nature synthesis

Article

https://doi.org/10.1038/s44160-025-00819-2

Chemoenzymatic platform with coordinated cofactor self-circulation for lignin valorization

Received: 26 August 2024

Accepted: 6 May 2025

Published online: 10 June 2025

Check for updates

Liangxu Liu[1,2], Xiaotong Wang[3,4], Yeyun Wu[4,5], Hengrun Li[1,2], Jiawei Wang[1,2], Chaofeng Li[1,2], Haotian Zheng[1,2] & Jun Ni[1,2,3]

Lignin valorization is vital for achieving economically viable and sustainable lignocellulosic biorefineries. However, the value-added utilization of lignin constituents is hampered by the requirement of expensive cofactors and low conversions. Here, by integrating coenzyme regeneration and cell-free expression, we report an in vitro multienzyme-coordinated expression with cofactor self-circulation (iMECS) strategy to achieve efficient lignin-to-molecule conversion. The iMECS system established a catalytic bridge with high atom economy for converting lignin waste into valuable aromatic compounds. Curcumin, vanillin and raspberry ketone were efficiently synthesized in a coenzyme-free manner, with a conversion of over 90%. We also demonstrated the flexibility of the iMECS platform, through which a wide array of phenylpropanoids could be easily obtained by enzyme swapping and pathway extension. By integrating chemical depolymerization with the iMECS system, lignin-rich agricultural waste was directly converted into valuable chemicals, and the overall catalytic efficiency was enhanced by up to 48-fold compared with the reported titres. This efficient, general platform can facilitate the utilization of lignocellulosic biomass, thereby promoting zero-waste biorefineries for a more sustainable future.

Biorefining lignocellulosic biomass into value-added products is a promising route for coupling petrochemical replacement with carbon mitigation[1–4]. Conventional biorefineries generate more than 200 million tonnes of lignin waste annually, and the largely unexploited waste is simply burned for heat, raising environmental concerns about carbon emissions[5,6]. Recently, the biological utilization of lignin and its hydrolytic aromatics has attracted extensive attention, which is key to realizing zero-waste and economically viable biorefineries[7–9]. However, the constant requirement of expensive cofactors (for example, adenosine triphosphate (ATP), reduced nicotinamide adenine dinucleotide phosphate (NADPH) and coenzyme A (CoA)) has seriously hindered the sustainable valorization of lignin into valuable products[6,10]. To mitigate this constraint, one potential approach is to introduce auxiliary cofactor regeneration enzymes, which often necessitates precise regulation of the cofactor metabolic network to prevent disruption of cellular homeostasis[11,12]. Moreover, the imbalance in catalytic activity between biosynthetic machinery and cofactor regeneration systems is a key barrier to efficient conversion, resulting in extremely limited yields of lignin-valorized molecules. For example, the productivity of raspberry ketone (RK) from lignin-derived *p*-coumaric acid (*p*CA) was less than 0.003 g l^{-1} h^{-1}, with a conversion of only 11.3% (refs. 13,14), and the productivity of curcumin was limited to 0.006 g l^{-1} h^{-1} (ref. 15). Currently, there is no general solution to overcome expensive cofactor requirements and unbalanced enzymatic activity simultaneously; thus, developing a universal strategy for improving lignin bioconversion routes is desirable.

The efficiency of individual molecular machines should be carefully balanced to obtain a coordinated cofactor self-circulation system

[1]State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China. [2]Zhangjiang Institute for Advanced Study, Shanghai Jiao Tong University, Shanghai, China. [3]Innovation Center for Synthetic Biotechnology, Lumy Biotechnology, Changzhou, China. [4]Research Center for Proteins and Bits, Lumy Biotechnology, Shanghai, China. [5]School of Life Sciences, Fudan University, Shanghai, China. e-mail: tearroad@sjtu.edu.cn

https://doi.org/10.1038/s44160-025-00819-2

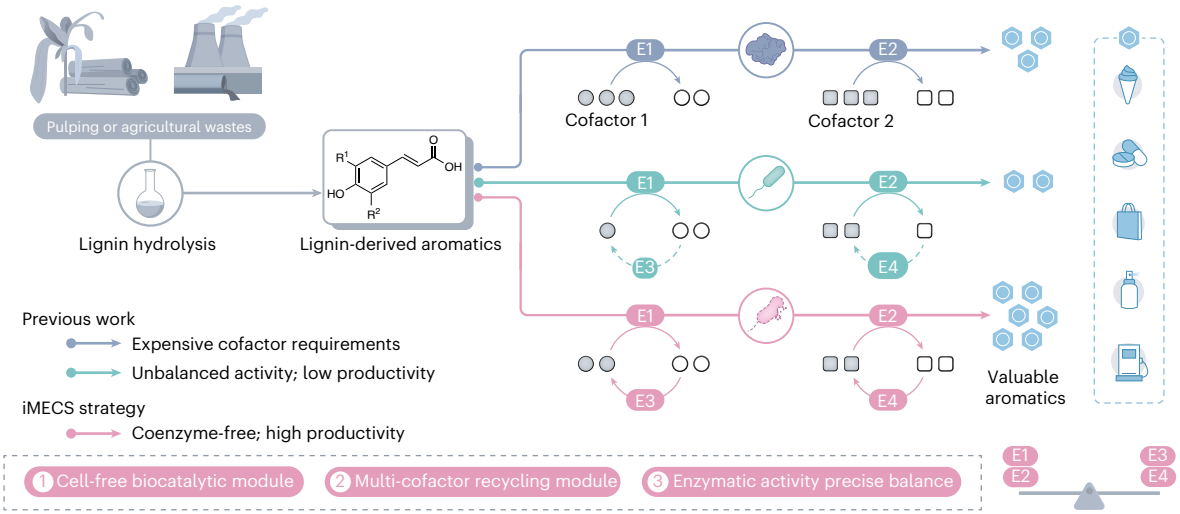

Fig. 1 | Schematic representation of distinct routes for lignin biovalorization. The lignin in pulping or agricultural wastes can be hydrolysed into lignin-derived aromatics, and then these aromatics can be biotransformed into target products. The purple arrow represents purified-enzyme-mediated transformation, which requires many expensive cofactors. The green arrow represents microbial transformation, and the imbalance of enzyme activity often leads to low yields. The pink arrow represents iMECS based on CFPS, which achieves the efficient lignin-to-molecule conversion without cofactors. The icons at the centre of the pathways represent (from top to bottom) enzymatic catalysis, whole-cell catalysis and cell-free catalysis, respectively. E1 and E2 refer to enzymes for conversion process, while E3 and E4 represent enzymes used for cofactor regeneration.

for lignin valorization. Innovative synthetic biology and metabolic engineering tools enable fine-tuning of gene expression for pathway balancing in vivo[11,16–18]. Nevertheless, optimizing the efficiency of numerous enzymes in parallel remains an extremely time-consuming and labour-intensive process, especially for nonlinear interacting pathways like lignin biorefinery in which multiple cofactors are involved[9,12]. Although a cell-free system with purified enzymes is an easily controllable candidate for biomanufacturing, it requires an expensive purification process[19]. Alternatively, by decoupling biosynthetic reactions from strict cellular regulations, cell-free protein synthesis (CFPS)-driven metabolic engineering offers tremendous flexibility in facilitating pathway balancing by adjusting the concentration of linear DNA or plasmids[20–22]. Moreover, the open nature of this system circumvents the limitations associated with aromatic toxicity, substrate uptake and product secretion[23]. We surmise that the CFPS-based approach holds appealing potential for orchestrating multiple molecular machines in lignin valorization. However, no attempts have been made to accelerate the conversion of lignin into hydrolytic aromatics.

