Synthetic Biology, 2024, **9(1)**, ysae013
DOI: https://doi.org/10.1093/synbio/ysae013
Advance Access Publication 20 September 2024
Research Article

In vitro transcription-based biosensing of glycolate for prototyping of a complex enzyme cascade

Sebastian Barthel[1,*,‡], Luca Brenker[1,‡], Christoph Diehl[1], Nitin Bohra[1,2], Simone Giaveri[1], Nicole Paczia[3], Tobias J. Erb[1,4,*]

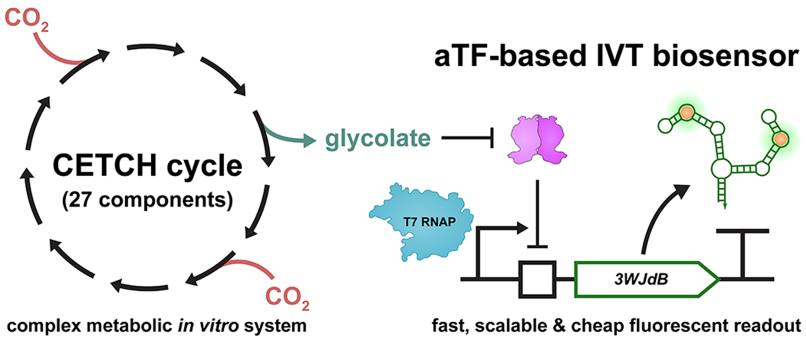
[1]Department of Biochemistry & Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Marburg 35043, Germany
[2]Max Planck School Matter to Life, Heidelberg 69120, Germany
[3]Core Facility for Metabolomics and Small Molecule Mass Spectrometry, Max Planck Institute for Terrestrial Microbiology, Marburg 35043, Germany
[4]Center for Synthetic Microbiology (SYNMIKRO), Philipps University Marburg, Marburg 35043, Germany
[‡]These authors contributed equally to this work.
*Corresponding authors. Tobias J. Erb, Department of Biochemistry & Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, Marburg 35043, Germany. E-mail: toerb@mpi-marburg.mpg.de; Sebastian Barthel, Department of Biochemistry & Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, Marburg 35043, Germany. E-mail: sebastian.barthel@mpi-marburg.mpg.de.

Abstract

In vitro metabolic systems allow the reconstitution of natural and new-to-nature pathways outside of their cellular context and are of increasing interest in bottom-up synthetic biology, cell-free manufacturing, and metabolic engineering. Yet, the analysis of the activity of such *in vitro* networks is very often restricted by time- and cost-intensive methods. To overcome these limitations, we sought to develop an *in vitro* transcription (IVT)-based biosensing workflow that is compatible with the complex conditions of *in vitro* metabolism, such as the crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle, a 27-component *in vitro* metabolic system that converts CO_2 into glycolate. As proof of concept, we constructed a novel glycolate sensor module that is based on the transcriptional repressor GlcR from *Paracoccus denitrificans* and established an IVT biosensing workflow that allows us to quantify glycolate from CETCH samples in the micromolar to millimolar range. We investigate the influence of 13 (shared) cofactors between the two *in vitro* systems to show that Mg^{2+}, adenosine triphosphate, and other phosphorylated metabolites are critical for robust signal output. Our optimized IVT biosensor correlates well with liquid chromatography–mass spectrometry–based glycolate quantification of CETCH samples, with one or multiple components varying (linear correlation 0.94–0.98), but notably at ∼10-fold lowered cost and ∼10 times faster turnover time. Our results demonstrate the potential and challenges of IVT-based systems to quantify and prototype the activity of complex reaction cascades and *in vitro* metabolic networks.

Keywords: *in vitro* transcription; allosteric transcription factors; biosensing; pathway prototyping; CETCH cycle; GlcR; screening method; *in vitro* metabolic system

Graphical Abstract



Submitted: 18 September 2024; **Received**: 27 August 2024; **Accepted**: 18 September 2024

Introduction

Synthetic biochemistry aims to reconstruct biological functions outside of a living cell ("cell-free"). Prominent examples are efforts to reconstitute natural (or new-to-nature) pathways from purified enzymes *in vitro* [1–9]. Such approaches not only allow us to study the fundamental design principles and function of metabolic networks [2–4] but also bear great application potential. For example, recent work demonstrated the cell-free conversion of the greenhouse gas carbon dioxide (CO_2) into polyketides, terpenes, and antibiotic precursors [5, 6] or the valorization of glucose and other low-cost precursors into monoterpenes and cannabinoids [7–9].

Compared to *in vivo* systems, cell-free metabolic networks are highly flexible in their composition, can be precisely modified and customized, and allow biochemical reactions to take place under "non-physiological" conditions [10]. Through cell-free systems, a rapid optimization of reaction compositions is possible without the need for molecular cloning, which minimizes time and cost. Consequently, lysate-based cell-free systems have been increasingly used to prototype pathways for the optimal combination and stoichiometry of individual enzymes and components [11–15]. In several cases, these optimized *in vitro* pathways could also be successfully implemented *in vivo*. Altogether, this showcases the capabilities of cell-free systems as a broad tool for *in vitro* and *in vivo* metabolic engineering purposes.

To optimize complex biological systems with minimal experimental effort, Pandi *et al.* recently reported a versatile workflow Machine-learning guided Experimental Trials for Improvement of Systems (METIS) that combines laboratory automation with active learning to explore the combinatorial space in iterative design–build–test cycles for (local) optima [16]. METIS successfully helped to improve several biological systems [3, 4, 17, 18], including an *in vitro* CO_2 fixation cycle of 27 different variables (crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle) [2, 19]. The CETCH cycle converts CO_2 into glycolate and could be improved by >10-fold through METIS. Although active learning–guided workflows are able to drastically minimize the number of samples screened, the screening phase still heavily relies on the use of costly and time-intensive instrumental analytics. In the case of the CETCH cycle, >3000 samples were analyzed by liquid chromatography–mass spectrometry (LC–MS), which requires 12 min per sample for glycolate quantification at a cost of ∼US$7 (Supplementary Note S1). We therefore set out to explore a low-cost and well-scalable *in vitro* transcription (IVT)-based biosensing method that would allow us to quantify activity of the *in vitro* metabolic system in the reaction mixture and thus increase throughput.

To establish such an IVT biosensing method, we turned our attention to a system called RNA Output Sensors Activated by Ligand Induction (ROSALIND), which was recently developed to detect pollutants in water samples [20]. The ROSALIND system consists of a linear DNA template encoding the sequence of an RNA aptamer ["Three-way Junction dimeric Broccoli" (*3WJdB*)] [21] (Fig. 1a). The *3WJdB* aptamer is expressed under the control of a T7 promoter and an operator sequence that is repressed by an allosteric transcription factor (aTF). Only in the presence of its specific effector, the aTF releases the operator sequence to allow *3WJdB* expression, which in turn results in a green fluorescent readout by stabilizing the fluorogenic dye DFHBI-1T in its fluorescent state.

Here, we demonstrate a ROSALIND-based biosensing workflow for the rapid prototyping and screening of complex *in vitro* metabolic systems, using the CETCH cycle as the proof of principle. We developed a glycolate-responsive sensor module to read out the glycolate-forming activity of the CETCH cycle and investigated the inhibitory effects of CETCH cycle components on the IVT system. We identified the availability of free Mg^{2+} as a critical factor for establishing highly robust and quantitative sensing of glycolate production. Notably, the IVT-based biosensing workflow reduces screening costs by an order of magnitude and reduces the analysis time of large sample sets from several days to ∼8 h. This work not only demonstrated that IVT-based reporter systems can be coupled to complex *in vitro* metabolic systems but also identified critical components and bottlenecks in setting up robust IVT-based screens under such complex and challenging conditions. Our work paves the way for the development of similar IVT-based reporter systems and further guides ongoing efforts to integrate *in vitro* metabolic networks and IVT–translation systems [22] toward constructing a synthetic cell [23, 24].

Materials and methods

Reagents

Unless stated differently, chemicals were purchased from Merck KGaA (Darmstadt, Germany) and Carl Roth GmbH (Karlsruhe, Germany). Commercial enzymes and bioreagents were purchased from New England Biolabs (Frankfurt am Main, Germany).

Strains and growth media

For molecular cloning, *Escherichia coli* NEB® turbo was grown in lysogeny broth supplemented with an appropriate antibiotic (100 μg/ml ampicillin or 34 μg/ml chloramphenicol). For protein production, either *E. coli* M15 [T7 RNA polymerase (RNAP)] or *E. coli* BL21-AI [MBP-GlcR (MGlcR)] was grown in terrific broth (TB) supplemented with 100 μg/ml ampicillin or 34 μg/ml chloramphenicol, respectively. All strains used are listed in Supplementary Table S1.

Assembly of plasmids and preparation of linear DNA templates

Oligonucleotides were purchased from Merck KGaA. Synthetic dsDNA was purchased from Twist Bioscience (South San Francisco, CA, USA). Sanger sequencing was performed by MicroSynth (Göttingen, Germany).

