentrations. This was most likely caused by factors affecting key limiting enzymes required for growth. For example, C1 + GLY-AUX and GLY-AUX rely on different transamination reactions to recover growth. Consequently, their growth rate differed significantly (Fig. 3A and Supplementary Fig. 5). The faster growth of C1 + GLY-AUX was probably due to an increased rate of glycine biosynthesis from glyoxylate caused by the additional engineered expression of a heterologous transaminase (*bhcA*). We could confirm this hypothesis by observing an improved growth rate of the GLY-AUX strain when *BhcA* was additionally produced in this strain background (Supplementary Fig. 6).

A functional glycolate dehydrogenase complex extends the sensors' detecting ability to glycolate

Once the ability of the sensors to detect glyoxylate was validated, we aimed to extend the sensing ability of the strains to glycolate. The glycolate dehydrogenase complex (GlcDEF) catalyzes the oxidation of glycolate to glyoxylate and is native to *E. coli*. Due to this activity, we expected glycolate to be converted to glyoxylate for subsequent formation of the selected biomass precursors depending on the metabolic context of each sensor strain. Accordingly, when cultivated on glycolate the strains showed comparable growth dependencies as observed with glyoxylate (Fig. 3C). Moreover, when plotting the maximum OD_{600} in response to the glycolate concentration provided, the sensor strains could cover a range of concentrations of three orders of magnitude (from 10 μM to 20 mM) (Fig. 3D). As had been done for glyoxylate, the maximum glycolate detection capacity was determined by growing the wild-type strain on glycolate as carbon source (Supplementary Fig. 7).

Altogether, the set of metabolic sensors showed a strict dependence on glyoxylate or glycolate for growth, a dependency that can be exploited to couple growth to modules forming either molecule. In the next sections, we showcase two exemplary applications of these metabolic sensors for the screening of enzymatic activities in vivo or for the measurement of extracellular glycolate in spent cultivation media.

In vivo testing of malate thiokinase and malyl-CoA lyase through glyoxylate sensing

Next, we investigated the use of our metabolic sensors for prototyping metabolic pathways (Fig. 4A). A versatile metabolic module employable in various pathways is the two-step reaction sequence catalyzed by malate thiokinase (MtkAB) and malyl-CoA lyase (Mcl) (Fig. 4B, C). The joint activity of these enzymes activates malate to malyl-CoA (at the expense of ATP) and then cleaves it to generate acetyl-CoA and glyoxylate. This module is part of natural metabolic routes, including natural and modified variants of the serine cycle³⁸, but is also required for the HydrOxyPropionyl-CoA/Acrylyl-CoA (HOPAC) cycle²¹, a synthetic route for CO_2 fixation so far demonstrated only in vitro (Fig. 4B). Moreover, this module here has been proposed in the context of the reverse glyoxylate shunt²², which in theory allows generation of two C2 compounds from one C4 moiety, and therefore holds potential for achieving higher product yields for C2-dependent productions from sugars²². Therefore, we reasoned that screening the combined MtkAB and Mcl activity using our metabolic sensors could create new opportunities for pathway engineering. In fact, while the module had previously been shown to rescue an acetyl-CoA auxotrophic strain³⁹, engineering of the full HOPAC cycle requires functionality in a glyoxylate auxotroph because the cycle product is glyoxylate. Hence, we chose to demonstrate the feasibility of the module in our sensor strains as a proof-of-principle for glyoxylate selections.

Growth in the metabolic sensors is expected to be a trade-off between flux capacity through the target module and the metabolic burden associated with its production. Using the growth rate of the metabolic sensor as a proxy for this trade-off, we proceeded by testing how the module capacity is affected by modulating the expression of the required genes (*mtkAB* and *mcl*). We hypothesized that if modules display different flux capacities, they might differ in their growth rate, yet they should not differ in their final biomass density because they have access to the same nutrients in the cultivation medium.

We cloned *mtkAB* and *mcl* under the translational control of RBSs, RBS^A, RBS^B, and RBS^C as strong, medium, and weak ribosome binding sites, respectively³. Then, we transformed this library into three metabolic sensors: C1 + GLY-AUX (Fig. 4D, E), 2OXO-AUX (Supplementary Fig. 9A–D), and UPP-AUX (Supplementary Fig. 9E, F). All C1 + GLY-AUX strains grew to the same maxOD_{600} (Fig. 4E), in line with our above-mentioned expectation that the growth rate would correlate with module flux, while the maxOD_{600} should be comparable (Fig. 4F). Indeed, different expression constructs resulted in different growth rates, with two constructs resulting in faster growth as a

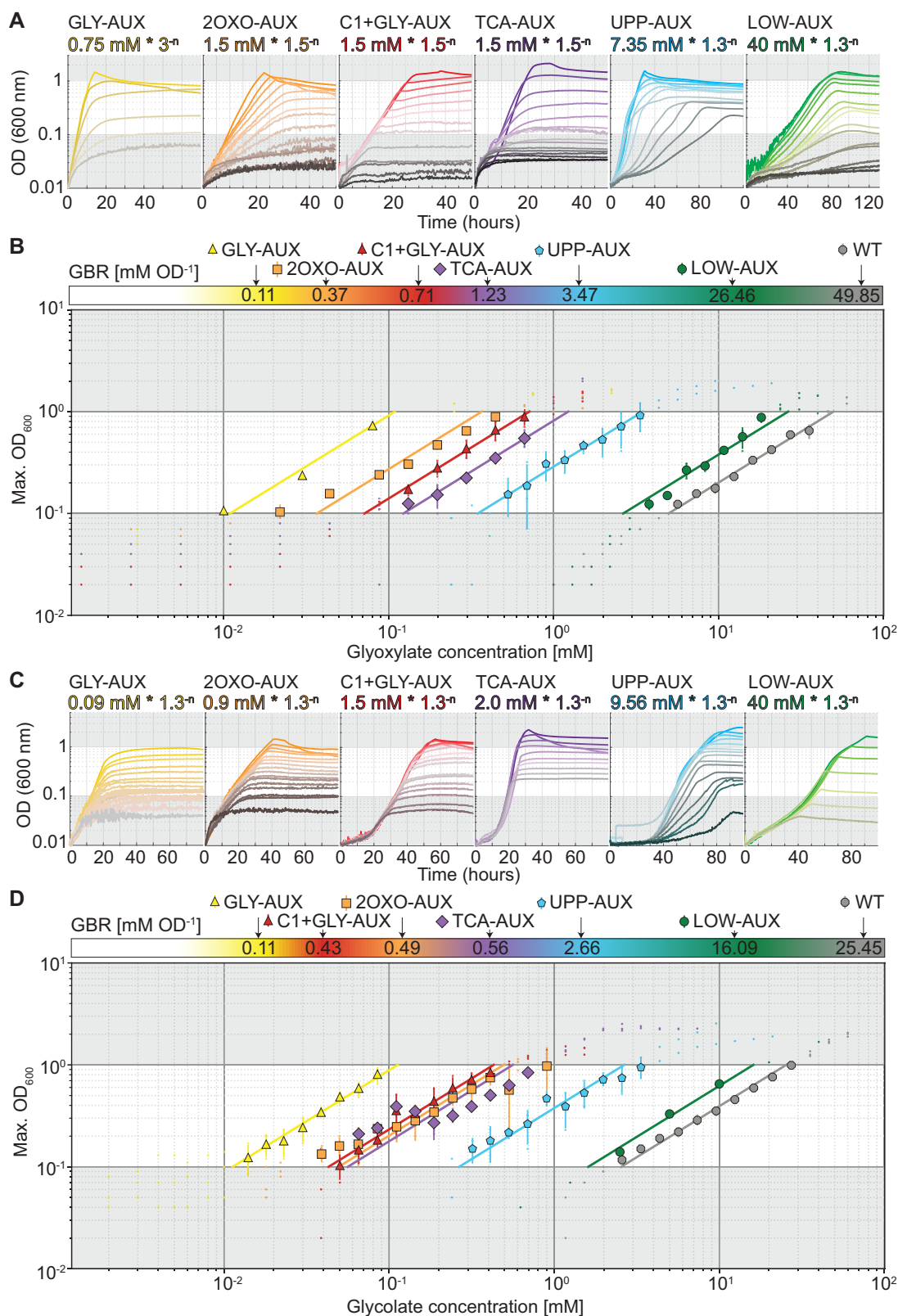
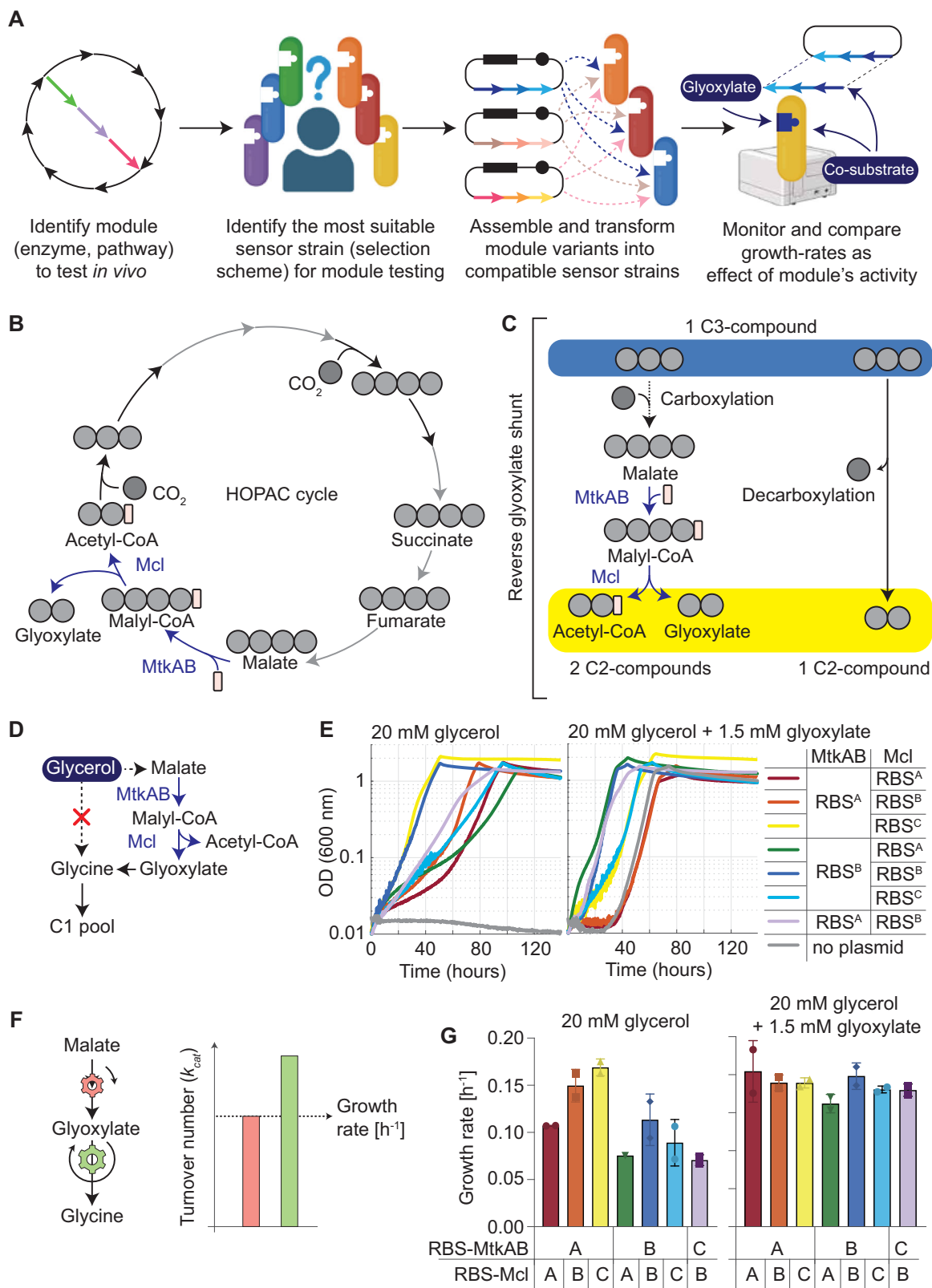