To address this issue, we report a broadly applicable strategy that seamlessly integrates coenzyme regeneration and cell-free expression into the lignin upgrading process, termed the in vitro multienzyme-coordinated expression with cofactor self-circulation (iMECS) strategy (Fig. 1). The lignin-to-molecule conversion pipeline is divided into three main steps: (1) selecting suitable enzymes for the cell-free biocatalytic module, (2) constructing a multi-cofactor recycling module and (3) fine-tuning DNA addition to achieve a precise balance of enzymatic activity. In this study, using an automated liquid-handling workflow, we rapidly screened 1,296 unique pathway combinations with more than 20 reaction component sets, which boosted curcumin productivity to 0.175 g l^{-1} h^{-1}, with a molar yield of 95.31%. We also demonstrated the flexibility of the iMECS platform, where RK and other phenylpropanoids could be easily obtained by enzyme swapping and pathway extension. Without the addition of external cofactors, the overall catalytic efficiency of the iMECS-based route was enhanced by up to 48-fold compared with the reported titres. Next, a tandem chemical depolymerization and coenzyme-free biotransformation process was implemented to produce valuable aromatics directly from lignin-rich agricultural waste. This general sustainable system with extraordinary efficiency is highly feasible for lignin valorization, which promotes the economic feasibility of next-generation biorefineries.

Results

Multidimensional optimization of lignin-to-molecule route

To evaluate the feasibility of the iMECS strategy, we chose curcumin, a therapeutic polyphenol with an estimated global market of US$151.9 million by 2027[24], as the initial target product. Curcumin is synthesized from ferulic acid (FA), a predominant lignin monomer, by coupling *p*-coumarate-CoA ligase (4CL) with curcuminoid synthase (CUS) (Fig. 2a)[25]. First, we established a biocatalytic module consisting of purified 4CL from *Arabidopsis thaliana* and CUS from *Oryza sativa*, and the necessary cofactors (ATP, CoA and malonyl-CoA) were adequately added to drive the biosynthetic route. However, only 28.1 ± 2.9 mg l^{-1} of curcumin was obtained from 1 mM FA (194 mg l^{-1}) in this one-pot two-enzyme system (Fig. 2b), and the considerable accumulation of FA suggests that 4CL may be the rate-limiting step. Therefore, we investigated alternative enzymes derived from various sources for FA conversion. Fortunately, feruloyl-CoA synthetase (FCS) was identified from *Caldimonas thermodepolymerans* (Fig. 2c and Supplementary Fig. 1), enabling the highest curcumin production of 115.3 ± 4.4 mg l^{-1}, which was 4.1-fold higher than that of plant 4CL (Fig. 2b). We surmised that its high catalytic activity and excellent thermal stability make 4CL an ideal molecular machine (Supplementary Fig. 2). To circumvent the stoichiometric consumption of cofactors, an in situ multi-cofactor recycling module was designed and introduced. In this regard, we used malonyl-CoA synthetase (MatB)[26] and polyphosphate kinase 2 Class III (PPK)[27] for malonyl-CoA and ATP replenishment with synchronous CoA recycling (Fig. 2a and Supplementary Fig. 3). The four-enzyme cascade yielded 126.4 mg l^{-1} of curcumin using substoichiometric amounts of cofactors and sacrificial substrates (malonate and polyphosphate) (Supplementary Fig. 4), demonstrating that the artificial multi-cofactor recycling module successfully cycled the cofactors. Nevertheless, eliminating expensive purification processes and improving overall efficiency remain challenging.

With the working multienzyme cascade in hand, we set out to establish an iMECS system for the efficient synthesis of curcumin without protein purification. A high-yielding *Escherichia coli* lysate-based cell-free transcription–translation system was used for the rapid