Plasmids were generated by Golden Gate Assembly using the modular cloning system proposed by Stukenberg *et al.* [25]. A mixture of 0.4 nM vector DNA and 4–8 nM insert DNA was assembled using 0.5 U/μl Esp3I or 1 U/μl BsaI-HFv2 and 40 U/μl T4 ligase in 1× T4 ligase buffer. Reactions were cycled 15 times for 1.5 min at 37°C and 3 min at 16°C. Enzymes were heat inactivated for 5 min at 50°C and 10 min at 80°C. The Golden Gate product was transformed into chemically competent *E. coli* NEB turbo cells, and individual clones were verified by Sanger sequencing using oligonucleotides oSB0021 and oSB0022.

All linear DNA templates were prepared by polymerase chain reaction (PCR) amplification from the respective plasmids using oligonucleotides oSB0021 and oSB0022 and Q5 DNA polymerase, following the vendor's instructions. All amplified DNA fragments were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), according to the vendor's instructions. DNA concentrations were calculated from absorbance measurements at 260 nm (A_{260}) using a NanoDrop2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). To increase the throughput of screening various T7 promoter–operator sequences, we

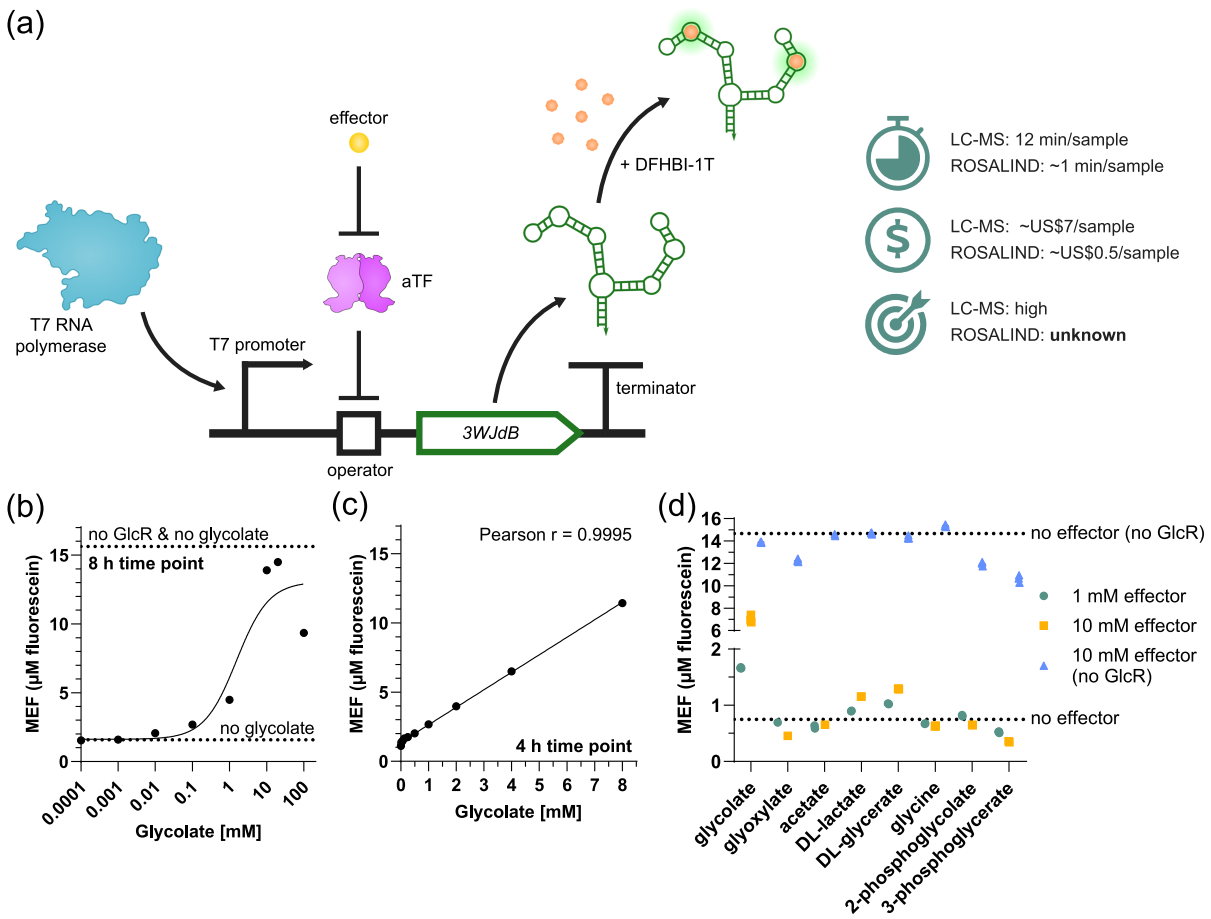


Figure 1. Characterization of a glycolate-responsive IVT-based biosensor module. (a) The ROSALIND system is based on the controlled expression of the *3WJdB* RNA aptamer and the correlating fluorescence signal of the *3WJdB*:DFHBI-1T complex. Expression is regulated by a transcriptional repressor that binds to an operator sequence downstream of a T7 promoter. The repression is lifted in a dose-responsive manner by binding an effector molecule to the aTF. This system allows faster and cheaper sample measurement in microtiter plates than analysis by LC–MS, but its precision is yet unknown. Time estimates refer to glycolate quantification. (b, c) Dose–response of the GlcR sensor module to glycolate in the concentration range of 100 nM to 100 mM glycolate (8 h time point, further time points shown in Supplementary Fig. S5e). The sensor shows an operational range from 10 µM to 20 mM glycolate with an excellent correlation between 16 µM and 8 mM glycolate and a 10-fold dynamic range. Note that IVT biosensing responses are highly time sensitive with better correlation at early time points and better sensitivity at late time points (Supplementary Fig. S5; Table S8). (d) Promiscuity assay of GlcR shows a low-level dose response to DL-lactate and DL-glycerate (4 h time point, further data shown in Supplementary Fig. S6). Raw fluorescence data are standardized to MEF (µM fluorescein). Data are the mean of 3 technical replicates ± SD. IVT output without the effector molecule and without MGlcR is shown as horizontal dotted lines.

developed a workflow (inspired by previous work [26, 27]), which is described in detail in Supplementary Note S2 and the Supplementary Methods. All plasmids, linear DNA templates, and oligonucleotides used are listed in Supplementary Tables S2, S3, and S4, respectively.

Protein production and purification

MGlcR was produced and purified as previously described by Schada von Borzyskowski *et al.* [28]. T7 RNAP was produced in an *E. coli* M15 strain harboring plasmid pQE30-T7 RNAP [29] (strain sAP94). First, a preculture was inoculated in TB, supplemented with 100 µg/ml ampicillin. Cells were grown to high density overnight at 37°C. The preculture was used the following day to inoculate a production culture in TB medium supplemented with 100 µg/ml ampicillin and antifoam reagent. The culture was grown in a baffled flask at 37°C until an optical density (OD_{600}) of 0.7 was reached. The culture was then cooled down to room temperature for 30 min, before inducing the culture with 0.5 mM isopropyl β-D-1-thiogalactopyranoside. Cells were grown overnight at 20°C. Cells were harvested at 4000 × *g* for 20 min at 12°C, and cell pellets were resuspended in twice their volume of Buffer A (50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 500 mM KCl) with 5 mM $MgCl_2$ and DNase I (Roche, Basel, Switzerland). Cells were lysed by sonication using a SonoplusGM200 (BANDELIN electronic GmbH & Co. KG, Berlin, Germany) equipped with a KE76 tip at 50% amplitude for 3× 1 min of 1-s on/off pulses. Lysates were cleared by centrifugation at 100 000 × *g* for 1 h at 8°C, and the supernatant was then filtered through 0.45 µm filters (Sarstedt, Nümbrecht, Germany). For affinity purification, an Äkta Start FPLC system (formerly GE Healthcare, now Cytiva, Marlborough, MA, USA) was used with two stacked 1-ml Ni-NTA columns (HiTrap HP, Cytiva). The cleared lysate was loaded onto the columns, which were equilibrated with Buffer A. The column was washed with Buffer A + 75 mM imidazole and eluted with Buffer A + 500 mM imidazole. The eluate was desalted using two stacked 5-ml HiTrap desalting columns (Sephadex G-25 resin, Cytiva) and protein elution buffer (25 mM Tris-HCl pH 7.4, 100 mM NaCl). Protein concentration was calculated from absorbance at

280 nm (A_{280}) on a NanoDrop2000, and respective extinction coefficients were calculated using ProtParam (https://web.expasy.org/protparam/). Purified T7 RNAP was aliquoted, flash-frozen in liquid nitrogen, and stored at −70°C. See Supplementary Table S9 for an overview of which batch of T7 RNAP was used in which experiment.