Fig. 3 | Growth profiles and dose-response relationship in the auxotrophic metabolic sensors. **A** Growth profile of the different metabolic sensors with varying concentrations of glyoxylate supplemented in the growth medium. For each strain, the highest concentration of glyoxylate was reported, together with the dilution factor used. **B** Relative distribution of the glyoxylate-to-biomass ratio (GBR) for the different strains, calculated by flux balance analysis using the medium-scale model, and correlation between the maximal OD₆₀₀ measured and

the initial glyoxylate concentration in the medium. This relationship highlights the linear trend of dose-response to glyoxylate at OD₆₀₀ between 0.1 and 1.0. **C**, **D** The same experiments as in A and B but focused on glyoxylate as the target metabolite. Experiments were conducted in 96-well plates in triplicates, which displayed identical growth curves ($\pm 5\%$) and were averaged. Source data are provided as a Source Data file.



reflection of the combined effect of high flux through MtkAB and Mcl and low metabolic burden (Fig. 4G). The requirement of a strong RBS^A for *mtkAB* and weaker RBSs (RBS^B or RBS^C) for *mcl* suggests that the rate-limiting step of the module is the step catalyzed by MtkAB. Although the translation-initiation rates are strongly influenced by the genetic context both upstream and downstream of the RBS, we conclude that the RBS^A-*mtkAB*-RBS^B-*mcl* and RBS^A-*mtkAB*-RBS^C-*mcl*

constructs provide the best trade-off and, therefore, the highest flux capacity to support growth through the module.

Notably, the two other tested selection strains were not suitable for such turnover estimations. The 20XO-AUX strains showed differences in the final biomass density, as well as in growth rates, and required supplementation with three times more succinate when MtkAB-Mcl was expressed (Supplementary Fig. 9B). The different

Fig. 4 | Application of the auxotrophic metabolic sensors for in vivo enzyme screening operations. **A** Workflow for using metabolic sensors with different designs and demands for the screening of metabolic modules. Created in BioRender. Orsi (2025) <https://BioRender.com/x91g989>. **B, C** Lumped architectures of the HydrOxyPropionyl-CoA/Acrylyl-CoA (HOPAC) cycle and reverse glyoxylate shunt. **D** Selection scheme of C1 + GLY-AUX sensor strain, including MtkAB and Mcl activities. **E** Growth profile of the RBS library was tested under selective (20 mM glycerol) and non-selective (20 mM glycerol + 1.5 mM glyoxylate) conditions within the C1 + GLY-AUX strain. In this type of application for the sensor strain, the growth rate is the proxy for the module activity (metabolic flux). Experiments were

conducted in 96-well plates in triplicates, which displayed identical growth curves ($\pm 5\%$) and were averaged. **F** Schematic presents how the turnover number of the rate-limiting enzyme determines the growth rate of the selection strain. If, e.g., the rate-limiting step is the conversion of malate into glyoxylate (as expected in our experimental setup), that step will dictate the growth rate of the strain. **G** Quantification of the growth rates for the different candidates under selective and non-selective conditions. Experiments were conducted in 96-well plates in duplicates and were averaged. The error bars correspond to the standard deviation calculated on those replicates. Created with Biorender.com. Source data are provided as a Source Data file.

behavior of 2OXO-AUX could be a consequence of the selection scheme architecture. With C1 + GLY-AUX the fate of the carbon provided by the substrate (glycerol) follows a linear path all the way to glyoxylate (Fig. 4D), whereas in 2OXO-AUX succinate is needed both for production of glyoxylate, as well as for its further condensation to form isocitrate (Supplementary Fig. 9C). Therefore, MtkAB-Mcl and isocitrate lyase may be competing for the succinate pool, which affects the pool of 2-oxoglutarate and therefore the ability to rescue strain growth (Supplementary Fig. 9A–D). The last design tested (UPP-AUX) did not result in the growth of any of the combinations screened (Supplementary Fig. 9E, F). This suggests that the demand from the strain is too high to be supported by the module through its own turnover rate (Supplementary Fig. 9F), which implies that the current module flux is incapable of supporting growth when the GBR is 0.13 or higher.

We thus demonstrated that different translation rates of MtkAB and Mcl affect the growth rates of the metabolic sensors. These results can be used as a basis to further investigate the corresponding fluxes for glyoxylate production in vivo. In the next section, to demonstrate a second application of the metabolic sensors, we focused on quantifying extracellular glycolate in the spent medium of microbial cultivations.

Detection of glycolate produced through photorespiration

Another aim of this study was to demonstrate the feasible use of metabolic sensors to monitor phenomena of ecological relevance. From an environmental perspective, glycolate synthesis occurs during phytoplankton blooms in the ocean⁴⁰. During this event, glycolate synthesis is caused by a metabolic process called photorespiration, which is known to limit CO₂ fixation via the Calvin-Benson-Bassham (CBB) cycle in phototrophic organisms^{41,42}. Here, photorespiration occurs through the promiscuous oxygenase activity of the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and limits the CO₂ fixation rate of this enzyme^{43,44}.