https://doi.org/10.1038/s44160-025-00819-2

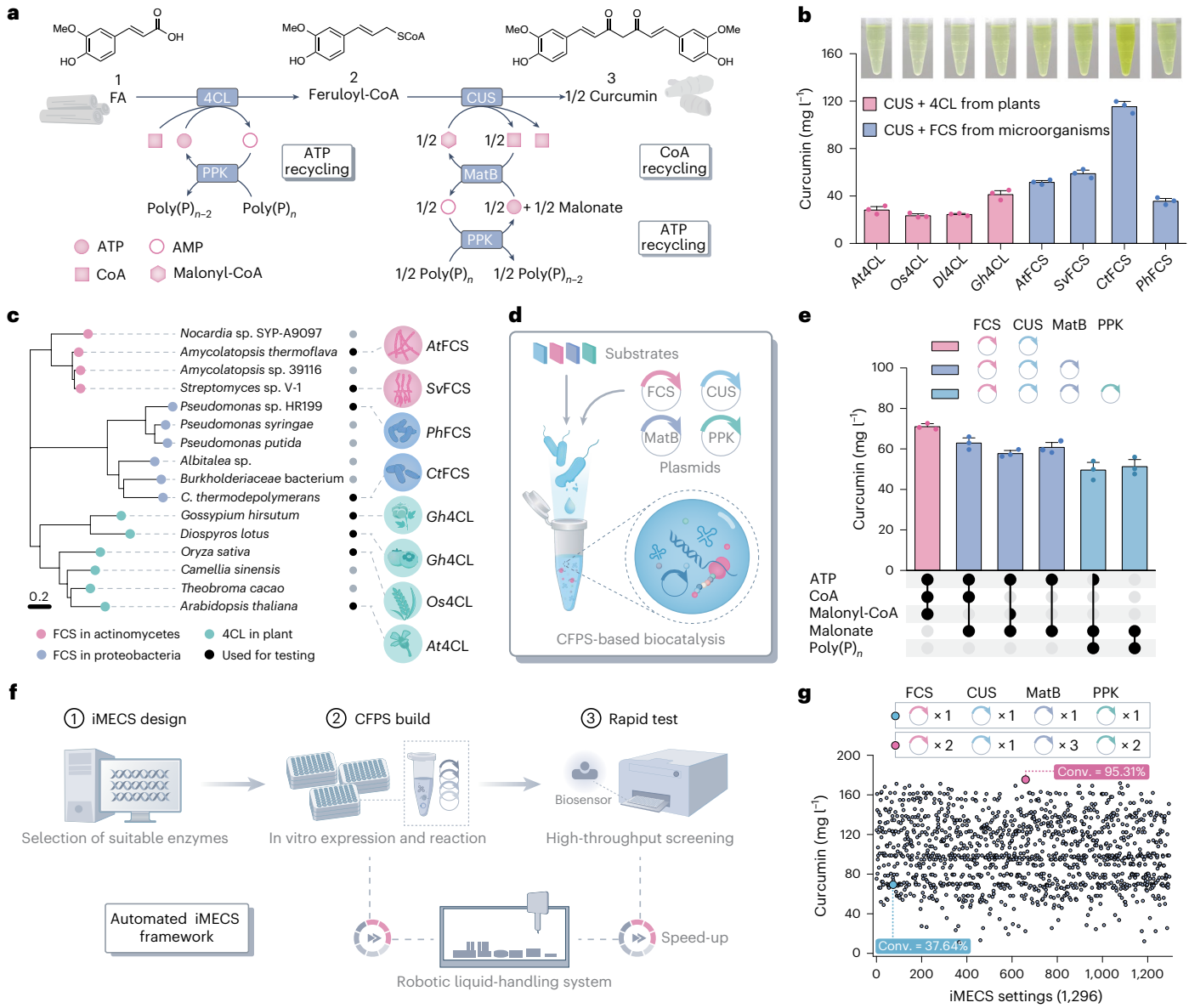

Fig. 2 | iMECS strategy for in vitro biosynthesis of curcumin from FA. **a**, An overview of the metabolic pathway for the formation of curcumin based on the iMECS strategy. **b**, The synthesis of curcumin using purified CUS and 4CL–FCS from various sources (marked as black circles in **c**). Values are shown as mean ± s.d. ($n = 3$ independent replicates); dots represent individual data points. **c**, Phylogenetic tree analysis of FCS–4CL indicating the evolutionary relativity and the homological degrees among the amino acid sequences of FCS–4CL. The sources of enzymes used in this work are marked as black circles. **d**, A schematic overview of the CFPS-based biocatalysis. The enzymatic pathway was assembled by adding the plasmid mixture to the CFPS system, followed by the addition of substrates for catalytic reactions. **e**, Curcumin production by CFPS-based biocatalysis. The black circles refer to excessive addition, while the black semicircles refer to substoichiometric addition (1/10 of the FA). All data are plotted for $n = 3$ independent experiments. **f**, A practical overview of the automated iMECS framework. The workflow consists of three parts: selection of suitable designs (design), in vitro protein expression and substrate transformation (build) and high-throughput screening (test). The protein biosensor used here is described in Supplementary Fig. 8. **g**, Curcumin end-point titres obtained in screening 1,296 plasmid combinations with different concentration ratios. Curcumin is measured at 1 h after the addition of FA. The conversion (Conv.) refers to the mol mol^{-1} measurement normalized to the theoretical maximum conversion. The initial concentration of FA in these experiments was 1 mM. Values are shown as mean ± s.d. ($n = 3$ independent replicates).

expression of 4CL, CUS, MatB and PPK using mix-and-match plasmids (Fig. 2d and Supplementary Fig. 5). After in vitro protein synthesis, 71.0 mg l^{-1} of curcumin was obtained from 1 mM FA with sufficient malonyl-CoA and ATP (Fig. 2e). Next, the cofactors were removed individually, and the corresponding curcumin production revealed that the cofactor recycling module was critical for the curcumin yield. Intriguingly, we found that the cofactor recycling reaction without replenishing cofactors also produced 51.2 mg l^{-1} of curcumin (Fig. 2e), indicating that residual cofactors in the cell lysate were sufficient to initiate the iMECS system. The distribution of ATP to adenosine monophosphate (AMP) in the PPK-mediated system was lower than in the system with excessive ATP addition, which may indicate that the cofactor recycling system could not regenerate ATP fast enough to match the curcumin biosynthesis (Supplementary Fig. 6). Here, although the iMECS strategy achieved the synthesis of curcumin without additional cofactors, the low conversion (<30%, mol mol^{-1} measurement normalized to theoretical maximum conversion) indicates the necessity for further optimization. After the first-layer optimization of sacrificial substrate concentrations (Supplementary Fig. 7), we sought to precisely balance enzyme activities by simply tuning the ratio of each plasmid. Inspired

by the in vitro prototyping and rapid optimization of biosynthetic enzymes (iPROBE) workflow[20], we automated the iMECS framework by combining a robotic liquid-handling system with an enzyme-coupled biosensor to accelerate the optimization process (Fig. 2f). We rapidly screened 1,296 unique pathway combinations in one day, and the loading ratio of FCS:CUS:MatB:PPK at 2:1:3:2 had the highest curcumin productivity of 0.175 g l^{-1} h^{-1} (Fig. 2g and Supplementary Table 1), which was more than 1,455% higher than that of typical whole-cell bioconversion (Supplementary Table 1). Correspondingly, the conversion of curcumin by the iMECS system reached 95.31% (Fig. 2g), far exceeding that reported in previous studies (Supplementary Table 1). These data demonstrate the superiority of the iMECS system in rapidly establishing a coenzyme-free route for the conversion of lignin-derived molecules.

Thermostable iMECS for energy-efficient vanillin synthesis

To evaluate the broad applicability of the iMECS system, we further explored the catalytic process for converting FA to vanillin, a widely used flavour compound with an estimated market of US$724.5 million[28]. FCS and enoyl-CoA hydratase (ECH) in *Streptomyces* sp. V-1 are responsible for this conversion: FCS ligates FA and CoA to form feruloyl-CoA with ATP consumption, and ECH subsequently converts feruloyl-CoA into vanillin and acetyl-CoA (Fig. 3a and Supplementary Fig. 9)[29]. The shunt product, acetyl-CoA, contains a high-energy thioester bond, which leads to energy wastage and hinders CoA recycling. To lower the ATP requirement, we designed a two-step energy extraction cascade consisting of phosphotransacetylase (PTA) and acetate kinase (ACK) as a 'molecular connector' between ATP and CoA (Fig. 3a). Specifically, PTA uses the by-product acetyl-CoA for CoA release, and the simultaneously produced acetyl phosphate can be perfectly used by ACK for ATP regeneration[30]. An additional enzyme, polyphosphate:AMP phosphotransferase (PAP)[31], which is responsible for energy and phosphate supply, was introduced into our design (Fig. 3a). With a full set of the five purified enzymes, the cofactors were successfully cycled to drive the conversion of 1 mM FA, and 94.4 ± 2.7 mg l^{-1} of vanillin was obtained using substoichiometric cofactors (Fig. 3b). The distribution of ATP to AMP in the five-enzyme system was higher than in the PPK-mediated system (Supplementary Fig. 6b), which may indicate that the newly constructed cofactor recycling system is more efficient. Here, the five-enzyme system realizes the energy-efficient synthesis of vanillin, and the energy extraction cascade can theoretically reduce ATP demand by 67% compared with the dual-enzyme system comprising FCS and ECH, as only 15.06 kJ mol^{-1} are invested in the ATP regeneration.