IVT assays

IVT reactions were typically set up, unless stated differently, by adding the following components at their final concentration: IVT buffer [40 mM HEPES (pH 7.8), 8 mM $MgCl_2$, 10 mM dithiothreitol, 20 mM NaCl, and 2 mM spermidine], 0.2 mM DFHBI-1T and 11.4 mM nucleoside triphosphates (rNTPs, Thermo Fisher Scientific), 0.015 U/μl thermostable inorganic pyrophosphatase (New England Biolabs), and 15 nM DNA template. MGlcR (750 nM) was added to the GlcR module (see Supplementary Fig. S12 for batch-to-batch variations). To ensure stability of rNTPs, stocks at 80 mM rNTPs are buffered in a 200 mM Tris base. The sample volume of an IVT reaction was 20 μl and prepared in replicates of $n = 3$. The reaction mixture (57.8 μl = 3.4 × 17 μl) was equilibrated at room temperature (RT) for 30 min, before adding the respective analyte in a 1:10 dilution (6.8 μl = 3.4 × 2 μl, e.g. effector molecule, individual CETCH components at an indicated concentration, or CETCH cycle sample) and >10 U/μl T7 RNAP (3.4 μl = 3.4 × 1 μl, see Supplementary Fig. S12 for batch-to-batch variations; activity reference is NEB T7 RNAP, #M0251S). The reaction mixtures were mixed by pipetting, and 3 × 20 μl was immediately transferred into a 384-well, black, optically clear, flat-bottom, nonbinding microtiter plate (Greiner Bio-One, Kremsmünster, Austria; catalog no.: 781906). Plates were centrifuged for 30 s in a small tabletop plate centrifuge (VWR, Radnor, PA, USA) before measurement. Reactions were characterized in triplicates on a plate reader (Infinite M200, Tecan, Männedorf, Switzerland) at 37°C, with 30 s of shaking before each fluorescence read at an excitation wavelength of 472 nm and an emission wavelength of 507 nm. Bottom reads of the plate allow for more precise measurements compared to top reads. To convert arbitrary fluorescence measurements to micromolar equivalents of fluorescein (MEFs), serial dilutions of a 12.5-μM stock of fluorescein standard (Invitrogen, catalog no.: F36915) were prepared in dH_2O and measured alongside each IVT assay. To convert arbitrary fluorescence units in MEF, (i) fluorescein fluorescence (in arbitrary units) was linearly regressed with fluorescein concentrations (in μM) and (ii) arbitrary fluorescence units were then divided by the slope of the linear fit. See the study by Jung *et al.* [20] for a detailed description of MEF standardization.

IVT assays with *E. coli* lysate were prepared from lysate of *E. coli* BL21 Star as previously described [30]. Twenty units of RNase inhibitor (NEB, #M0314S) was added to the initial titration of lysate in IVT. Lysate samples were filtered through a molecular weight cut-off (MWCO) membrane using 3- and 10-kDa Amicon filters [Merck Millipore, catalog no.: UFC500308 (3 kDa), UFC501008 (10 kDa)] for 30 min at 14 000 × *g* and 4°C.

CETCH cycle assays

The production and purification of enzymes were performed as previously described by Sundaram *et al.* [5]. To test whether stopping reactions by removing enzymes through MWCO filtration yields the same glycolate concentration as stopping reactions by protein precipitation with formic acid, we ran a single CETCH cycle assay (Day 7, Condition 15 [16]) in an 80-μl volume (1.5-ml microcentrifuge tube; started with 100 μM propionyl-coenzyme A (CoA) substrate; 300 rpm shaking in a thermoshaker for 3 h at 30°C; see concentrations in Supplementary Table S6). Two 9-μl samples were quenched with 1 μl of 50% formic acid, and two 25-μl samples were filtered through a 10-kDa MWCO plate (PALL AcroPrep Advance 96-well filter plate; 350 μl, Omega 10K MWCO, catalog no.: 8034) by centrifugation (15 min, 2272 × *g*, 20°C). Two microliters of the samples was diluted in 18 μl of ddH_2O and used for quantification via LC–MS (method previously described by Pandi *et al.* [16]).

To generate different glycolate concentrations under constant buffer and cofactor conditions, we ran CETCH cycle assays in which only the methylsuccinyl-CoA oxidase (Mco) concentration was titrated. Six reaction mixtures of Condition 15 (see Supplementary Table S6) were prepared in a 125-μl volume with different concentrations of Mco: 2×, 1×, 0.5×, 0.25×, 0.1×, and no Mco (1× = 26 μM). Reaction mixtures were prepared in 1.5-ml microcentrifuge tubes, started with 100 μM propionyl-CoA, and shaken for 3 h at 30°C and 300 rpm in a thermoshaker. Samples were filtered, and glycolate was quantified as described earlier. Filtered samples were stored at −20°C.

To prepare CETCH cycle samples with varying buffer and cofactor conditions, samples were prepared in a 150-μl volume (1.5-ml microcentrifuge tube; started with the 100 μM propionyl-CoA substrate; 500-rpm shaking in a thermoshaker for 4 h at 30°C). Concentrations of individual CETCH cycle components were varied in the following ranges: HEPES (75–200 mM, pH 7.4–7.8), $MgCl_2$ (2.5–17.5 mM), creatine phosphate (CP) (5–60 mM), sodium bicarbonate (2.5–60 mM), sodium formate (10–60 mM), CoA (0–5 mM), coenzyme B_{12} (0–0.1 mM), adenosine triphosphate (ATP) (3–5 mM), dihydronicotinamide-adenine dinucleotide phosphate (NADPH) (2.5–10 mM), propionyl-CoA oxidase (0.10–9.57 μM), crotonyl-CoA carboxylase/reductase (0.62–2.78 μM), epimerase (0.74–6.70 μM), methylmalonyl-CoA mutase (0.61–2.89 μM), succinyl-CoA reductase (3.49–13.08 μM), succinic semialdehyde reductase (0.55–4.97 μM), 4-hydroxybutyryl-CoA synthetase (0.53–12.28 μM), 4-hydroxybutyryl-CoA dehydratase (0.73–3.64 μM), ethylmalonyl-CoA mutase (0.86–2.88 μM), Mco (26.01–46.54 μM), mesaconyl-CoA hydratase (0.28–2.84 μM), malyl-CoA/citramalyl-CoA lyase (2.79–14.73 μM), catalase (KatE, 2.46–8.21 μM), formate dehydrogenase (7.28–40.77 μM), creatine kinase (0.78–3.14 μM), carbonic anhydrase (0.02–0.13 μM), and glyoxylate/succinic semialdehyde reductase (3.31–5.25 μM). For a detailed overview of the involved enzymes, see the study by Sundaram *et al.* [5] and refer to Supplementary Table S6 for details. All assays were started with 0.1 mM propionyl-CoA. After 4 h, samples were filtered through 10-kDa MWCO spin filters (Amicon Ultra 0.5 ml, Merck Millipore, catalog no.: UFC501008), by centrifuging at 14 000 × *g* and 4°C for 20 min. Glycolate from filtrates was quantified as described earlier (with the minor difference that 10 μM internal ^{13}C-glycolate standard was used), and samples were stored at −20°C.

Overview of the offline IVT biosensing workflow (all details are described earlier)

(i) Prepare and run CETCH samples.
(ii) Filter CETCH samples through a 10-kDa MWCO membrane (plate or spin column-based).
(iii) Prepare dilution series NIST-traceable fluorescein standard and transfer to a 384-well plate.
(iv) Prepare ROSALIND reaction mixture on ice:
  (a) omit the CETCH sample and T7 RNAP
  (b) prepare in a 3.4× scale to prepare three replicates per sample.
(v) Aliquot 57.8 μl in PCR tube strips, and equilibrate at RT for 30 min to ensure good repression by GlcR.

(vi) Add 6.8 µl of CETCH samples to respective wells → 1:10 dilution of the CETCH sample in the IVT sensor.
(vii) Add 3.4 µl of T7 RNAP to each well.
(viii) Mix by pipetting up and down a volume of 40 µL.
(ix) Transfer 20 µl in triplicates in a 384-well plate, centrifuge the plate, and start plate reader measurement.
(x) Linearly regress fluorescein standard data to calculate MEF values from arbitrary units for data analysis.

Statistical analysis

All data presented are the mean of $n = 3$ technical replicates ± SD, unless otherwise stated in the respective caption. Pearson's correlations were calculated using GraphPad Prism 10.1.0 based on the average values of replicates. We chose Pearson's correlation over Spearman correlation to calculate the linear, rather than monotonic, relationship between LC–MS-quantified samples and sensor output. We used Student's *t*-test ($n = 2$) or one-way analysis of variance test ($n > 2$) with a significance threshold of $P > .05$ to statistically compare the datasets.

Websites

ProtParam was used to calculate the molecular weight and theoretical extinction coefficients of all proteins used in this study (https://web.expasy.org/protparam/) [31].

Results

Establishing GlcR as a glycolate-responsive IVT-based biosensor module

To develop IVT-based reporter systems for *in vitro* metabolic networks, we chose the CETCH cycle as the model system, because of its pioneering role in synthetic CO_2 fixation, its biological complexity (involving a total of 17 enzymes and 10 nonenzyme components, such as cofactors), and its recent use in a METIS-assisted optimization workflow in which 1000 different combinations (3000 samples) had been already tested [16].