Photorespiration has been associated with the secretion of glycolate in photoautotrophic microorganisms in the range of $\mu\text{g/L}$ ($<1\text{ mM}$, in open waters) to g/L ($>30\text{ mM}$, under controlled conditions)^{45,46}. Therefore, we reasoned that the sensitivity range of our sensor strains should be able to detect glycolate in the spent medium of phototroph cultures (Fig. 5A). Importantly, for this purpose, we reasoned that the presence of glycolate will restore the growth of our metabolic sensors and that the maxOD₆₀₀ will be proportional to the concentration of glycolate in the media.

We considered two phototrophic organisms where glycolate secretion was already demonstrated: *Synechococcus elongatus*⁴⁶ (which grows in freshwater) and *Chlamydomonas reinhardtii*⁴⁵ (which grows in salt water). We observed that only the low-salt medium used for *S. elongatus* cultivation was compatible with our *E. coli* strains (Supplementary Fig. 10A, B), and therefore we further investigated glycolate secretion by this cyanobacterium using both wild-type and *Accm* *S. elongatus* strains. The *Accm* strain lacks the carboxysome, a micro-compartment that permits increased local CO₂ concentrations around RuBisCO and enhances the enzyme's carboxylation rate (therefore the

Accm strain is known to exhibit higher rates of RuBisCO oxygenation in ambient CO₂ conditions⁴⁷).

In the experimental setup, we performed cultivation as follows: (i) high CO₂ concentrations (3%) for three days (as negative control where photorespiration should be minimal); (ii) alternatively, we cultivated the strains at 3% CO₂ for three days to reach high biomass densities, followed by an additional three days at 0.5% CO₂ to stimulate photorespiration. We collected samples at the end of these cultivations, removed cyanobacterial cells by centrifugation at 13,000 $\times g$, and mixed spent medium in a 1:1 ratio with fresh M9 medium for the application of the different metabolic sensors. We then monitored the growth of the *E. coli* strains until the stationary phase was reached (Fig. 5B) and used the maxOD₆₀₀ (adjusted to a 1:1 dilution) to infer glycolate concentration from glyoxylate-to-OD₆₀₀ correlation as discussed above (Fig. 3D). As the control, we also measured the glycolate concentrations in the spent media samples using conventional analytics (HPLC). As expected, the *Accm* strain grown under photorespiratory conditions yielded the highest glycolate concentration, whereas the wild-type strain grown constantly at high CO₂ resulted in the lowest concentration (Fig. 5B). This result was confirmed by the maxOD₆₀₀ values in all three sensor strains, which thus suggests that optical density measurements can serve as at least a coarse estimate of the relative glycolate concentration within the samples.

We continued by determining which of the three strains was most precise in sample detection. From the HPLC data, we could confirm that the concentration of the samples ranged from 10 μM to 50 μM glycolate (Fig. 5B), which is within the linear range of our most sensitive GLY-AUX strain (Fig. 3D). In fact, when plotting the sensor-derived concentration estimates of glycolate against those measured with HPLC, we could determine a slope of 0.93 for the GLY-AUX measurements ($R^2 = 0.992$) (Fig. 5C). On the other hand, since the other two strains 2OXO-AUX and TCA-AUX presented a linear range for glyoxylate-to-OD₆₀₀ correlation only at higher glycolate concentrations (Fig. 3D), we observed a greater discrepancy between HPLC-derived and sensor-derived measurements for these strains, which resulted in a slope of 2.37 for 2OXO-AUX and 0.75 for TCA-AUX. Therefore, when glycolate concentrations lay within the linear range of detection of the sensor, measurement of the strain's maxOD₆₀₀ can be a reliable readout of the concentration of the real target molecule in the medium.

Discussion

AMS are valuable tools for several biological applications, including synthetic biology and environmental monitoring. Yet their implementation is not trivial because the created auxotrophic phenotypes must be robust in various contexts. Achieving such phenotypes requires deep knowledge of the host's metabolism. In our study, we streamlined the workflow for creating such sensor strains by combining the *in-silico* design of AMS with their engineering and physiological characterization together with in vivo applications. We selected glyoxylate as the metabolite of interest around which we created a set of sensor strains because this metabolite is not essential for growth and therefore creating ad hoc growth-coupled designs requires

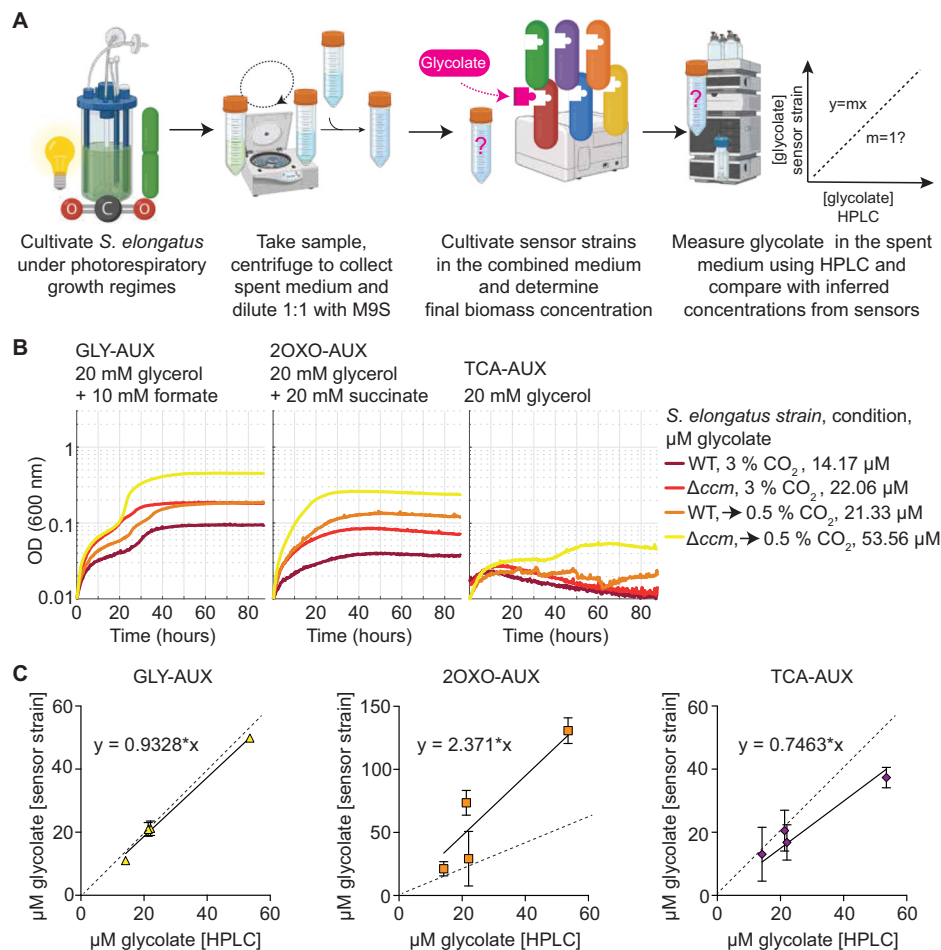


Fig. 5 | Application of the auxotrophic metabolic sensors for the determination of an extracellular metabolite (glycolate) in the spent medium of cultivation (photoautotrophic growth). **A** Workflow depicting the steps required for the use of metabolic sensors for the determination of the extracellular metabolite in the spent medium of phototrophic *Synechococcus elongatus* strains grown under photorespiratory regimes. Created in BioRender. Orsi (2025) <https://BioRender.com/x91g989>. **B** Growth profile of three metabolic sensors (GLY-AUX, 2OXO-AUX, and TCA-AUX) on the spent medium for the *S. elongatus* strains. The measured glycolate concentrations are shown for each of the conditions tested (cultivation at a constant 3% CO_2 , or transition from 3% to 0.5% CO_2 to induce photorespiration in

a wild-type (WT), or carboxysome mutant (Δccm) background. Experiments were conducted within 96-well plates and were performed in triplicates, which displayed identical growth curves ($\pm 5\%$), and hence were averaged. **C** Correlation between the data obtained from HPLC analyses (x-axis) and the concentrations determined by observing the final biomass concentration of different metabolic sensors (y-axis). The plotted data corresponds to the average of the three biological replicates, and the error bars represent the standard deviation on those measurements. A good correlation should correspond to a slope close to 1. Source data are provided as a Source Data file.

rewiring of the host's metabolism. Through this workflow, we created a set of six sensor strains available off-the-shelf for the same screening purpose. We showcased their application in two practical cases where growth rate or final optical density served as a proxy for prototyping a segment of an engineered metabolic pathway and for environmental monitoring, respectively. Having access to multiple strains allowed us to successfully accomplish both these screening investigations.