Next, we attempted to further optimize the efficiency of vanillin and avoid enzyme purification using the iMECS system. For the initial FA conversion, a cell-free dual plasmid system for FCS and ECH expression was used, which produced 94.9 mg l^{-1} of vanillin with a yield of 0.62 mol mol^{-1} FA (Fig. 3c). Notably, a certain amount of vanillyl alcohol accumulated in the reaction solution (Supplementary Fig. 10a), and several non-specific alcohol dehydrogenases (ADHs) in *E. coli* lysate may be responsible for this undesired by-product. Previously, a temperature-directed biocatalytic process was developed to avoid side reactions in *E. coli* whole-cell catalysis[29]. This strategy inactivated the endogenous ADHs through high temperature (50 °C), while thermostable exogenous enzymes could maintain their activities to complete catalytic synthesis[29]. Inspired by this methodology, we propose that a temperature-based method could be applied in a cell-free system. This method would deactivate the endogenous mesophilic enzymes present in *E. coli* extracts at higher temperatures, while the expressed thermostable enzymes would remain active (Fig. 3d). The thermostability of PTA, ACK and PAP at 50 °C should meet the requirements of the temperature-directed strategy (Supplementary Fig. 10). However, reports on the use of thermostable ECH for vanillin biosynthesis are limited. Because *Ct*FCS with excellent thermal stability was obtained, we surmised that *Ct*ECH, encoded by the *CtECH* gene downstream of the *CtFCS* gene in *C. thermodepolymerans*, may be a suitable candidate (Supplementary Fig. 11). Indeed, the *Ct*ECH with a more compact structure was relatively stable at 50 °C, and about 60% of the initial activity was detected after 24 h of incubation (Fig. 3e). Then, the upgraded cell-free dual plasmid system for FCS and ECH expression was constructed; as expected, 106.0 mg l^{-1} of vanillin was obtained without detectable vanillyl alcohol when the reaction temperature increased from 30 °C to 50 °C (Fig. 3c). By further introducing a cofactor recycling module containing an energy extraction cascade, the five-plasmid-mediated thermostable iMECS system synthesized 61.7 mg l^{-1} of vanillin without the addition of cofactors (Fig. 3c). Similar to the pure enzyme reaction, the distribution of ATP to AMP followed the same trend (Supplementary Fig. 6c). Subsequently, under an optimized FCS:ECH:PTA:ACK:PAP loading ratio of 2:3:1:2:2, the highest vanillin productivity reached 1.79 g l^{-1} h^{-1}, and the conversion improved from 44.05% to 98.21% (Fig. 3f and Supplementary Table 2). The pH of the iMECS system remained relatively stable during the catalytic process (Supplementary Fig. 12). Collectively, these results demonstrate the efficiency of the iMECS system in the coenzyme-free biosynthesis of vanillin, and the thermostable cell-free strategy described herein can be widely used for the elimination of by-products.

Triple-cofactor recycling promotes RK biosynthesis

Based on the encouraging results for curcumin and vanillin biosynthesis, we next aimed to demonstrate the applicability of the iMECS strategy in a more complex lignin-to-molecule pathway. We selected the multi-cofactor-dependent pathway from lignin-derived *p*CA to RK as a model, because RK is one of the most expensive natural flavouring components with high demand[32]. The RK pathway can be divided into two modules: an upstream 4-hydroxybenzylidene acetone (4HBA) synthesis module and a downstream reduction module (Fig. 4a). The upstream module consisting of FCS and benzalacetone synthase (BAS) is similar to the curcumin biosynthetic pathway, requiring ATP, CoA and malonyl-CoA as cofactors[32]. The downstream module involves NADPH-dependent RK/zingerone synthase (RZS), which reduces 4HBA to RK via a redox mechanism (Fig. 4a and Supplementary Fig. 13). First, we tested the upstream module separately from the downstream module by externally supplying sufficient cofactors. As FCS exhibits excellent activity towards *p*CA, we screened for BAS that could enhance the 4HBA titre using a cell-free dual plasmid expression system. Strikingly, the *Rp*BAS from *Rhemu palmatum* enabled a 4HBA production of 103.7 ± 4.9 mg l^{-1} from 1 mM *p*CA, which was at least 2.2-fold higher than other BAS (Fig. 4b). Substituting serine 331 with valine in *Rp*BAS has previously been shown to increase its catalytic activity in vivo[14]; however, the mutated *Rp*BAS diminished 4HBA production in the cell-free system (Fig. 4b). Next, we attempted to eliminate the need to add cofactors by incorporating the multi-cofactor recycling modules (MatB and PPK) used in curcumin production. As expected, after the in vitro protein synthesis of FCS, BAS, MatB and PPK, 82.3 ± 5.8 mg l^{-1} of 4HBA was obtained without replenishing cofactors (Fig. 4b), implying that the cofactors were successfully cycled to drive the upstream module.

To construct the downstream module, several RZS from different sources were tested using a CFPS system. *Ec*CurA (where CurA is curcumin reductase) from *E. coli* showed the best activity in 4HBA reduction, and the titre of RK reached 141.8 ± 3.9 mg l^{-1} with sufficient NADPH added (Fig. 4c). Four different NADPH regeneration systems were installed to replace the redox cofactor addition[33], and the formate dehydrogenase (FDH)- and glucose dehydrogenase (GDH)-mediated systems showed better compatibility with the CurA-catalysed route (Fig. 4d). Considering the irreversible reaction and inexpensive sacrificial substrate, we used FDH from *Candida boidinii* with D195Q/Y196R mutations[34] to replenish the pool of NADPH in the cell-free system. The downstream module containing the CurA–FDH dual plasmid system synthesized 113.6 ± 4.0 mg l^{-1} of RK without the addition of NADPH (Fig. 4d). The distribution of NADPH to $NADP^+$ in the FDH-mediated