The end product of the CETCH cycle is glycolate, which is produced from CO_2. We therefore set out to construct a glycolate-responsive sensor module from a ROSALIND DNA template and a suitable aTF. A glycolate-responsive transcription factor, GlcC from *E. coli*, was previously described. However, this protein acts as a transcriptional activator [32–34], which made it incompatible with the T7 promoter–based IVT system of ROSALIND, which strictly relies on transcriptional repression. Thus, we turned our attention to another aTF from *Paracoccus denitrificans* that was recently reported to regulate glycolate assimilation in the β-hydroxyaspartate cycle [35]. This GntR family transcriptional repressor, named GlcR, is encoded by *pden4400*, binds the intergenic sequence *pden4399-4400*, and was shown to unbind in the presence of glycolate [28], which made the protein an interesting candidate for our envisioned IVT-based biosensor.

We purified GlcR as a fusion protein with N-terminal maltose-binding protein (MGlcR) and confirmed its binding and unbinding from the intergenic sequence between *pden4400* and *pden4397-4399* in the absence and presence of glycolate, respectively, by electrophoretic mobility shift assays (EMSA) (Supplementary Fig. S1a). To identify fragments carrying a putative operator site (*glcO*, Supplementary Fig. S1b), we next split the 150 base pair (bp)–long intergenic sequence into six fragments, with each fragment composed of ∼60 bp in length and an ∼30-bp overlap with neighboring fragments. The sixth fragment also encoded the first 51 bp of *pden4399*. EMSA showed that MGlcR bound to four of the six fragments (fragment nos 2–5), and in particular to fragment #3, which was bound by MGclR ∼2- to 4-fold stronger as fragment nos 2, 4, and 5, indicating that GlcR has multiple operator sites.

We then focused on the putative operator site in fragment no. 3. However, further splitting of fragment no. 3 into 20- and 30-bp-long fragments completely abolished MGlcR binding, suggesting that the operator site spans >30 bp (Supplementary Fig. S1c). We next removed base pairs in 4-bp steps from the 5′ end of fragment no. 3 and prepared eight ROSALIND templates encoding putative *glcO* sequences between 60- and 32-bp length (named according to their length, i.e. the 60-bp-long sequence was named $glcO_{60}$) as part of a P_{T7}-*glcO*-3*WJdB* expression cassette (Supplementary Note S2). When tested for 3*WJdB* expression in the presence or absence of MGlcR, all eight constructs showed repression between 3-fold ($glcO_{48}$ and $glcO_{52}$) and up to 8-fold and 16-fold in the case of $glcO_{60}$ and $glcO_{36}$, respectively (Supplementary Fig. S2a).

We continued with the two best sensor constructs, $glcO_{36}$ and $glcO_{60}$, and titrated MGlcR over a constant DNA template concentration (25 nM). At 1.25 µM MGlcR (i.e. 50× aTF:DNA template ratio) $glcO_{36}$ showed an ∼80-fold repression, while $glcO_{60}$ required 5 µM MGlcR (i.e. 200× aTF:DNA template ratio) to reach a similar level of repression (Supplementary Fig. S2b). At these concentrations, both constructs showed an ∼3-fold derepression with 10 mM glycolate (Supplementary Fig. S2c). We selected $glcO_{36}$ as the final construct and sought to further increase the sensitivity of the system by reducing the total aTF concentration. It has recently been shown that removing excess aTF molecules, which act as effector chelators, improves the sensitivity of ROSALIND sensor modules [20]. When lowering the DNA template to 15 or 5 nM DNA [and keeping the aTF:DNA ratio at 50× (i.e. 0.75 µM and 0.25 µM MGlcR)], glycolate-induced derepression increased by 6- and 10-fold, respectively, while the signal was only reduced by 8 and 50%, respectively, resulting in a good balance between glycolate sensitivity and total output signal (Supplementary Fig. S2d). We tested the influence of the T7 RNAP preparation on the signal (Supplementary Fig. S3) and confirmed that the sensor was functional in HEPES buffer between pH 7.2 and 7.8, the buffer conditions of the CETCH cycle (Supplementary Fig. S4). As standard conditions for all subsequent IVT-based biosensing experiments, we chose HEPES buffer pH 7.8 with 15 nM $glcO_{36}$, 750 µM MGlcR, and in-house T7 RNAP, which we refer to as the GlcR sensor module.

The GlcR sensor module is operational over three orders of magnitude

To determine the operational range of the GlcR sensor module, we tested the response of the sensor to glycolate concentrations over six orders of magnitude (from 100 to 100 mM), which defined a limit of detection at 10 µM glycolate and an upper limit at 20 mM (Fig. 1b; Supplementary Fig. S5a and e; Table S8). We observed inhibition at glycolate concentrations of >20 mM. The GlcR sensor module showed an excellent linear response between 16 µM and 8 mM (Pearson's $r = 1.0$) and a dynamic range of 10.3-fold after 4 h of incubation (Fig. 1c; Supplementary Fig. S5c and d).

We also tested the specificity of the GlcR sensor module with seven structurally and context-related small organic acids: glyoxylate, acetate, DL-lactate, DL-glycerate, glycine, 2-phosphoglycolate, and 3-phospho-D-glycerate. Notably, the sensor was highly specific for glycolate and showed no response with other C2 acids, including glyoxylate and the amino acid glycine. However, we observed a low dose-responsive signal with DL-lactate and DL-glycerate (1.5- and 1.7-fold to 10 mM effector, respectively; in comparison, glycolate: 9.4-fold), indicating some promiscuity of GlcR with C3 alpha-hydroxy acids (Fig. 1d; Supplementary Fig. S6). However, since these C3 acids are not part of the CETCH cycle,

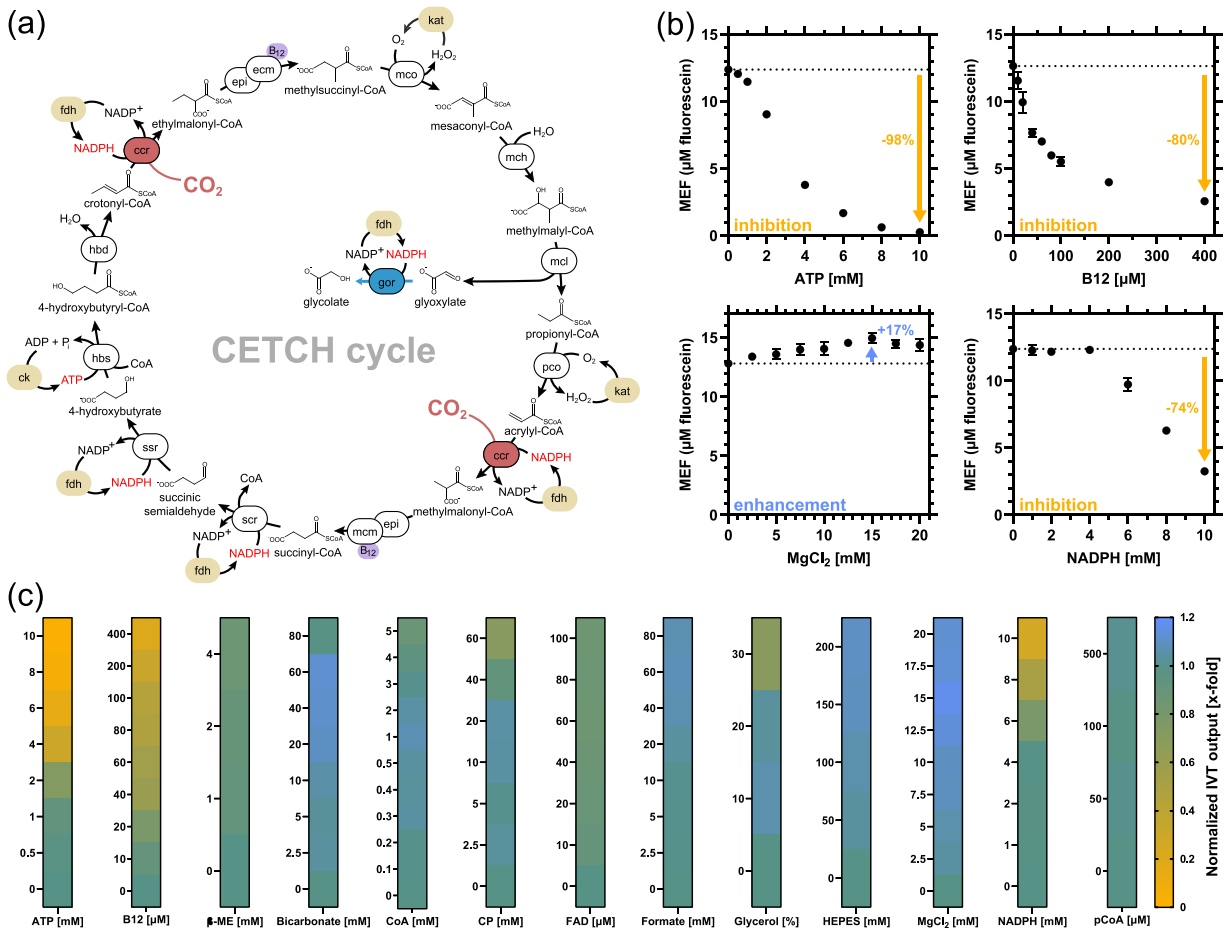


Figure 2. Individual influence of CETCH cycle components on IVT in the absence of GlcR. (a) Reaction sequence of the CETCH cycle to convert CO_2 into glycolate [2, 19]. Ten nonenzymatic components are actively involved in the cycle, and three additional components are required to maintain enzyme activity during storage. (b) Titration of ATP, coenzyme B_{12}, $MgCl_2$, and NADPH concentrations in IVT reactions shows a dose-dependent influence of each component on IVT output after 4 h. Detailed data for nine additional components are shown in Supplementary Fig. S7. Raw fluorescence data are standardized to MEF (µM fluorescein). Data are the mean of $n = 3$ technical replicates ± SD. (c) Heat map describing the influence of all 13 nonenzyme components of the CETCH cycle on IVT output [as shown in (b) and Supplementary Fig. S7]. Data are normalized to IVT output in the absence of the respective component to indicate inhibition, enhancement, and no effect of the screened component at the indicated concentration.

we concluded that the GlcR sensor module could be used for the envisioned IVT-based biosensing system.