To implement the sensor construction workflow, we used a medium-scale metabolic model¹⁴ to predict enzyme inactivations that force growth to be dependent on glyoxylate. The procedure for finding suitable deletion candidates had been previously applied to design strains that depend on CO_2 fixation via RuBisCO¹², glycerate sensor strains²⁵, or synthetic methylotrophs^{48,49}. Yet, in all these cases, computational analysis relied on a metabolic core model⁵⁰ which covered only glycolysis/gluconeogenesis, the pentose phosphate pathway, the TCA cycle, and oxidative phosphorylation. In contrast, the larger size of the model used in this study allowed us to explore more designs than with the core *E. coli* model while reducing computation time and unfeasible solutions typical of genome-scale models⁵¹. Thanks to the algorithm, we could identify three groups of design strategies

(glycolytic disruptions, interruption of TCA anaplerosis, and transamination of glyoxylate to glycine) which we further engineered in the sensor strains. In particular, the glyoxylate to glycine transamination group of solutions was possible thanks to this medium-scale model because it relied on enzyme inactivations within the amino acid biosynthetic network.

We needed to supplement the algorithm's list of candidate enzymes with additional gene deletions to effectively implement robust AMS. While the algorithm allowed us to rapidly identify non-trivial growth-coupling strategies, our reliance on further manual interventions underscores a limitation of the current pipeline that could be subjected to further optimization in the future. Most of these interventions were inspired by data available in databases such as EcoCyc (<https://ecocyc.org/>)⁵². We applied the following types of manual interventions to complement the model: deletion of latent metabolic pathways which may become activated, e.g., ΔmgsA ^{27,28} in LOW-AUX; addition of heterologous modules to relieve metabolic demands, i.e., C1 moieties, as in the case of the C1 module from the reductive glycine pathway³³ in GLY-AUX; limitation in reaction annotations from a thermodynamic perspective, as in the case of *maeAB*,

since ME1 and ME2 were considered irreversible in metabolic models; limitation in annotations as an effect of promiscuous reactions, such as in the case of *ltaE* (in C1 + GLY-AUX and GLY-AUX) and *prpC* (in 2OXO-AUX), which can complement inactivation of *glyA*³² and *gltA*³⁰, respectively. Ultimately, in vivo characterization of the designed strains confirmed the model predictions in terms of the ranked sensitivity of the strains towards glyoxylate (and glycolate).

The main application of AMS is for testing enzyme modules by measuring growth as a proxy for their activity^{3,4}, where growth reflects a trade-off between the module's flux capacity and the metabolic burden for its synthesis. The two-carbon molecule glyoxylate is a product of CO₂ fixation pathways, such as the serine cycle or the in vitro crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle and the HOPAC cycle^{21,53}, and can be synthesized through the reverse glyoxylate shunt from pyruvate bypassing its decarboxylation to acetyl-CoA^{22,54}. We tested MtkAB and Mcl as key components of the reverse glyoxylate shunt and HOPAC cycle^{21,22}, with indications that a stronger expression of MtkAB is required compared to Mcl. Experimental data from testing different modules suggested that the choice of the selection scheme (and therefore metabolic sensor) also has an impact on the ability to screen for modules. We demonstrated that strains with high metabolic demands or competitive nodes for substrate utilization might not be suitable for module testing. The next steps in this research direction include adding upstream modules of the HOPAC cycle to the selection strains, such as those that enable the conversion of methylmalonyl-CoA to malate, to facilitate in vivo testing²¹. Shifting the focus to other CO₂ fixation or C1-assimilation pathways to engineer, the generation of sensor strains for, e.g., pyruvate, acetyl-CoA, glycerone phosphate, and glyceraldehyde 3-phosphate would provide the community with additional platforms where to test natural and synthetic designs in surrogate hosts. In principle, this approach could also be used for product formation.

Furthermore, we demonstrated how the estimation of extracellular metabolites (glycolate) in spent cultivation medium is most precise when using AMS whose linear range of detection overlapped with the target molecule concentration in the medium. In the literature, a similar approach was used to determine mevalonate in spent medium but that approach used growth rate as the benchmark⁵⁵. Instead, here we demonstrated how the final biomass concentration can also be used as a proxy for determining the concentration of the target molecule in the spent medium, when interpolated with the compound-to-OD₆₀₀ correlation, as also recently demonstrated using whole cell biosensors⁵⁶. The glycolate AMS we created exhibited a linear range of detection within the concentration range 0.01–20 mM, thus adding to a list of transcription-based [0.1–200 mM range⁵⁷ or 0.01–20 mM range⁵⁸] and enzyme-based [0.01–1 mM range⁵⁹] biosensors previously described in the literature. We argue that this simple setup for measuring the molecule of interest can be of use when performing high-throughput screening using multi-well-plate setups. Assuming an average duration of 0.5 h per sample for an HPLC method targeting organic acid detection, screening at least 60 samples with our detection method results in faster analysis turnover. Our sensors reached their final biomass in approximately 30 h, making this the total analysis duration for glycolate detection in spent fermentation media—equivalent to the time required to analyze 60 samples sequentially via HPLC. Additionally, using cell turbidity rather than fluorescence signals, as in transcription-based biosensors, simplifies the readout process.

We postulate that using a computational algorithm complemented by manual interventions will streamline the creation of growth-coupled designs for various metabolites of interest. The AMS generated in the presented study will allow us to explore novel designs for synthetic pathways leading to the synthesis of glyoxylate. Moreover, the auxotrophic phenotype can be exploited in this way for the evolution of (new) key enzymatic steps leading to the synthesis of the

target metabolite^{60–63}, as well as in high-throughput automated setups^{64–66}. Since glyoxylate is described as a key intermediate of some protometabolic pathways^{15–18}, we believe our sensor strains could create new possibilities for explorative studies within this field of research. Finally, with glycolate as a marker for photorespiration, we encourage the use of such AMS for the high-throughput study of photorespiration by other organisms, which has broad implications for agriculture^{67–69} and environmental studies^{41,70}.

Methods

The stoichiometric model

In order to identify promising knockout strategies in *E. coli* that create glyoxylate auxotrophy at different dependency levels, we used the recently published iCH360 model¹⁴. iCH360 is a subnetwork of the much larger genome-scale model, focusing on central metabolic subsystems that carry relatively high flux, are central to maintaining and reproducing the cell, and provide precursors and energy to engineered metabolic pathways. This medium-sized model, by doing away with low-flux and secondary pathways and enzymes, facilitates applications like ours since these redundant reactions typically create bypasses that are unlikely to be relevant in vivo and greatly complicate the search. However, since we wanted to design auxotrophic strains to a non-standard carbon source (glyoxylate), we had to augment the iCH360 model with two metabolites and 5–6 reactions for it to be able to deal with glyoxylate metabolism:

Metabolites:

- Extracellular glyoxylate (glx_e): formula = C₃H₃O₄
- 2-Hydroxy-3-oxopropanoate (2h3oppa_n_c): formula = C₂H₃O₃

Note that 2-Hydroxy-3-oxopropanoate and tartronate semialdehyde are synonyms.

Reactions:

- Glyoxylate exchange (EX_glx_e): glx_e ⇌
- Glyoxylate transport (glx_t): glx_e + h_c ⇌ glx_c + h_c
- Aspartate-glyoxylate transaminase (BHC): asp_L_c + glx_c ⇌ oaa_c + gly_c
- glyoxylate carboligase (GLXCL): glx_c + h_c ⇌ 2h3oppa_n_c + co2_c
- tartronate semialdehyde reductase (TRSARr): 2h3oppa_n_c + h_c + nadh_c ⇌ glyc_R_c + nad_c
- formate-tetrahydrofolate ligase (FTHFLi*): for_c + atp_c + thf_c ⇌ 10fthf_c + adp_c + pi_c

* The FTHFLi reaction was added only in the strain named C1 + GLY-AUX which utilizes formate as the source for all the C1 carbon metabolism.

Systematic search for growth-coupled designs

After establishing a suitable stoichiometric model as described in the previous section, we applied a search algorithm for identifying auxotrophic knockout strategies.

First, ATP maintenance reaction, as the algorithm is designed to only calculate the marginal dependence on glyoxylate (at low growth rates) and the maintenance reaction is not relevant for that calculation. Then, we set the bounds of the glucose exchange flux to 0 (instead of the default lower bound of −10 mmol/gCDW/h), and replaced it with succinate or glycerol as the abundant carbon source, by setting the lower bound to −1000 mmol/gCDW/h.