https://doi.org/10.1038/s44160-025-00819-2

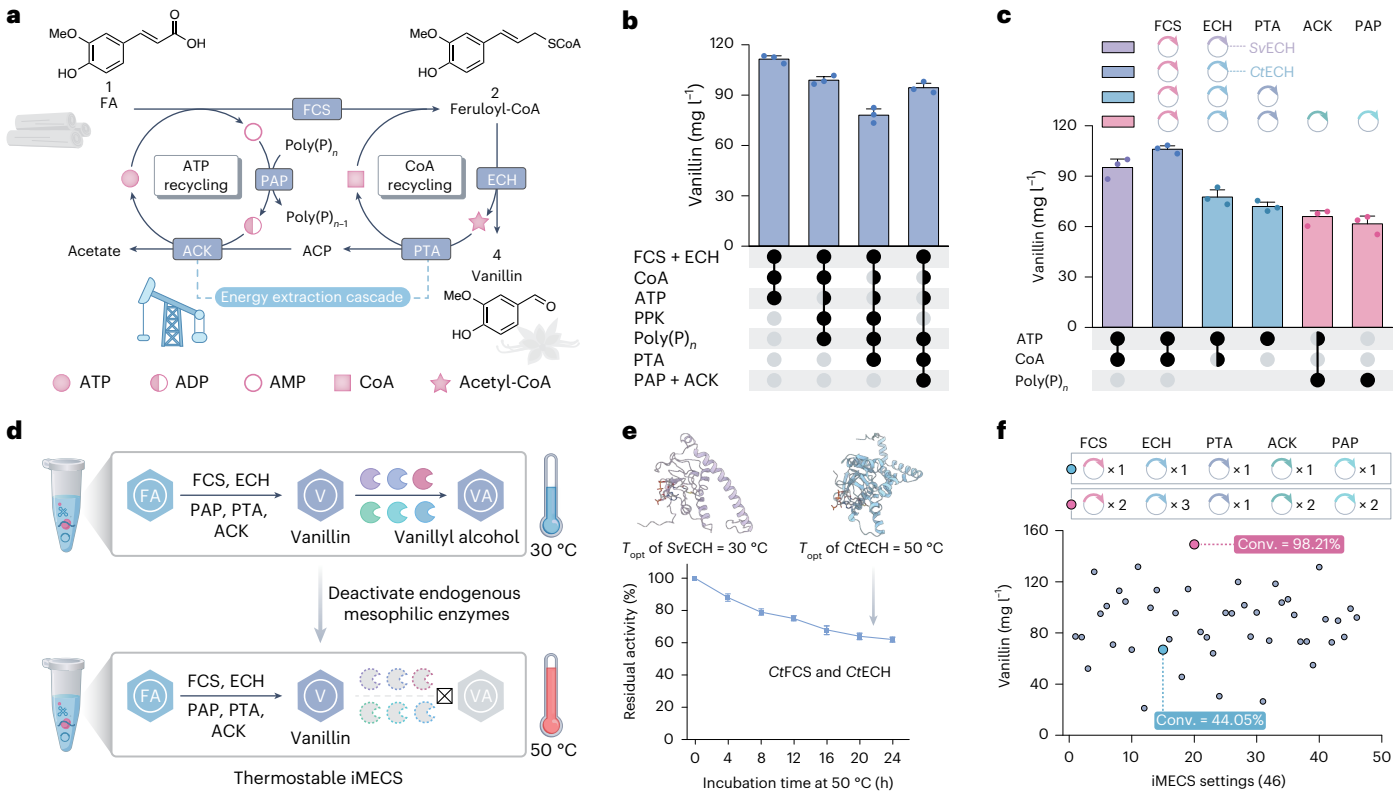

Fig. 3 | Thermostable iMECS for the conversion of FA to vanillin. a, An overview of metabolic pathway for the formation of vanillin based on iMECS. **b**, Synthesis of vanillin using purified enzymes. The black circles refer to excessive addition, while the black semicircles refer to substoichiometric addition (1/10 of the FA; initial concentration of FA was 1 mM). **c**, Vanillin production by CFPS-based biocatalysis. The different coloured bars represent which plasmids were added to different cell-free systems. **d**, A schematic representation of thermostable cell-free strategy to eliminate the redundant activities of endogenous ADHs. FA was converted to vanillin and vanillyl alcohol at 30 °C, while only vanillin was obtained at 50 °C. **e**, The thermal stability of FCS and ECH from *C. thermodepolymerans*. The *Ct*ECH has a more compact structure than *Sv*ECH from *Streptomyces* sp. V-1. Structural simulation analysis of enzymes was performed by AlphaFold2. **f**, Vanillin end-point titres obtained from screening 46 plasmid combinations with different concentration ratios. All plasmids were directly added to the *E. coli* lysate. The initial concentration of FA in these experiments was 1 mM. Vanillin is measured 5 min after the addition of FA. Values are shown as mean ± s.d. ($n = 3$ independent replicates). ACP, acetyl phosphate.

system was lower than in the excess NADPH system (Supplementary Fig. 6d), which may be the reason for its relatively lower RK production.

Next, we set out to establish an iMECS system for the efficient synthesis of RK by integrating the upstream and downstream modules. Based on the entire pathway combination scanning, the optimal six-plasmid-mediated system yielded 118.9 mg l^{-1} of RK without replenishing cofactors (Fig. 4e). Previous studies have demonstrated that fusion enzymes with spatial proximity can promote cascade efficiency[35,36], and we surmise that the fusion of CurA and FDH can accelerate catalytic efficiency and cofactor regeneration. Moreover, the fusion strategy can reduce the complexity of the iMECS system to simplify the process of fine-tuning plasmid addition (Fig. 4f), which is particularly important for systems that lack rapid product detection methods. Here, fusion with the GGGGS linker showed better performance (Supplementary Fig. 14a), and we conducted a new round of combination testing to optimize the plasmid addition ratio. With these efforts, the iMECS system with CurA–FDH fusion in the absence of any additional cofactor displayed a productivity of 0.15 g l^{-1} h^{-1} with a 91.56% conversion (Fig. 4g and Supplementary Table 3), which was 48 times higher than yields reported for other biosynthetic methods (Supplementary Table 3). These results further demonstrate the superiority of the iMECS system, which has great potential for the synthesis of more complex products from lignin-derived aromatics. Notably, the fusion strategy may not always be successful. For example, although the purified FCS–BAS fusion performed better than the separate enzymes (Supplementary Fig. 14b), the RK titre from iMECS was diminished by 8.3-fold when we fused FCS with BAS (Fig. 4g).

Chemoenzymatic synthesis of aromatics from lignin

Lignin-derived FA and *p*CA are key precursors for the biosynthesis of a wide variety of phenylpropanoids with important pharmacological activity but extremely low content in natural sources[37,38]. Having validated the iMECS system for efficiently converting FA and *p*CA, we investigated whether our system could be used as a broad platform to produce other valuable phenylpropanoids. To produce resveratrol, a typical stilbene precursor widely used in pharmaceuticals and nutraceuticals, the BAS in the 4HBA formation module was simply swapped with stilbene synthase (Fig. 5a). Similarly, by swapping BAS with chalcone synthase and chalcone isomerase, naringenin, a gatekeeper flavonoid with diverse biological activities[39], can be synthesized from *p*CA (Fig. 5a and Supplementary Fig. 15). With the cocktails of equal-molar plasmids, we produced 57.9 mg l^{-1} of resveratrol and 122.3 mg l^{-1} of naringenin from 1 mM *p*CA in the absence of additional cofactors (Fig. 5b). Next, by fine-tuning plasmid addition for the precise balancing of enzymatic activity, the iMECS systems yielded 211.3 ± 5.7 mg l^{-1} of resveratrol and 229.1 ± 8.5 mg l^{-1} of naringenin, with conversions of 92.58% and 84.2%, respectively (Fig. 5b and Supplementary Fig. 16). Similar to RK production, the productivities were higher than those reported so far (Supplementary Tables 4 and 5). These results again demonstrate that the loading ratio of plasmids is a key factor in yield optimization,