Nonenzyme components of CETCH influence online IVT analysis

To assess whether online (i.e. direct) biosensing of CETCH samples would be possible with our IVT biosensor, we next tested the compatibility of IVT with components of the CETCH cycle. The CETCH cycle consists of 27 components—17 enzymes and 10 nonenzyme components, some of which are shared between CETCH and the IVT (e.g. $MgCl_2$ and ATP). Overall, these nonenzyme components include substrates (bicarbonate and propionyl-CoA), cofactors (ATP, $MgCl_2$, NADPH, CoA, and coenzyme B_{12}), metabolites for energy supply (formate and creatine phosphate), and a buffer reagent (HEPES) (Fig. 2a). In addition to these 10 nonenzyme components of the cycle, all CETCH enzymes are kept in 20% glycerol, and Pco and Mco are stored additionally with flavin adenine dinucleotide (FAD). We also tested β-mercaptoethanol (β-ME) because it is widely used to store enzymes.

We investigated the individual influence of these 13 different nonenzyme components on our basic IVT system (without the GlcR sensor module, each component titrated separately). For the 10 nonenzyme components of the CETCH cycle, we tested concentrations in our IVT were previously used during METIS-assisted optimization of the cycle by Pandi *et al.* [16]. For β-ME, FAD, and glycerol, we sampled a wider range of concentrations (Supplementary Table S5). Notably, 8 of the 13 nonenzyme components inhibited the IVT reaction, with ATP, NADPH, and coenzyme B_{12} showing strong inhibition of up to 74–98% at high concentrations (Fig. 2b and c; Supplementary Fig. S7). ATP inhibition was partially due to an increased competition between ATP and the other three rNTPs (Supplementary Fig. S8). However, the majority of ATP inhibition seemed to be caused by chelation of free magnesium ions, which are essential for T7 RNAP activity, similar to the inhibition of phi29 DNA polymerase by rNTPs [36]. We further speculated that NADPH, CP, CoA, and FAD followed a similar inhibitory mechanism. High concentrations of glycerol were inhibited by 30%. CoA, FAD, and β-ME showed minor inhibition between 12% and 15%, while propionyl-CoA showed no effect. In contrast, four nonenzyme components, bicarbonate, $MgCl_2$, formate, and HEPES increased IVT output by up to 15% (Fig. 2b and c; Supplementary Fig. S7). While additional Mg^{2+} probably directly increased T7 RNAP activity, the mechanisms behind the slightly increased IVT output of the other three components remained unknown.

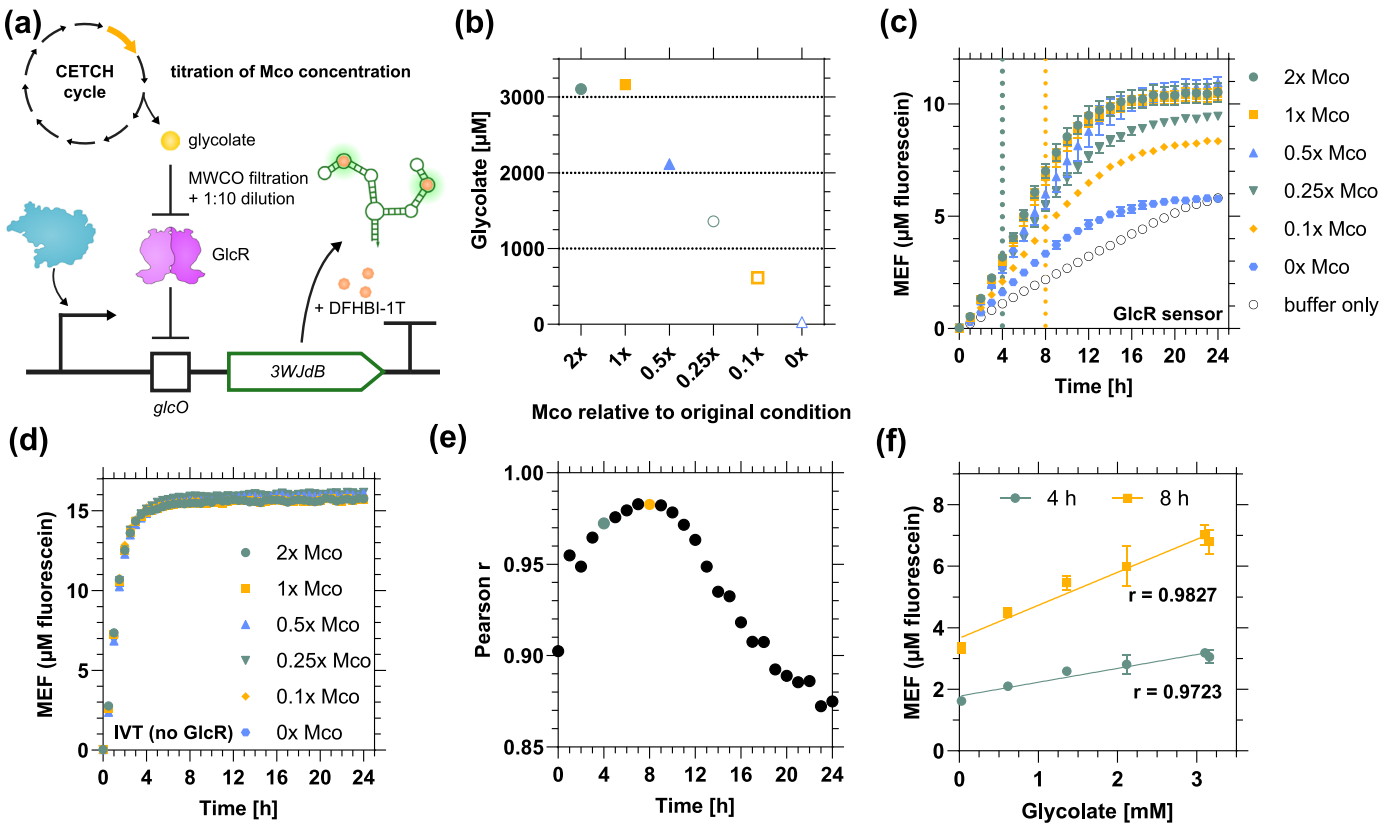


Figure 3. Glycolate sensing from CETCH cycle samples with a single component, the concentration of enzyme Mco, varied. (a) Schematic of the experimental setup. (b) LC–MS quantification of glycolate from six CETCH cycle samples with the titrated Mco concentration measured in technical triplicates. (c) Time course of glycolate measurement using the GlcR sensor module. Time points shown in (f) are indicated as vertical dotted lines. (d) Time course of IVT measurement in the absence of GlcR showing no differences in inhibition by CETCH cycle samples. (e, f) Correlation between GlcR sensor module output and LC–MS quantification [as shown in (b)] over time. The quality of the correlation is time sensitive and worsens as soon as the first IVT reactions plateau. Data of 4-h and 8-h time points are exemplarily shown in (f). Raw fluorescence data are standardized to MEF (µM fluorescein). Data are the mean of $n = 3$ technical replicates ± SD (c, d, f).

Overall, the complex (and partially adverse) effects of the different nonenzyme components on IVT showed that online measurements cannot be simply used for IVT-based biosensing under complex varying conditions of the CETCH cycle. We thus decided to work with an offline workflow, in which samples are quenched and diluted 1:10 before analysis to minimize the effects of the nonenzyme components on our IVT biosensor.

Establishing offline IVT sensing of CETCH samples with one enzyme component varied

We next developed an offline biosensing workflow, in which CETCH cycle variants are run first, and their output is analyzed using our IVT biosensor in a subsequent step (Fig. 3a). To quench CETCH reactions, we separated small molecules from enzymes by filtration through a 10-kDa molecular weight cutoff (MWCO) membrane before analysis of the filtrate (Supplementary Fig. S9) in a 1:10 dilution.

As a proof of concept, we measured glycolate production from quenched CETCH samples, in which only one enzyme was varied. To that end, we titrated Mco, a critical enzyme known to limit the productivity of the CETCH cycle [16]. We ran six CETCH cycle reactions (Day 7, Condition 15 in the study by Pandi *et al.* [16], Supplementary Table S6), with different Mco concentrations (0–52 µM). These six conditions yielded different glycolate concentrations (Fig. 3b), which our offline IVT biosensor was able to quantify with high correlations. The time course showed excellent correlation at 4 h ($r = 0.97$) and between 7 and 9 h ($r > 0.98$), after which the IVT reactions started to plateau and other factors became limiting (Fig. 3c, e, and f). Overall, these results demonstrated that our offline IVT-based biosensor workflow is able to screen the productivity of CETCH cycle variants with varying enzyme concentrations.