The next step is to iterate through all possible single or double knockouts of reactions from the list of central reactions—i.e., reactions that exist in the core model of *E. coli*⁵⁰ and the ones added to iCH360 (excluding the glyoxylate uptake itself). In some cases, we lumped together two reactions (denoted as REACTION1|REACTION2), either because they are arranged in one linear pathway with no branchpoints, catalyzed by the same enzyme, or the same chemical reaction with different cofactors (catalyzed by two isoenzymes):

- G6PDH2r
- PGL
- PGI
- PFK
- FBP
- FBA
- TPI
- PGK
- GAPD
- PGM
- ENO
- PYK
- PPS
- PDH
- PFL
- GND
- RPE
- RPI
- TKT1|TKT2
- TALA
- PPC
- PPK
- ME1|ME2
- SUCDi
- FUM
- CS
- ACONTa|ACONTb
- ICDHy
- ICL
- MALS
- AKGDH
- SUCOAS
- MDH
- ALCD2x
- ACALD
- GLXCL|TRSARr
- GHMT2r
- BHC

For each possible single or double knockout, we ran a series of flux-balance analyses (FBA) using the cobrapy toolbox⁷¹. We first tested whether it can grow at all on succinate. If it could, we ran another FBA with a high abundance of succinate, and a limiting amount of glyoxylate. The ratio between the maximal growth rate and the glyoxylate uptake rate represents that GBR. The same is repeated with glycerol instead of succinate.

We then normalized the GBR value by dividing them by the GBR of a wild-type cell growing on glyoxylate alone (without succinate nor glycerol). This brings the values to generally be between 0 and 1, except for a few cases where the KO makes the cell require even more glyoxylate (typically, this is not directly related to the glyoxylate itself, but rather a disruption that makes biosynthesis less efficient and therefore requiring more ATP). The results are summarized in Supplementary Fig. 1. Based on these results, we selected 5 designs that span the range of relevant GBR values (Supplementary Table 1).

We found it sufficient to consider only single and double knockout strains with either succinate or glycerol as an additional carbon source, as the 5 chosen strains spanned a large enough range of GBR values. Nevertheless, we ran the same procedure also for sets of 3 and 4 knockouts, and the results can be found in files stored in the Git repository (see Data and materials availability).

Strains and plasmids used in this study

All strains and plasmids employed in this study are listed in Supplementary Data 1. *E. coli* SJ488 (bacterial strain with catalog number

#68246 on addgene), which carries recombinases downstream an arabinose inducible promoter and a flippase downstream a rhamnose-dependent promoter⁷², was the base strain for all engineering efforts and was used as wildtype reference whenever required. NEB5 α cells were used for cloning. When testing the strains on glyoxylate, in some cases we generated ‘glyoxylate-only’ variants in which the glycolate dehydrogenase complex (GlcDEF) was removed. Although not essential for the strain’s function, this deletion was made to prevent possible futile cycles in the interconversion of glyoxylate to glycolate when growing on glyoxylate.

Gene deletions

Gene deletions were performed by λ -red recombineering or through Cas9-mediated base editing to introduce premature STOP codons⁷³. For λ -red recombineering, knockout cassettes were constructed by amplifying the antibiotic resistance cassette from pKD3 (chloramphenicol resistance, specified by ‘Cap’ in primer name) or pKD4 (kanamycin resistance) (both vectors were a gift from Barry L. Wanner; Addgene plasmids #45604 or #45605) with ‘KO’-primers introducing 50 bp overhangs with homology to the respective target locus using PrimeStar GXL polymerase (Takara Bio) or OneTaq polymerase (New England Biolabs). These ‘KO’-primers were designed following the KEIO knockout collection homology arms sequences⁷⁴. After purifying the cassette by PCR purification using the GeneJet PCR purification kit (Thermo Scientific, Dreieich, Germany), the target strain was electroporated with the fragment. For this, the cells were inoculated in LB medium and grown to an OD₆₀₀ of 0.3–0.5, when recombinase expression was induced by adding 15 mM L-arabinose and further incubation for 45 min at 37 °C. After harvesting the cells by centrifugation (13,000 \times g, 30 s, 2 °C) and washing them three times with ice-cold 10% glycerol, the cells were electroporated with ~300 ng of the deletion cassette (1 mm cuvette, 1.8 kV, 25 μ F, 200 Ω). The cells were plated on LB plates containing the relevant antibiotic. Successful gene deletion in grown colonies was confirmed by verifying the target locus size in a PCR using ‘ver’-primers (Supplementary Table 2) and DreamTaq polymerase (Thermo Scientific, Dreieich, Germany). To remove the antibiotic resistance cassette, a 2 mL strain culture was grown to an OD₆₀₀ of 0.5, when 50 mM L-rhamnose was added to induce flippase expression, followed by incubation for \geq 3 h at 30 °C. To confirm successful cassette removal, the colonies were tested for antibiotic sensitivity and the target locus size was confirmed by PCR with ‘KO-Ver’ primers and DreamTaq polymerase (Thermo Scientific, Dreieich, Germany).

For Cas9-mediated base editing, we followed the protocol provided by the original work⁷³. We constructed pMBEC plasmids containing up to six single-guide RNAs (sgRNAs) for targeting up to three genes (two sgRNA/gene) per round of mutation. The spacer sequences for premature STOP codon introduction were identified using the CRISPy-web tool (<https://crispy.secondarymetabolites.org/>)⁷⁵ and are listed in Supplementary Table 3.

Gene integration by P1 transduction

To reintroduce *glcDEF* in the TCAAUX strain²⁹, *glcB* was deleted by P1 transduction⁷⁶, thus transferring the *glcB* deletion locus with a kanamycin resistance from a Δ *glcB* donor strain (JW2943) from the KEIO collection⁷⁴ and the adjacent *glcDEF* wildtype genes. After the transduction, the cells were plated on LB-kanamycin plates containing 20 mM sodium citrate. Successful reintroduction of *glcDEF* was confirmed by PCR using the verification primers previously used to verify the deletion of *glcDEFGB* (Supplementary Data 1). Furthermore, the growth of the strain with glycolate instead of glyoxylate was confirmed in tubes.

Plasmid-based gene expression

For plasmid-based gene expression, the pZ-ASS plasmid (p15A origin of replication and strong promoter) was used³. Malate thiokinase (*mtkAB*

from *Methylococcus capsulatus*, as described by Liao et al.²²) and the codon-optimized variant of malyl-CoA lyase (*mcl* from *Rhodobacter spaeroides*) was obtained from pTE3262 (unpublished, both plasmids were provided by Dr. Shanshan Luo, for gene sequences see Supplementary Data 2). All genes were amplified with primers introducing ribosomal binding sites “A”, “B” or “C”⁷⁷ as indicated by the primer name, the coherent overlapping primers were used to amplify the pZ-ASS backbone (Supplementary Data 1). In front of all genes, a spacer of the sequence “TAATAGAAATAATTTTGTCTTAACCTTTA” was introduced. In addition, between *mtkB* and *mtkA* another spacer of the sequence “TCTAGAGCTAGCGTTGATCGGAGGTTCTGTTAAGTAACTGAACCC” was introduced, while between *mtkA* and *mcl* a spacer of the sequence “TGTCGTTAGTGACGCTTACCTCTT” was introduced to achieve non-redundant fragment overhangs. The fragments needed to obtain pZ-ASS-MtkBA-Mcl plasmids with combinations or different ribosomal binding sites were assembled by a HiFi DNA assembly as described for the HiFi assembly protocol (New England Biolabs, Ipswich, Massachusetts). After transforming the HiFi NEB5 α cells with the plasmid, the cells were plated on LB Streptomycin plates. Successful plasmid assembly was confirmed by whole-plasmid sequencing (plasmidsaurus, Eugene, Oregon).

Routine strain cultivation

For routine strain handling, Lysogeny broth (LB) medium (composed of 1% NaCl, 0.5% yeast extract, 1% tryptone) was used. When appropriate, antibiotics (kanamycin (25 μ g/mL), ampicillin (100 μ g/mL), streptomycin, (100 μ g/mL), or chloramphenicol (30 μ g/mL)) were added.

Plate reader experiments

For growth tests, M9 minimal medium without antibiotics was used (50 mM Na₂HPO₄, 20 mM KH₂PO₄, 20 mM NH₄Cl, 2 mM MgSO₄, 1 mM NaCl, 134 μ M EDTA, 100 μ M CaCl₂, 13 μ M FeCl₃·6H₂O, 6.2 μ M ZnCl₂, 1.62 μ M H₃BO₃, 0.76 μ M CuCl₂·2H₂O, 0.42 μ M CoCl₂·2H₂O, 0.081 μ M MnCl₂·4H₂O). Carbon sources were added as indicated in the text. For “relaxing” conditions of overnight precultures, 1.5 mM glyoxylate or glycolate was supplemented, depending on the experiment. For the experiment, 2 mL cultures were harvested by centrifugation (10,000 \times g, 30 s) and washed three times with “selective” M9 minimal medium in the absence of glyoxylate or glycolate. Then, the washed cells were diluted to a final OD₆₀₀ of 0.01 in 96-well microtiter plates (Nunc Delta Surface, Thermo Scientific). The medium composition (relaxing, selective, or relaxing with dilutions of glyoxylate or glycolate) was adjusted for each experiment. Each well of the 96 well plates contained 150 μ L of culture and 50 μ L mineral oil (Sigma-Aldrich) to prevent evaporation but allow gas exchange. A BioTek Epoch 2 plate reader (BioTek, Bad Friedrichshall, Germany) was used to monitor the growth of technical duplicates at 37 °C by measuring the absorbance (at 600 nm) every -10 min with intermittent orbital and linear shaking. During the analysis with MATLAB, blank measurements were subtracted and OD₆₀₀ values were converted to cuvette OD₆₀₀ values by multiplying with a factor of 4.35 which had previously been established for the instrument.