https://doi.org/10.1038/s44160-025-00819-2

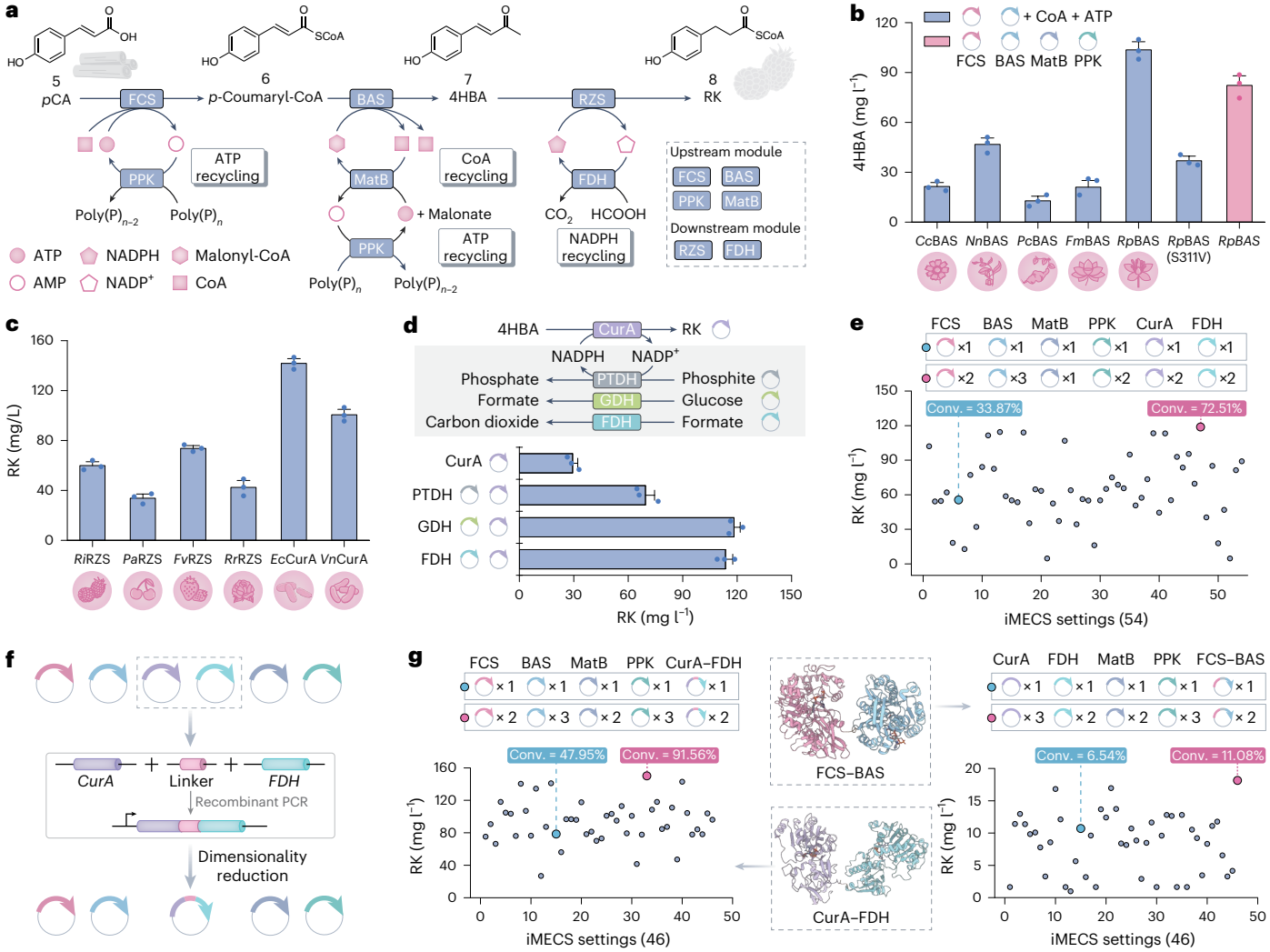

Fig. 4 | iMECS-based biosynthesis of RK with triple-cofactor recycling. **a**, An overview of metabolic pathway for the conversion of *p*CA to RK. **b**, CFPS-based synthesis of 4HBA using BAS from various sources. Values are shown as mean ± s.d. ($n = 3$ independent replicates); dots represent individual data points. **c**, CFPS-based synthesis of RK from 4HBA by using RZS from various sources. Excessive NADPH was added to the reaction system. **d**, Different NADPH regeneration systems were compared for the biosynthesis of RK from 1 mM 4HBA. Values are shown as mean ± s.d. ($n = 3$ independent replicates). **e**, RK end-point titres obtained from screening 54 plasmid combinations with different concentration ratios. RK is measured at 1 h after the addition of *p*CA. **f**, A schematic representation for dimensionality reduction of plasmid combination by fusion expression strategy. Fusion expression was achieved through recombinant PCR, reducing the complexity of the iMECS system by simplifying the process of fine-tuning plasmid addition. **g**, RK end-point titres obtained from screening 46 plasmid combinations with different concentration ratios. All plasmids were directly added to the CFPS system. The initial concentration of *p*CA in these experiments was 1 mM. The fusion of CurA and FDH can increase the yield of RK, while the fusion of FCS and BAS reduces the yield. Structural simulation analysis of the fusions was performed by AlphaFold2. Values are shown as mean ± s.d. ($n = 3$ independent replicates). PTDH, phosphite dehydrogenase.

emphasizing the importance of pathway balancing with iMECS. Owing to its good efficiency and flexibility, the iMECS system can be quickly expanded to the biosynthesis of a series of valuable aromatics through pathway extension and molecular swapping.

With our primary goal of lignin valorization in mind, we investigated whether the iMECS system could be utilized for the conversion of lignin-rich agricultural waste into valuable aromatic compounds. This approach would substantially reduce overall costs, as purification steps are not required for FA and *p*CA. To achieve this, we attempted to develop a hybrid chemical–iMECS route for lignin valorization, which capitalizes on the efficient chemical depolymerization of lignin and the selective biological conversion of aromatic monomers (Fig. 5c). Initially, FA and *p*CA were efficiently released from different agricultural residues via mild alkaline hydrolysis (Supplementary Fig. 17). Next, various value-added chemicals were directly obtained from the neutralized and concentrated hydrolysates (initial concentrations are equivalent to 1 mM FA or *p*CA) using the iMECS system. Some enzymes in these systems exhibited certain promiscuity, potentially converting both FA and CA simultaneously (Supplementary Fig. 18). To avoid by-products and the impact on catalytic efficiency, waste hydrolysates rich in *p*CA or FA were used to synthesize corresponding valuable aromatics, respectively (Fig. 5d). Specifically, FA-rich hydrolysates of corncob, peanut straw, sesame straw, wheat straw and sawdust were suitable for curcumin and vanillin biosynthesis, whereas RK, resveratrol and naringenin were produced from the *p*CA-rich hydrolysates of sugarcane bagasse, spent coffee grounds, soybean straw and rapeseed straw. The titres and productivities of target aromatics were 77.5–173.9 mg l^{-1} and 1.86–36.50 g l^{-1} day^{-1}, respectively, from the depolymerized lignin stream, without detoxifying or washing the pretreated biomass (Fig. 5d). The tolerance to toxic substrates or products is one of the greatest advantages of cell-free systems, and the conversions of these potentially toxic hydrolysates by iMECS could