Optimizing offline sensing of CETCH samples with multiple components varied

We next tested whether our offline IVT biosensor workflow was able to quantify glycolate concentrations from CETCH samples of highly diverse composition (enzymes and nonenzyme components varied, Fig. 4a). We prepared six CETCH samples with known productivity (Day 7 in the study by Pandi *et al.* [16], as specified in Supplementary Table S6) and compared their IVT readout with LC–MS-based quantification of glycolate (Fig. 4b). Overall, the GlcR module and the LC–MS-based method showed $r = 0.77$ after 4 h (Fig. 4c and d; Supplementary Fig. S10), with Sample 2 underestimating and Sample 3 overestimating the actual glycolate concentrations, respectively.

Strikingly, in both samples, the concentration of (free) $MgCl_2$ seemed to be the critical factor. Sample 3 contained the highest $MgCl_2$ concentration (17.5 mM) and low concentrations of Mg^{2+}-binding cofactors (6.95 mM ATP, NADPH, and CoA). In contrast, Sample 2 contained the lowest $MgCl_2$ concentration (2.5 mM) and

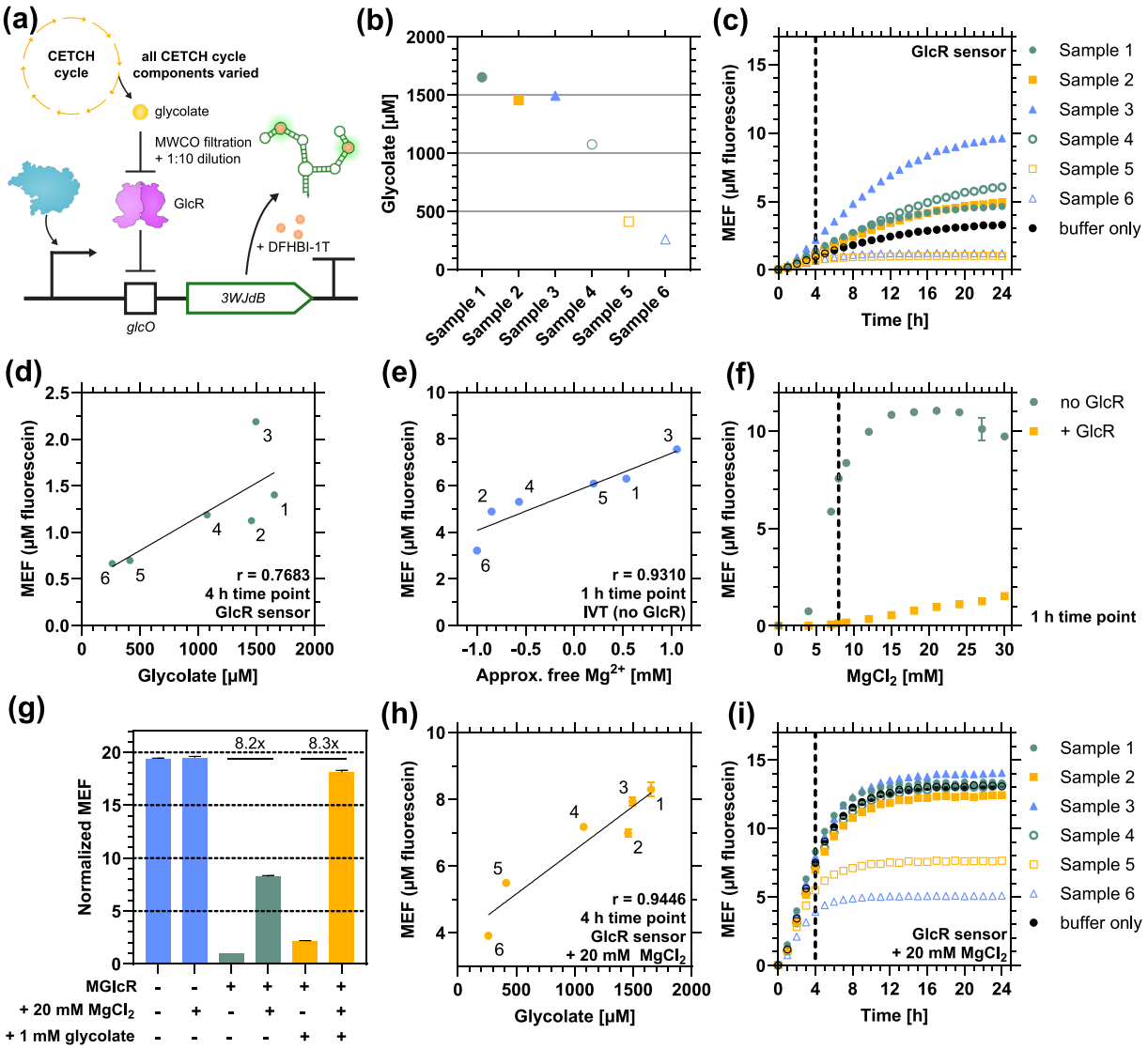


Figure 4. Glycolate sensing from CETCH cycle samples with varied concentrations of nonenzyme components and enzymes. (a) Schematic of the experimental setup. (b) LC–MS quantification of glycolate from six CETCH cycle samples of different compositions (detailed compositions listed in Supplementary Table S6), measured in technical triplicates. (c, i) Time course of glycolate measurement with the GlcR module without and with the addition of 20 mM $MgCl_2$, respectively. (d, h) Correlation between GlcR module output [4-h time point, indicated as dashed lines in (c, i)] and LC–MS quantification [as shown in (b)]. See Supplementary Fig. S10b for correlation coefficients of hourly time points. (e) Correlation between free Mg^{2+} and IVT output in the absence of GlcR. Free Mg^{2+} is approximated as the concentration difference of added $MgCl_2$ and Mg^{2+} binding to estimate the change in free Mg^{2+} upon addition of the CETCH sample to the GlcR sensor module. (f) Titration from 0 to 30 mM $MgCl_2$ showed a dose-dependent IVT output in the presence and absence of GlcR. In the presence of 750 nM GlcR, the leakiness of the repressed module increased linearly, whereas in the absence of GlcR, the IVT output was bell-shaped. (g) Effect of additional 20 mM $MgCl_2$ on the constitutive repressed and derepressed state of the GlcR sensor module. Elevated $MgCl_2$ concentrations increased the leakiness of the module but did not affect the response to 1 mM glycolate. Data shown are 4-h time points, normalized to data with only MGlcR added. (c–i) Raw fluorescence data are standardized to MEF (µM fluorescein). Data are the mean of $n = 3$ technical replicates ± SD.

high amounts of Mg^{2+}-binding cofactors (11 mM ATP, NADPH, and CoA). Because Mg^{2+} is the cofactor of T7 RNAP and its availability is essential for IVT (see earlier), we speculated that Mg^{2+} availability was strongly affecting the readout in these samples. This was supported by the fact that when we measured the effect of the six CETCH samples on the IVT system without the GlcR module, IVT output correlated well with the approximated concentration of free Mg^{2+} (Fig. 4e, $r = 0.93$).

We therefore decided to increase the overall Mg^{2+} concentration in our biosensor system to minimize the effect of CETCH cycle samples on Mg^{2+} availability. We examined the Mg^{2+} dependence of the IVT system and the GlcR module in the range of 0–30 mM $MgCl_2$ (Fig. 4f). Between 15 and 25 mM $MgCl_2$, the derepressed IVT system showed a broad plateau, while the Mg^{2+} effect on the repression of the system by GlcR was relatively small, indicating that this $MgCl_2$ concentration range was useful for robust sensing.

Indeed, when increasing the $MgCl_2$ concentrations from 8 mM to 20 mM in our IVT-based system, this significantly improved the correlation between GlcR module output and LC–MS to $r = 0.94$ (Fig. 4h; Supplementary Fig. S10), albeit at some increase in total output signal (Fig. 4i), caused by a higher baseline expression of the system (Fig. 4g). Overall, however, this setup established our GlcR IVT biosensor as a reliable glycolate quantification system

that worked robustly across different conditions. This was further confirmed by probing *E. coli* lysate spiked with glycolate (Supplementary Fig. S11), demonstrating the possibility of using our IVT-based biosensing also in bacterial lysates that have become an important platform for pathway prototyping, recently [11, 12, 14, 37].

Discussion

Here, we explored the potential of IVT-based biosensors in cell-free manufacturing and in particular synthetic biochemistry. As a proof of concept, we tested the sensing of glycolate production from the CETCH cycle, a synthetic CO_2 fixation cycle, which was only possible by LC–MS analysis, thus far. Our experiments demonstrate that IVT-based biosensing of complex samples is feasible, with excellent and robust correlations.