¹³C labeling experiments

For stationary isotope labeling of proteinogenic amino acids ¹³C₂-glyoxylate ([1,2-¹³C]-glyoxylic acid monohydrate, LGC Standards) was used as a tracer. As a control, sodium glyoxylate monohydrate (Sigma Aldrich) was used as an unlabeled substrate. All experiments were performed in biological triplicate ($n = 3$) as follows. Strains were grown in 4 mL LB. Subsequently, 40 μ L of culture was transferred into 4 mL of fresh M9 medium supplemented with labeled or unlabeled tracer plus the additional co-substrate (glycerol or succinate). To reduce the effect of LB carryover, once the OD₆₀₀ reached a level of 0.8, 40 μ L of grown culture were transferred again into 4 mL of the same M9

medium, and at the late exponential phase, the equivalent of 1 mL of culture at OD₆₀₀ of 1 was harvested. Samples were then processed to analyze proteinogenic amino acid mass-isotopomers through GC-MS⁷⁸. For analysis of amino acid isotopomer data, we considered fragments containing the full carbon backbone of the amino acids of interest⁷⁹. Raw data from the GCMS was integrated using SmartPeak⁸⁰. Processed data was further corrected for the natural abundance of isotopes in the derivatization agents used for GCMS analysis⁸¹.

Proteomics

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD054993. C1 + GLY-AUX and a wild type were cultivated in 10 mL of M9 + 20 mM glycerol + 1.5 mM glyoxylate or M9 + 20 mM glycerol + 5 mM glycine at 37 °C shaking at 180 rpm. Cells from three biological replicates were harvested at an equivalent of 1 mL of OD 3 in the late exponential phase by centrifugation for 1 min at 11,000 \times g. The cells were washed two times with phosphate buffer (12 mM phosphate buffer, 2.7 mM KCl, 137 mM NaCl, pH = 7.4) before being flash-frozen in liquid nitrogen for subsequent storage at -70 °C. To isolate the proteome, cell pellets were resuspended in 2% sodium lauroyl sarcosinate (SLS, in 100 mM ammonium bicarbonate) and heat incubated at 95 °C for 15 min. For DNA shearing, the cells were then sonicated for 30 s (Vial Tweeter, Hielscher). After quantifying the total protein in a BCA assay, the samples were treated with 5 mM Tris (2-carboxy-ethyl) phosphine (TCEP) at 90 °C for 15 min, followed by protein alkylation using 10 mM Iodoacetamide for 30 min in the dark at 25 °C. For protein digestion, 50 μ g of total protein was incubated with 1 μ g of porcine trypsin (Promega) in the presence of 0.5% SLS at 30 °C overnight. After digestion, SLS was precipitated by adding a final concentration of 1.5% trifluoroacetic acid (TFA, Thermo Fischer Scientific). Peptides were desalted by using C18 solid phase extraction cartridges (Macherey-Nagel). Cartridges were prepared by adding acetonitrile (ACN), followed by equilibration with 0.1% TFA. Peptides were loaded on equilibrated cartridges, washed with 5% ACN and 0.1% TFA containing buffer, eluted with 50% ACN and 0.1% TFA, and finally dried. The peptides were resuspended in 100 μ L 0.1% TFA and the peptide mixtures were then analyzed by LC-MS on an Exploris 480 instrument connected to an Ultimate 3000 RSLC nano and a nanospray flex ion source (Thermo Scientific). A capillary column (75 μ m \times 42 cm) packed in-house with C18 resin (2.4 μ m, Dr. Maisch) was used for peptide separation. The following separating gradients were used: 94% solvent A (0.15% formic acid) and 6% solvent B (99.85% acetonitrile, 0.15% formic acid) to 35% solvent B over 60 min at a flow rate of 300 nL/min. DIA-MS acquisition method was performed with: spray voltage set to 2.3 kV, funnel RF level at 45, and heated capillary temperature at 275 °C. For DIA experiments MS1 resolution was set to 120,000 at m/z 200 and the full MS AGC target was 300% with max. injection time (IT) of 50 ms. The mass range was set to 350–1400. AGC target value for fragment spectra was set at 3000%. 49 windows of 15 Da were used with an overlap of 1 Da. Resolution was set to 15,000 and IT to 22 ms. Stepped HCD collision energy of 25%, 27.5%, and 30% was used. MS1 data was acquired in profile, and MS2 DIA data in centroid mode.

Analysis of DIA data was performed using DIA-NN version 1.8⁸², using the UniProt protein database from *Escherichia coli* K12, and added sequences for the SJ488 λ -red recombineering machinery (Phage recombinase gam (UniProt ID NP_040618.1), Phage recombinase beta (UniProt ID WP_000100844), Phage recombinase exo (UniProt ID IAVQ_A), Flippase (UniProt ID P03870.1)). Full tryptic digest was allowed with three missed cleavage sites, and oxidized methionines and carbamidomethylated cysteines. Match between runs and remove likely interferences were enabled. The neural network classifier was set to the single-pass mode, and protein inference was based on genes. The quantification strategy was set to any LC (high accuracy).

Cross-run normalization was set to RT-dependent. Library generation was set to smart profiling. DIA-NN outputs were further evaluated using a SafeQuant version modified to process DIA-NN outputs⁸³.

Determining extracellular glycolate in spent medium using LC-MS/MS

Chromatographic metabolite separation was performed on an Agilent Infinity II 1290 HPLC system using a Kinetex EVO C18 column (150 × 2.1 mm, 3 μm particle size, 100 Å pore size, Phenomenex) connected to a guard column of similar specificity (20 × 2.1 mm, 3 μm particle size, Phenomenex) at a constant flow rate of 0.1 mL/min with 2 μL injection volume and 0.1% formic acid in water as mobile phase A and 0.1% formic acid in methanol (Honeywell, Morristown, New Jersey, USA) as phase B at 25 °C. The mobile phase profile consisted of the following steps and linear gradients: 0–4 min constant at 0% B; 4–6 min from 0% to 100% B; 6–7 min constant at 100% B; 7–7.1 min from 100% to 0% B; 7.1–12 min constant at 0% B. An Agilent 6495 mass spectrometer was used in negative mode with an electrospray ionization source and the following conditions: ESI spray voltage 2000 V, nozzle voltage 500 V, sheath gas 300 °C at 11 L/min, nebulizer pressure 50 psig and drying gas 80 °C at 16 L/min.

Compounds were identified and quantified based on their mass transition, retention time, and peak area compared to an external standard curve using the MassHunter software (Agilent, Santa Clara, CA, USA). The parameters mass transitions, collision energies, Cell accelerator voltages, and Dwell times have been optimized using chemically pure standards and are given in Supplementary Table 4.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. The mass spectrometry proteomics data (PXD054993) is accessible through the link <https://www.ebi.ac.uk/pride/archive/projects/PXD054993>. A reporting summary for this Article is available as a Supplementary Information file. Source data are provided in this paper. All strains presented in the manuscript can be obtained for academic research from the corresponding author upon request. Source data are provided in this paper.

Code availability

All scripts, data files, and result plots can be found on our GitLab repository: <https://gitlab.com/elad.noor/glyoxylate-auxotrophy>.