https://doi.org/10.1038/s44160-025-00819-2

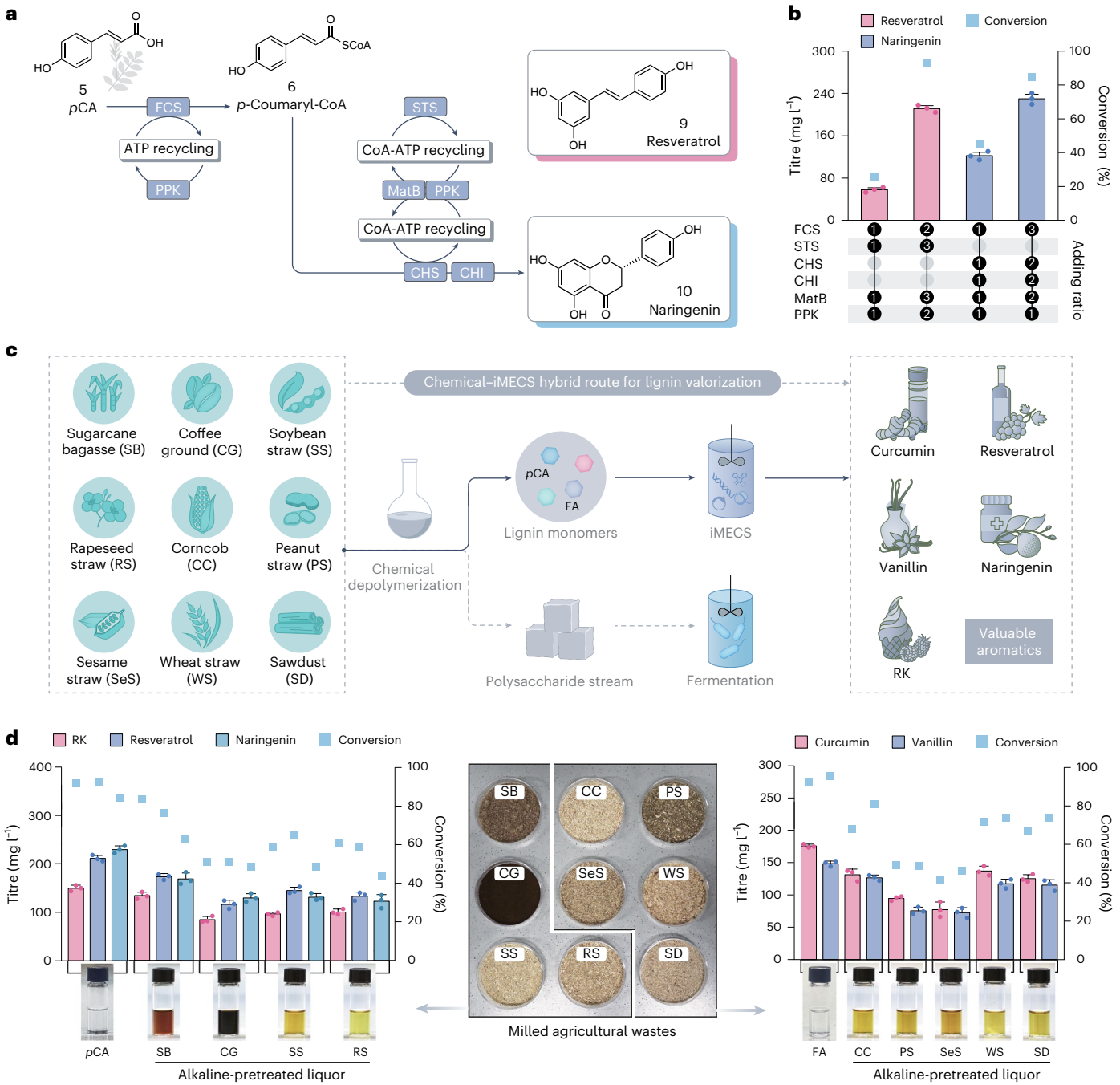

Fig. 5 | Chemical–iMECS hybrid route for lignin valorization. a, An overview of the metabolic pathway for the conversion of *p*CA to resveratrol and naringenin. **b**, The iMECS strategy for in vitro biosynthesis of resveratrol and naringenin from *p*CA. The numbers in the black circle represent the addition ratios of different plasmids. The yield of resveratrol and naringenin obtained by adding plasmids at equal and optimal ratios in the CFPS-based catalytic system is exhibited here. The blue squares represent the corresponding conversion. **c**, A conceptual schematic of the chemical–iMECS hybrid route for lignin-to-molecule conversion. Biomass waste fractionation can yield streams enriched in lignin and polysaccharides, which can be converted along parallel processes. The chemical–iMECS hybrid route combines the chemical depolymerization of lignin and selective biological conversion of lignin-derived aromatics for lignin valorization. **d**, Biosynthesis of valuable aromatics from alkaline-pretreated liquor. The alkaline-pretreated liquor was prepared from various milled agricultural wastes, concentrated and added to the CFPS system. The initial addition of different hydrolysates was equivalent to 1 mM FA or *p*CA. Vanillin is measured 5 min after the catalytic reaction, while other aromatic products are measured 1 h after the catalytic reaction. Values are shown as mean ± s.d. ($n = 3$ independent replicates). STS, stilbene synthase; CHS, chalcone synthase; CHI, chalcone isomerase.

reach up to 83.38% (Fig. 5d). Notably, without any additional cofactors, the productivities achieved from agricultural waste here were 1.23- to 35.89-fold higher than values reported using purified FA and *p*CA as the substrate (Supplementary Tables 1–5). The technology developed in this study is poised to serve as a universal platform for the value-added utilization of other lignin-rich wastes.

Discussion

Lignin valorization into chemicals is key to achieving more sustainable and competitive lignocellulosic biorefineries while also addressing environmental concerns arising from the combustion of lignin-rich waste streams[40,41]. However, the requirement for expensive cofactors and limited conversions seriously hamper the value-added utilization

of lignin[6,10]. We successfully tackled the long-standing challenges using the iMECS system, a broadly applicable strategy developed in this study. By integrating chemical depolymerization with iMECS, lignin-rich agricultural waste can be directly converted into valuable phenylpropanoids without any additional cofactors. Here, the iMECS system established a catalytic bridge with superior atom economy for converting lignin waste into useful aromatic compounds, preserving the aromatic structure of natural resources. The hybrid chemical–iMECS system exhibits a 1.23- to 35.89-fold enhancement in productivity compared with the most efficient previously reported whole-cell routes (Supplementary Tables 1–5). We demonstrated the flexibility of the iMECS platform, through which a wide array of phenylpropanoids could be easily obtained by enzyme swapping and pathway extension. Previous studies have developed interesting and meaningful routes for utilizing lignin, including the use of engineered microorganisms to funnel multiple lignin components into bulk chemicals with large market sizes[42,43]. Although single phenylpropanoids have a relatively small volume compared with the total amount of lignin (approximately 300 billion tonnes), the enormous range of commercially available aromatics collectively represents a multi-billion-dollar market.

In the past decade, rapid advancements in synthetic biology have driven the biosynthesis of complex chemicals. Given the multiple advantages such as self-replication, metabolic diversity, easy recycling and enzyme-protected environment, whole-cell systems (in vivo) harbouring modified metabolic pathways have been developed to produce a variety of valuable chemicals, materials and pharmaceuticals[44–46]. However, the iterative design–build–test cycle of metabolic engineering remains costly and time-consuming, especially for nonlinear pathways with multienzyme combinations[47,48]. For example, the commercial semi-synthesis of artemisinin requires more than 150 person-years, mostly for delicate pathway balancing[49]. Instead, the CFPS-based prototyping platform offers tremendous flexibility to facilitate design–build–test cycles of only a few days[50], and we showed that efficient lignin-to-molecule routes can be quickly established via the iMECS system. As CFPS translates well into an in vivo system[20,21,51], iMECS has the potential to guide the reconstitution of the lignin-to-molecule pathway in vivo. We have also developed several universal strategies for improving the efficiency of the iMECS system. The thermostable cell-free strategy was used to eliminate the by-products of vanillin synthesis, and this strategy can be extended to almost all cell-free systems facing endogenous enzyme problems, thus avoiding complex metabolic remodelling of the chassis. In the future, the rational protein design and biological funnel strategy could be used for the elimination of other by-products caused by the promiscuity of enzymes[42,52,53]. The fusion enzyme strategy was used to improve the catalytic efficiency and reduce the complexity of the iMECS system; however, we do not recommend arbitrary fusion of enzymes that cross-affect other parts of the system. For instance, the RK titre was diminished by 8.3-fold when FCS was fused with BAS. The cofactor regeneration systems were systematically designed here to produce target molecules, and we envision a small library of cofactor recycling modules that can be used in conjunction with a wide variety of cell-free biocatalytic modules.