Key to establish our biosensor was the finding that IVT-based sensing is highly sensitive to (shared) components that are commonly used in *in vitro* systems, including ATP, NADPH, and other nucleotide-based and phosphorylated cofactors, which strongly inhibit IVT activity. These results are in line with recent studies that investigated the inhibition of reconstituted transcription, translation, and DNA replication systems [36]. In our study, we extend these findings by providing a fitness landscape of T7 RNAP-based transcription in the presence of common metabolic cofactors. These data are not only relevant for efforts to establish other IVT-based biosensors, but might also be helpful for efforts of integrating metabolic and IVT–translation systems toward constructing a synthetic cell [22–24], and to prototype genetic (RNA) circuits for cell-free biosensing and biocomputing under complex conditions and within artificial compartments [38–42].

On a more practical note, our work shows that IVT biosensing offers several advantages over classical metabolic quantification methods, because it allows us to improve throughput and cost efficiency and does not rely on expensive analytical instrumentation. For example, during METIS-assisted optimization of the CETCH cycle, Pandi *et al.* screened glycolate production in a 384-well format with LC–MS [16]. At a sample analysis time of ∼12 min, complete data analysis takes ∼80 h. In contrast, analysis with the GlcR module in the 384-well format is finished within 4–8 h, which is at least 10 times faster. In addition, the cost per sample for LC–MS is ∼US$7 (2020–2 average, without instrument purchase included), while IVT-based biosensing reagent costs are ∼US$0.5 for a 20-µl reaction mixture, which makes our system a more cost- and resource-efficient alternative to LC–MS measurements. Increased throughput at lower cost, however, comes at the expense of some accuracy, as various components can interfere with the IVT measurement, which makes LC–MS the unmatched tool for highly precise quantifications (see Supplementary Note S1 for a detailed comparison of IVT-based biosensing and LC–MS and Supplementary Table S7 for IVT reagent costs).

We note that the applicability of our approach depends on the availability of suitable aTF–operator pairs for sensing the metabolite of interest. Today, only a few of aTFs are sufficiently characterized and collected in curated databases such as GroovDB [43]. However, systems biology approaches have proven to be powerful in identifying the role of aTFs *in vivo* [44–48] and are complemented recently through computational approaches that were able to successfully predict new aTF–operator pairs [49–52] and enzymes converting nondetectable metabolites into detectable ones [53, 54]. At the same time, efforts to engineer aTFs for new effector specificities and enhanced properties are increasing [48, 55–57], which will hopefully increase the repertoire and availability of aTFs to construct new IVT biosensors in the future.

Acknowledgments

The authors would like to thank Lennart Schada von Borzyskowski, Katharina Kremer, Amir Pandi, Blake Rasor, and Scott Scholz for helpful discussions and Peter Claus for technical assistance with the operation of the LC–MS instrument. We also thank Lennart Schada von Borzyskowski and Katharina Kremer for providing the GlcR sequence and the intergenic sequence *pden4399-4400* and Amir Pandi for providing the *E. coli* strain sAP94 and his feedback on the manuscript.

Author contributions

Sebastian Barthel, Luca Brenker, Tobias J. Erb (Conceptualization), Sebastian Barthel, Luca Brenker (Methodology), Sebastian Barthel, Luca Brenker, Christoph Diehl, Nitin Bohra, Simone Giaveri (Investigation), Sebastian Barthel (Visualization), Sebastian Barthel, Tobias J. Erb (Writing—original draft); Sebastian Barthel, Tobias J. Erb (Writing—review & editing), Tobias J. Erb (Funding acquisition), Nitin Bohra (Resources), and Sebastian Barthel, Tobias J. Erb (Supervision).

Supplementary data

Supplementary data is available at *SYNBIO* Online.

Conflict of interest: None declared.

Funding

This project was funded by Max Planck Society general funding. S.G. is grateful to the European Molecular Biology Organization postdoctoral fellowship (ALTF 162-2022). N.B. conducted his research within the Max Planck School Matter to Life supported by the German Federal Ministry of Education and Research (BMBF) in collaboration with the Max Planck Society.

Data availability

Data underlying this article are available in FigShare at dx.doi.org/10.6084/m9.figshare.25869418. Plasmids were deposited with Addgene (pTE5400: #221189, pTE5417: #221191, pTE5418: #221192, pTE5419: #221193, and pTE5423: #221194).

References

1. Bowie JU, Sherkhanov S, Korman TP *et al.* Synthetic biochemistry: the bio-inspired cell-free approach to commodity chemical production. *Trends Biotechnol* 2020;**38**:766–78. https://doi.org/10.1016/j.tibtech.2019.12.024
2. Schwander T, Schada von Borzyskowski L, Burgener S *et al.* A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* 2016;**354**:900–4. https://doi.org/10.1126/science.aah5237
3. McLean R, Schwander T, Diehl C *et al.* Exploring alternative pathways for the in vitro establishment of the HOPAC cycle for synthetic CO_2 fixation. *Sci Adv* 2023;**9**:eadh4299. https://doi.org/10.1126/sciadv.adh4299
4. Luo S, Diehl C, He H *et al.* Construction and modular implementation of the THETA cycle for synthetic CO_2 fixation. *Nat Catal* 2023;**6**:1228–40. https://doi.org/10.1038/s41929-023-01079-z