References

- Downs, D. M. Understanding microbial metabolism. *Annu Rev. Microbiol.* **60**, 533–559 (2006).
- Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
- Wenk, S., Yishai, O., Lindner, S. N. & Bar-Even, A. An engineering approach for rewiring microbial metabolism. *Methods Enzymol.* **608**, 329–367 (2018).
- Orsi, E., Claassens, N. J., Nikel, P. I. & Lindner, S. N. Growth-coupled selection of synthetic modules to accelerate cell factory development. *Nat. Commun.* **12**, 5295 (2021).
- Schulz-Mirbach, H. et al. On the flexibility of the cellular amination network in *E. coli*. *Elife*. <https://doi.org/10.7554/eLife.77492> (2022).
- Gleizer, S. et al. Conversion of *Escherichia coli* to generate all biomass carbon from CO₂. *Cell* **179**, 1255–1263 (2019).
- Chen, F. Y. H., Jung, H. W., Tsuei, C. Y. & Liao, J. C. Converting *Escherichia coli* to a synthetic methylotroph growing solely on methanol. *Cell* **182**, 933–946 (2020).
- Kim, S. et al. Growth of *E. coli* on formate and methanol via the reductive glycine pathway. *Nat. Chem. Biol.* **16**, 538–545 (2020).
- Wenk, S. et al. Evolution-assisted engineering of *E. coli* enables growth on formic acid at ambient CO₂ via the Serine Threonine Cycle. *Metab Eng.* **88**, 14–24 (2025).
- Cros, A., Alfaro-Espinoza, G., De Maria, A., Wirth, N. T. & Nikel, P. I. Synthetic metabolism for biohalogenation. *Curr. Opin. Biotechnol.* **74**, 180–193 (2022).
- Satanowski, A. et al. Awakening a latent carbon fixation cycle in *Escherichia coli*. *Nat. Commun.* **11**, 5812 (2020).
- Antonovsky, N. et al. Sugar synthesis from CO₂ in *Escherichia coli*. *Cell* **166**, 115–125 (2016).
- Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. *Physiology of the Bacterial Cell: a Molecular Approach* (Sinauer Associates, 1990).
- Corrao, M., He, H., Liebermeister, W., Noor, E. & Bar-Even, A. A compact model of *Escherichia coli* core and biosynthetic metabolism. Preprint at <https://doi.org/10.48550/arXiv.2406.16596> (2024).
- Yadav, M., Pulletikurti, S., Yerabolu, J. R. & Krishnamurthy, R. Cyanide as a primordial reductant enables a protometabolic reductive glyoxylate pathway. *Nat. Chem.* **14**, 170–178 (2022).
- Krishnamurthy, R. & Liotta, C. L. The potential of glyoxylate as a prebiotic source molecule and a reactant in protometabolic pathways—the glyoxylose reaction. *Chem* **9**, 784–797 (2023).
- Pulletikurti, S., Yadav, M., Springsteen, G. & Krishnamurthy, R. Prebiotic synthesis of α-amino acids and orotate from α-ketoacids potentiates transition to extant metabolic pathways. *Nat. Chem.* **14**, 1142–1150 (2022).
- Springsteen, G., Yerabolu, J. R., Nelson, J., Rhea, C. J. & Krishnamurthy, R. Linked cycles of oxidative decarboxylation of glyoxylate as protometabolic analogs of the citric acid cycle. *Nat. Commun.* <https://doi.org/10.1038/s41467-017-02591-0> (2018).
- Schwander, T., Schada Von Borzyskowski, L., Burgener, S., Cortina, N. S. & Erb, T. J. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **354**, 900–904 (2016).
- Ren, J. et al. An unnatural pathway for efficient 5-aminolevulinic acid biosynthesis with glycine from glyoxylate based on retro-biosynthetic design. *ACS Synth. Biol.* **7**, 2750–2757 (2018).
- McLean, R. et al. Exploring alternative pathways for the in vitro establishment of the HOPAC cycle for synthetic CO₂ fixation. *Sci. Adv.* <https://doi.org/10.1126/sciadv.adh4299> (2023).
- Mainguet, S. E., Gronenberg, L. S., Wong, S. S. & Liao, J. C. A reverse glyoxylate shunt to build a non-native route from C4 to C2 in *Escherichia coli*. *Metab. Eng.* **19**, 116–127 (2013).
- Vuoristo, K. S., Sanders, J. P. M., Weusthuis, R. A., Mars, A. E. & Eggink, G. Metabolic engineering of TCA cycle for production of chemicals. *Trends Biotechnol.* **34**, 191–197 (2015).
- Yang, P., Liu, W., Chen, Y. & Gong, A. D. Engineering the glyoxylate cycle for chemical bioproduction. *Front. Bioeng. Biotechnol.* **10**, 1066651 (2022).
- Aslan, S., Noor, E., Benito Vaquerizo, S., Lindner, S. N. & Bar-Even, A. Design and engineering of *E. coli* metabolic sensor strains with a wide sensitivity range for glycerate. *Metab. Eng.* **57**, 96–109 (2020).
- Schada Von Borzyskowski, L. et al. Marine proteobacteria metabolize glycolate via the β-hydroxyaspartate cycle. *Nature* **575**, 500–504 (2019).
- Long, C. P. & Antoniewicz, M. R. Metabolic flux responses to deletion of 20 core enzymes reveal flexibility and limits of *E. coli* metabolism. *Metab. Eng.* **55**, 249–257 (2019).
- Iacometti, C. et al. Activating silent glycolysis bypasses in *Escherichia coli*. *BioDesign Res.* **2022**, 1–17 (2022).
- Claassens, N. J. et al. Replacing the Calvin cycle with the reductive glycine pathway in *Cupriavidus necator*. *Metab. Eng.* **62**, 30–41 (2020).