5. Sundaram S, Diehl C, Cortina NS *et al*. A modular in vitro platform for the production of terpenes and polyketides from CO_2. *Angew Chem Int Ed* 2021;**60**:16420–25. https://doi.org/10.1002/anie.202102333
6. Diehl C, Gerlinger PD, Paczia N *et al*. Synthetic anaplerotic modules for the direct synthesis of complex molecules from CO_2. *Nat Chem Biol* 2023;**19**:168–75. https://doi.org/10.1038/s41589-022-01179-0
7. Valliere MA, Korman TP, Woodall NB *et al*. A cell-free platform for the prenylation of natural products and application to cannabinoid production. *Nat Commun* 2019;**10**:565. https://doi.org/10.1038/s41467-019-08448-y
8. Valliere MA, Korman TP, Arbing MA *et al*. A bio-inspired cell-free system for cannabinoid production from inexpensive inputs. *Nat Chem Biol* 2020;**16**:1427–33. https://doi.org/10.1038/s41589-020-0631-9
9. Korman TP, Opgenorth PH, Bowie JU. A synthetic biochemistry platform for cell free production of monoterpenes from glucose. *Nat Commun* 2017;**8**:1–8. https://doi.org/10.1038/ncomms15526
10. Claassens NJ, Burgener S, Vögeli B *et al*. A critical comparison of cellular and cell-free bioproduction systems. *Curr Opin Biotechnol* 2019;**60**:221–29. https://doi.org/10.1016/j.copbio.2019.05.003
11. Karim AS, Dudley QM, Juminaga A *et al*. In vitro prototyping and rapid optimization of biosynthetic enzymes for cell design. *Nat Chem Biol* 2020;**16**:912–19. https://doi.org/10.1038/s41589-020-0559-0
12. Vögeli B, Schulz L, Garg S *et al*. Cell-free prototyping enables implementation of optimized reverse β-oxidation pathways in heterotrophic and autotrophic bacteria. *Nat Commun* 2022;**13**:3058. https://doi.org/10.1038/s41467-022-30571-6
13. Liew FE, Nogle R, Abdalla T *et al*. Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale. *Nat Biotechnol* 2022;**40**:335–44. https://doi.org/10.1038/s41587-021-01195-w
14. Dudley QM, Karim AS, Nash CJ *et al*. In vitro prototyping of limonene biosynthesis using cell-free protein synthesis. *Metab Eng* 2020;**61**:251–60. https://doi.org/10.1016/j.ymben.2020.05.006
15. Kelwick R, Ricci L, Chee SM *et al*. Cell-free prototyping strategies for enhancing the sustainable production of polyhydroxyalkanoates bioplastics. *Synth Biol* 2018;**3**:ysy016. https://doi.org/10.1093/synbio/ysy016
16. Pandi A, Diehl C, Yazdizadeh Kharrazi A *et al*. A versatile active learning workflow for optimization of genetic and metabolic networks. *Nat Commun* 2022;**13**:3876. https://doi.org/10.1038/s41467-022-31245-z
17. Sakai A, Jonker AJ, Nelissen FHT *et al*. Cell-free expression system derived from a near-minimal synthetic bacterium. *ACS Synth Biol* 2023;**12**:1616–23. https://doi.org/10.1021/acssynbio.3c00114
18. Morini L, Sakai A, Vibhute MA *et al*. Leveraging active learning to establish efficient in vitro transcription and translation from bacterial chromosomal DNA. *ACS Omega* 2024;**9**:19227–35. https://doi.org/10.1021/acsomega.4c00111
19. Miller TE, Beneyton T, Schwander T *et al*. Light-powered CO_2 fixation in a chloroplast mimic with natural and synthetic parts. *Science* 2020;**368**:649–654. https://doi.org/10.1126/science.aaz6802
20. Jung JK, Alam KK, Verosloff MS *et al*. Cell-free biosensors for rapid detection of water contaminants. *Nat Biotechnol* 2020;**38**:1451–59. https://doi.org/10.1038/s41587-020-0571-7
21. Alam KK, Tawiah KD, Lichte MF *et al*. A fluorescent split aptamer for visualizing RNA–RNA assembly in vivo. *ACS Synth Biol* 2017;**6**:1710–21. https://doi.org/10.1021/acssynbio.7b00059
22. Giaveri S, Bohra N, Diehl C *et al*. Integrated translation and metabolism in a partially self-synthesizing biochemical network. *Science* 2024;**385**:174–78. https://doi.org/10.1126/science.adn3856
23. Rothschild LJ, Averesch NJH, Strychalski EA *et al*. Building synthetic cells—from the technology infrastructure to cellular entities. *ACS Synth Biol* 2024;**13**:974–97. https://doi.org/10.1021/acssynbio.3c00724
24. Schwille P, Spatz J, Landfester K *et al*. MaxSynBio: avenues towards creating cells from the bottom up. *Angew Chem Int Ed* 2018;**57**:13382–92. https://doi.org/10.1002/anie.201802288
25. Stukenberg D, Hensel T, Hoff J *et al*. The Marburg collection: a Golden Gate DNA assembly framework for synthetic biology applications in vibrio natriegens. *ACS Synth Biol* 2021;**10**:1904–19. https://doi.org/10.1021/acssynbio.1c00126
26. Sun ZZ, Yeung E, Hayes CA *et al*. Linear DNA for rapid prototyping of synthetic biological circuits in an Escherichia coli based TX-TL cell-free system. *ACS Synth Biol* 2014;**3**:387–97. https://doi.org/10.1021/sb400131a
27. Lehr F-X, Gaizauskaite A, Lipińska KE *et al*. Modular Golden Gate assembly of linear DNA templates for cell-free prototyping. 2023. https://doi.org/10.48550/arXiv.2310.13665
28. Schada von Borzyskowski L, Hermann L, Kremer K *et al*. Multiple levels of transcriptional regulation control glycolate metabolism in Paracoccus denitrificans. *mBio* 2024;**15**:e01524. https://doi.org/10.1128/mbio.01524-24
29. Shimizu Y, Inoue A, Tomari Y *et al*. Cell-free translation reconstituted with purified components. *Nat Biotechnol* 2001;**19**:751–55. https://doi.org/10.1038/90802
30. Rasor BJ, Vögeli B, Jewett MC *et al*. Cell-free protein synthesis for high-throughput biosynthetic pathway prototyping. In: Karim AS, Jewett MC (eds), *Methods in Molecular Biology*. New York, NY: Springer US, 2022, 199–215.
31. Gasteiger E, Hoogland C, Gattiker A *et al*. Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed.), *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press, 2005, 571–607.
32. Pellicer MT, Badía J, Aguilar J *et al*. glc locus of Escherichia coli: characterization of genes encoding the subunits of glycolate oxidase and the glc regulator protein. *J Bacteriol* 1996;**178**:2051–59. https://doi.org/10.1128/jb.178.7.2051-2059.1996
33. Pellicer MT, Fernandez C, Badía J *et al*. Cross-induction of glc and ace Operons of Escherichia coli attributable to pathway intersection: characterization of the glc promoter. *J Biol Chem* 1999;**274**:1745–52. https://doi.org/10.1074/jbc.274.3.1745
34. Xu S, Zhang L, Zhou S *et al*. Biosensor-based multi-gene pathway optimization for enhancing the production of glycolate. *Appl Environ Microbiol* 2021;**87**:e00113–21. https://doi.org/10.1128/AEM.00113-21
35. Schada von Borzyskowski L, Severi F, Krüger K *et al*. Marine Proteobacteria metabolize glycolate via the β-hydroxyaspartate cycle. *Nature* 2019;**575**:500–4. https://doi.org/10.1038/s41586-019-1748-4
36. Seo K, Ichihashi N. Investigation of compatibility between DNA replication, transcription, and translation for in vitro central dogma. *ACS Synth Biol* 2023;**12**:1813–22. https://doi.org/10.1021/acssynbio.3c00130
37. Rasor BJ, Vögeli B, Landwehr GM *et al*. Toward sustainable, cell-free biomanufacturing. *Curr Opin Biotechnol* 2021;**69**:136–44. https://doi.org/10.1016/j.copbio.2020.12.012
38. Takahashi MK, Chappell J, Hayes CA *et al*. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription–translation (TX-TL) systems. *ACS Synth Biol* 2015;**4**:503–15. https://doi.org/10.1021/sb400206c
39. Boyd MA, Thavarajah W, Lucks JB *et al*. Robust and tunable performance of a cell-free biosensor encapsulated in lipid

vesicles. *Sci Adv* 2023;**9**:eadd6605. https://doi.org/10.1126/sciadv.add6605
40. Sharon JA, Dasrath C, Fujiwara A *et al*. Trumpet is an operating system for simple and robust cell-free biocomputing. *Nat Commun* 2023;**14**:2257. https://doi.org/10.1038/s41467-023-37752-x
41. Schoenmakers LLJ, Yewdall NA, Lu T *et al*. In vitro transcription–translation in an artificial biomolecular condensate. *ACS Synth Biol* 2023;**12**:2004–14. https://doi.org/10.1021/acssynbio.3c00069
42. Gonzales DT, Yandrapalli N, Robinson T *et al*. Cell-free gene expression dynamics in synthetic cell populations. *ACS Synth Biol* 2022;**11**:205–15. https://doi.org/10.1021/acssynbio.1c00376
43. d'Oelsnitz S, Love JD, Diaz DJ *et al*. GroovDB: a database of ligand-inducible transcription factors. *ACS Synth Biol* 2022;**11**:3534–37. https://doi.org/10.1021/acssynbio.2c00382
44. Lempp M, Farke N, Kuntz M *et al*. Systematic identification of metabolites controlling gene expression in E. coli. *Nat Commun* 2019;**10**:4463. https://doi.org/10.1038/s41467-019-12474-1
45. Donati S, Kuntz M, Pahl V *et al*. Multi-omics analysis of CRISPRi-knockdowns identifies mechanisms that buffer decreases of enzymes in *E. coli* metabolism. *Cell Systems* 2021;**12**:56–67.e6. https://doi.org/10.1016/j.cels.2020.10.011
46. Gagarinova A, Hosseinnia A, Rahmatbakhsh M *et al*. Auxotrophic and prototrophic conditional genetic networks reveal the rewiring of transcription factors in Escherichia coli. *Nat Commun* 2022;**13**:4085. https://doi.org/10.1038/s41467-022-31819-x
47. Rodionova IA, Gao Y, Monk J *et al*. A systems approach discovers the role and characteristics of seven LysR type transcription factors in Escherichia coli. *Sci Rep* 2022;**12**:7274. https://doi.org/10.1038/s41598-022-11134-7
48. Pearson AN, Incha MR, Ho CN *et al*. Characterization and diversification of AraC/XylS family regulators guided by transposon sequencing. *ACS Synth Biol* 2024;**13**:206–19. https://doi.org/10.1021/acssynbio.3c00441
49. Hanko EKR, Paiva AC, Jonczyk M *et al*. A genome-wide approach for identification and characterisation of metabolite-inducible systems. *Nat Commun* 2020;**11**:1213. https://doi.org/10.1038/s41467-020-14941-6
50. Hanko EKR, Joosab Noor Mahomed TA, Stoney RA *et al*. TFB-Miner: a user-friendly command line tool for the rapid mining of transcription factor-based biosensors. *ACS Synth Biol* 2023;**12**:1497–507. https://doi.org/10.1021/acssynbio.2c00679
51. d'Oelsnitz S, Love JD, Ellington AD *et al*. Ligify: automated genome mining for ligand-inducible transcription factors. *ACS Synth Biol* 2024;**13**:2577–86. https://doi.org/10.1021/acssynbio.4c00372
52. d'Oelsnitz S, Stofel SK, Love JD *et al*. Snowprint: a predictive tool for genetic biosensor discovery. *Commun Biol* 2024;**7**:1–9. https://doi.org/10.1038/s42003-024-05849-8
53. Delépine B, Libis V, Carbonell P *et al*. SensiPath: computer-aided design of sensing-enabling metabolic pathways. *Nucleic Acids Res* 2016;**44**:W226–31. https://doi.org/10.1093/nar/gkw305
54. Pandi A, Koch M, Voyvodic PL *et al*. Metabolic perceptrons for neural computing in biological systems. *Nat Commun* 2019;**10**:3880. https://doi.org/10.1038/s41467-019-11889-0
55. Snoek T, Chaberski EK, Ambri F *et al*. Evolution-guided engineering of small-molecule biosensors. *Nucleic Acids Res* 2020;**48**:e3. https://doi.org/10.1093/nar/gkz954
56. d'Oelsnitz S, Kim W, Burkholder NT *et al*. Using fungible biosensors to evolve improved alkaloid biosyntheses. *Nat Chem Biol* 2022;**18**:981–89. https://doi.org/10.1038/s41589-022-01072-w
57. d'Oelsnitz S, Nguyen V, Alper HS *et al*. Evolving a generalist biosensor for bicyclic monoterpenes. *ACS Synth Biol* 2022;**11**:265–72. https://doi.org/10.1021/acssynbio.1c00402

Synthetic Biology, 2024, 9(1), ysae013 , DOI: https://doi.org/10.1093/synbio/ysae013, Advance Access Publication 20 September 2024, Research Article
Submitted: 18 September 2024; **Received**: 27 August 2024; **Accepted**: 18 September 2024