30. Gerike, U., Hough, D. W., Russell, N. J., Dyall-Smith, M. L. & Danson, M. J. Citrate synthase and 2-methylcitrate synthase: structural, functional and evolutionary relationships. *Microbiology* **144**, 929–935 (1998).
31. Guzmán, G. I. et al. Model-driven discovery of underground metabolic functions in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **112**, 929–934 (2015).
32. Liu, J. Q. et al. Gene cloning, biochemical characterization and physiological role of a thermostable low-specificity L-threonine aldolase from *Escherichia coli*. *Eur. J. Biochem.* **255**, 220–226 (1998).
33. Yishai, O., Goldbach, L., Tenenboim, H., Lindner, S. N. & Bar-Even, A. Engineered assimilation of exogenous and endogenous formate in *Escherichia coli*. *ACS Synth. Biol.* **6**, 1722–1731 (2017).
34. Cusa, E., Obradors, N., Baldomà, L., Badía, J. & Aguilar, J. Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in *Escherichia coli*. *J. Bacteriol.* **181**, 7479–7484 (1999).
35. Rintoul, M. R. et al. Regulation of the *Escherichia coli* allantoin regulon: coordinated function of the repressor AllR and the activator AllS. *J. Mol. Biol.* **324**, 599–610 (2002).
36. Yishai, O., Bouzon, M., Döring, V. & Bar-Even, A. In vivo assimilation of one-carbon via a synthetic reductive glycine pathway in *Escherichia coli*. *ACS Synth. Biol.* **7**, 2023–2028 (2018).
37. Wenk, S. et al. An “energy-auxotroph” *Escherichia coli* provides an in vivo platform for assessing NADH regeneration systems. *Biotechnol. Bioeng.* **117**, 3422–3434 (2020).
38. Yu, H. & Liao, J. C. A modified serine cycle in *Escherichia coli* converts methanol and CO₂ to two-carbon compounds. *Nat. Commun.* <https://doi.org/10.1038/s41467-018-06496-4> (2018).
39. Yu, H., Li, X., Duchoud, F., Chuang, D. S. & Liao, J. C. Augmenting the Calvin–Benson–Bassham cycle by a synthetic malyl-CoA-glycerate carbon fixation pathway. *Nat. Commun.* <https://doi.org/10.1038/s41467-018-04417-z> (2018).
40. Lau, W. W. Y., Keil, R. G. & Armbrust, E. V. Succession and diel transcriptional response of the glycolate-utilizing component of the bacterial community during a spring phytoplankton bloom. *Appl. Environ. Microbiol.* **73**, 2440–2450 (2007).
41. Bauwe, H., Hagemann, M. & Fernie, A. R. Photorespiration: players, partners and origin. *Trends Plant Sci.* **15**, 330–336 (2010).
42. Dellero, Y., Jossier, M., Schmitz, J., Maurino, V. G. & Hodges, M. Photorespiratory glycolate-glyoxylate metabolism. *J. Exp. Bot.* **67**, 3041–3052 (2016).
43. Zhu, X.-G., Long, S. P. & Ort, D. R. Improving photosynthetic efficiency for greater yield. *Annu. Rev. Plant Biol.* **61**, 235–261 (2010).
44. Federici, F., Orsi, E. & Nikel, P. I. From rags to riches: exploiting the calvin-benson-bassham cycle for biomanufacturing. *Chem-CatChem.* <https://doi.org/10.1002/cctc.202300746> (2023).
45. Taubert, A., Jakob, T. & Wilhelm, C. Glycolate from microalgae: an efficient carbon source for biotechnological applications. *Plant Biotechnol. J.* **17**, 1538–1546 (2019).
46. Yang, F., Zhang, J., Cai, Z., Zhou, J. & Li, Y. Exploring the oxygenase function of form II Rubisco for production of glycolate from CO₂. *AMB Express* **11**, 1–15 (2021).
47. Kaplan, A. & Reinhold, L. CO₂ concentrating mechanisms in photosynthetic microorganisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 539–570 (1999).
48. Keller, P. et al. Methanol-dependent *Escherichia coli* strains with a complete ribulose monophosphate cycle. *Nat. Commun.* <https://doi.org/10.1038/s41467-020-19235-5> (2020).
49. Keller, P. et al. Generation of an *Escherichia coli* strain growing on methanol via the ribulose monophosphate cycle. *Nat. Commun.* <https://doi.org/10.1038/s41467-022-32744-9> (2022).
50. Orth, J. D., Fleming, R. M. T. & Palsson, B. Ø. Reconstruction and use of microbial metabolic networks: the core *Escherichia coli* metabolic model as an educational guide. *EcoSal Plus.* <https://doi.org/10.1128/ecosalplus.10.2.1> (2010).
51. Monk, J. M. et al. Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional environments. *Proc. Natl. Acad. Sci. USA* **110**, 20338–20343 (2013).
52. Karp, P. D. et al. The EcoCyc database. *EcoSal Plus* <https://doi.org/10.1128/ecosalplus.esp-0002-2023> (2023).
53. Schulz-Mirbach, H., Dronsella, B., He, H. & Erb, T. J. Creating new-to-nature carbon fixation: a guide. *Metab. Eng.* **82**, 12–28 (2024).
54. Orsi, E., Claassens, N. J., Nikel, P. I. & Lindner, S. N. Optimizing microbial networks through metabolic bypasses. *Biotechnol. Adv.* **60**, 108035 (2022).
55. Pflieger, B. F., Pitera, D. J., Newman, J. D., Martin, V. J. J. & Keasling, J. D. Microbial sensors for small molecules: development of a mevalonate biosensor. *Metab. Eng.* **9**, 30–38 (2007).
56. Hernández-Sancho, J. M. et al. A versatile microbial platform as a tunable whole-cell chemical sensor. *Nat. Commun.* <https://doi.org/10.1038/s41467-024-52755-y> (2024).
57. Xu, S., Zhang, L., Zhou, S. & Deng, Y. Biosensor-based multigene pathway optimization for enhancing the production of glycolate. *Appl. Environ. Microbiol.* <https://doi.org/10.1128/aem.00113-21> (2021).
58. Barthel, S. et al. *In Vitro Transcription-Based Biosensing of Glycolate for Prototyping of a Complex Enzyme Cascade* (Cold Spring Harbor Laboratory, 2024).
59. Tsiafoulis, C. G., Prodromidis, M. I. & Karayannis, M. I. Development of amperometric biosensors for the determination of glycolic acid in real samples. *Anal. Chem.* **74**, 132–139 (2002).
60. Chen, K. & Arnold, F. H. Engineering new catalytic activities in enzymes. *Nat. Catal.* **3**, 203–213 (2020).
61. Wang, Y. et al. Directed evolution: methodologies and applications. *Chem. Rev.* **121**, 12384–12444 (2021).
62. McLure, R. J., Radford, S. E. & Brockwell, D. J. High-throughput directed evolution: a golden era for protein science. *Trends Chem.* **4**, 278–291 (2022).
63. Molina, R. S. et al. *In vivo* hypermutation and continuous evolution. *Nat. Rev. Methods Primers.* <https://doi.org/10.1038/s43586-022-00130-w> (2022).
64. Yu, T., Boob, A. G., Singh, N., Su, Y. & Zhao, H. In vitro continuous protein evolution empowered by machine learning and automation. *Cell Syst.* **14**, 633–644 (2023).
65. Li, Z., Deng, Y. & Yang, G. Y. Growth-coupled high throughput selection for directed enzyme evolution. *Biotechnol. Adv.* **68**, 108238 (2023).
66. Orsi, E., Schada Von Borzyskowski, L., Noack, S., Nikel, P. I. & Lindner, S. N. Automated in vivo enzyme engineering accelerates biocatalyst optimization. *Nat. Commun.* <https://doi.org/10.1038/s41467-024-46574-4> (2024).
67. Walker, B. J., Vanlocke, A., Bernacchi, C. J. & Ort, D. R. The costs of photorespiration to food production now and in the future. *Annu. Rev. Plant Biol.* **67**, 107–129 (2016).
68. Garcia, A. et al. Enhancing crop yields through improvements in the efficiency of photosynthesis and respiration. *N. Phytol.* **237**, 60–77 (2023).
69. Betti, M. et al. Manipulating photorespiration to increase plant productivity: recent advances and perspectives for crop improvement. *J. Exp. Bot.* **67**, 2977–2988 (2016).
70. Behrenfeld, M. J. & Boss, E. S. Resurrecting the ecological underpinnings of ocean plankton blooms. *Annu. Rev. Mar. Sci.* **6**, 167–194 (2014).
71. Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRApy: constraints-based reconstruction and analysis for python. *BMC Syst. Biol.* **7**, 74 (2013).
72. Jensen, S. I., Lennen, R. M., Hergård, M. J. & Nielsen, A. T. Seven gene deletions in seven days: fast generation of *Escherichia coli*

- strains tolerant to acetate and osmotic stress. *Sci. Rep.* <https://doi.org/10.1038/srep17874> (2015).
73. Volke, D. C., Martino, R. A., Kozaeva, E., Smania, A. M. & Nickel, P. I. Modular (de)construction of complex bacterial phenotypes by CRISPR/nCas9-assisted, multiplex cytidine base-editing. *Nat. Commun.* <https://doi.org/10.1038/s41467-022-30780-z> (2022).
 74. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006 0008 (2006).
 75. Blin, K., Pedersen, L. E., Weber, T. & Lee, S. Y. CRISPy-web: an online resource to design sgRNAs for CRISPR applications. *Synth. Syst. Biotechnol.* **1**, 118–121 (2016).
 76. Thomason, L. C., Costantino, N. & Court, D. L. *E. coli* genome manipulation by P1 transduction. *Curr. Protoc. Mol. Biol.* **79**, 1.17.11–11.17.18 (2007).
 77. Zelcbuch, L. et al. Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.* **41**, e98–e98 (2013).
 78. Donati, S. et al. An automated workflow for multi-omics screening of microbial model organisms. *npj Syst. Biol. Appl.* <https://doi.org/10.1038/s41540-023-00277-6> (2023).
 79. Long, C. P. & Antoniewicz, M. R. High-resolution ¹³C metabolic flux analysis. *Nat. Protoc.* **14**, 2856–2877 (2019).
 80. Kutuzova, S. et al. SmartPeak automates targeted and quantitative metabolomics data processing. *Anal. Chem.* **92**, 15968–15974 (2020).
 81. Wahl, S. A., Dauner, M. & Wiechert, W. New tools for mass isotopomer data evaluation in ¹³C flux analysis: mass isotope correction, data consistency checking, and precursor relationships. *Biotechnol. Bioeng.* **85**, 259–268 (2004).
 82. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
 83. Ahrné, E., Molzahn, L., Glatter, T. & Schmidt, A. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *PROTEOMICS* **13**, 2567–2578 (2013).
- identified additional deletion targets with input from S.N.L. to A.B.E. E.O., H.S.M., C.C., A.S., H.P., S.A., N.G., R.V., F.S., and A.M. constructed the strains used in this study. E.O., H.S.M., and C.C. characterized the sensor strains. H.S.M. generated data points for the quantitative sensitivity comparisons. E.O. and H.S.M. assembled results from all authors and wrote the original manuscript draft with help from E.N., H.H., and P.I.N. E.O. generated labeling samples, which were processed and analyzed by K.S.J. and S.D. N.P. quantified metabolites and analyzed the data. A.K. cultivated *S. elongatus* strains and T.C. cultivated *C. reinhardtii* strains. T.G. analyzed proteomic samples. E.O., A.B.E., T.E., and P.I.N. acquired funding. A.B.E., T.E., and P.I.N. supervised the research.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-57407-3>.

Correspondence and requests for materials should be addressed to Enrico Orsi or Pablo I. Nickel.

Peer review information *Nature Communications* thanks Aditya Kunjapur and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Acknowledgements

The authors thank Jörg Kahnt for technical assistance with proteomics and Beau Dronsella for valuable feedback on the strain designs. This work was supported by The Novo Nordisk Foundation grant NNF20CC0035580. E.O. acknowledges support from the European Union through the Marie Skłodowska-Curie grant agreement 101065339. H.S.-M. acknowledges support from the Bosch Research Foundation and the Joachim Herz Foundation. The financial support from The Novo Nordisk Foundation through grants NNF10CC1016517, NNF18CC0033664, and NNF23OC0083631 to P.I.N. is likewise gratefully acknowledged.

Author contributions

E.O. and H.S.M. conceptualized the study. E.N. and H.H. predicted knockout combinations and selection demands in silico. C.C. and E.O.