Notably, the *E. coli*-based cell-free system exhibits linear scalability over a 10^6 change in reaction volume, ranging from laboratory-scale (microlitre volume) to industrial-scale (100 l volume), proving the potential for industrial scale-up applications[54]. However, as an initial proof of concept, the iMECS system will require additional technical developments to achieve industrial applications. Cost-effectiveness may be a key challenge for the industrial scale-up. Recently, numerous efforts have been made to remove cost barriers in cell-free manufacturing, such as using inexpensive reagents, reactor optimization and linear DNA templates[55,56]. Now, we can see that the cost of cell-free systems has decreased by several orders of magnitude in just 10 years, and past and ongoing efforts have already transformed the cell-free system into a more viable platform for biosynthesis[23,57,58]. The biosynthetic potential of iMECS could be further enhanced by replacing conventional high-energy phosphate donors with ubiquitous energy sources (for example, electricity or sunlight) in the future[59,60]. Meanwhile, the permeable membrane bioreactor integrated with pneumatic valves can be used to extend the lifetime of cell-free systems by continuously removing the inhibitory molecules and replenishing essential components[61]. Moreover, membrane bioreactors and hydrogel-based compartments can be used to simplify the in situ product separation[62]. In addition, implementing other cutting-edge technologies, such as industrial automation, metabolic modelling and machine-learning algorithms, in iMECS may greatly improve biomanufacturing efficiency.

In summary, the hybrid chemical–iMECS strategy provides a general coenzyme-free platform for the efficient conversion of lignin into a wide array of valuable aromatics, thereby facilitating the further development of sustainable biorefineries. A key benefit of iMECS-mediated biomanufacturing is that using lignin-rich agricultural waste as feedstock decouples biomanufacturing from sugar prices. In addition, the combustion treatment of lignin waste streams releases a large amount of CO_2. Lignin recycling not only avoids this aspect of carbon emissions but also reduces a substantial portion of greenhouse gas emissions compared with the traditional petroleum-based synthesis of aromatics. With the continuously decreasing cost of CFPS[56,63], the viability of the cell-free production of high-value targets is encouraging. In addition to these initial proof-of-concept demonstrations, we hope that the iMECS-mediated platform will boost the economic feasibility of next-generation biorefineries.

Methods

All experimental procedures and analytical techniques are detailed in the Supplementary Information, which include chemicals, strains and plasmids (Supplementary Method 1); culture conditions and enzyme purification (Supplementary Method 2); enzyme activity assay of *Ct*FCS and *Ct*ECH (Supplementary Method 3); CFPS reactions (Supplementary Method 4); automated iMECS framework (Supplementary Method 5); in vitro bioconversion of lignin monomer (Supplementary Method 6); pretreatment of agricultural wastes (Supplementary Method 7); and analysis and quantification of products and cofactors (Supplementary Method 8).

Reporting summary

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

Data availability

The data supporting the finding of the study are available in the article and its Supplementary Information. Source data are provided with this paper.

References

1. Aggarwal, N. et al. Microbial engineering strategies to utilize waste feedstock for sustainable bioproduction. *Nat. Rev. Bioeng.* **2**, 155–174 (2024).
2. Wrasman, C. J. et al. Catalytic pyrolysis as a platform technology for supporting the circular carbon economy. *Nat. Catal.* **6**, 563–573 (2023).
3. Manker, L. P. et al. Sustainable polyesters via direct functionalization of lignocellulosic sugars. *Nat. Chem.* **14**, 976–984 (2022).
4. Gao, J. et al. Engineering co-utilization of glucose and xylose for chemical overproduction from lignocellulose. *Nat. Chem. Biol.* **19**, 1524–1531 (2023).
5. Kim, J., Nguyen, T. V. T., Kim, Y. H., Hollmann, F. & Park, C. B. Lignin as a multifunctional photocatalyst for solar-powered biocatalytic oxyfunctionalization of C–H bonds. *Nat. Synth.* **1**, 217–226 (2022).

6. Ni, J., Wu, Y. T., Tao, F., Peng, Y. & Xu, P. A coenzyme-free biocatalyst for the value-added utilization of lignin-derived aromatics. *J. Am. Chem. Soc.* **140**, 16001–16005 (2018).
7. Elmore, J. R. et al. Production of itaconic acid from alkali pretreated lignin by dynamic two stage bioconversion. *Nat. Commun.* **12**, 2261 (2021).
8. Werner, A. Z. et al. Lignin conversion to β-ketoadipic acid by *Pseudomonas putida* via metabolic engineering and bioprocess development. *Sci. Adv.* **9**, eadj0053 (2023).
9. Erickson, E. et al. Critical enzyme reactions in aromatic catabolism for microbial lignin conversion. *Nat. Catal.* **5**, 86–98 (2022).
10. Wilkes, R. A. et al. Complex regulation in a *Comamonas* platform for diverse aromatic carbon metabolism. *Nat. Chem. Biol.* **19**, 651–662 (2023).
11. Chen, R. et al. Engineering cofactor supply and recycling to drive phenolic acid biosynthesis in yeast. *Nat. Chem. Biol.* **18**, 520–529 (2022).
12. Galman, J. L., Parmeggiani, F., Seibt, L., Birmingham, W. R. & Turner, N. J. One-pot biocatalytic in vivo methylation-hydroamination of bioderived lignin monomers to generate a key precursor to L-DOPA. *Angew. Chem. Int. Ed.* **61**, e202112855 (2022).
13. Chang, C., Liu, B., Bao, Y., Tao, Y. & Liu, W. Efficient bioconversion of raspberry ketone in *Escherichia coli* using fatty acids feedstocks. *Microb. Cell Fact.* **20**, 68 (2021).
14. Laurel, M., Mojzita, D., Seppänen-Laakso, T., Oksman-Caldentey, K. M. & Rischer, H. Raspberry ketone accumulation in *Nicotiana benthamiana* and *Saccharomyces cerevisiae* by expression of fused pathway genes. *J. Agric. Food Chem.* **71**, 13391–13400 (2023).
15. Wu, J. et al. Metabolic engineering for improved curcumin biosynthesis in *Escherichia coli*. *J. Agric. Food Chem.* **68**, 10772–10779 (2020).
16. Zhan, C. et al. Reprogramming methanol utilization pathways to convert *Saccharomyces cerevisiae* to a synthetic methylotroph. *Nat. Catal.* **6**, 435–450 (2023).
17. Zhang, J. et al. A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature* **609**, 341–347 (2022).
18. Ploessl, D. et al. A repackaged CRISPR platform increases homology-directed repair for yeast engineering. *Nat. Chem. Biol.* **18**, 38–46 (2022).
19. Guo, Y. et al. One-pot biocatalytic synthesis of rac-syringaresinol from a lignin-derived phenol. *ACS Catal.* **13**, 14639–14649 (2023).
20. Karim, A. S. et al. In vitro prototyping and rapid optimization of biosynthetic enzymes for cell design. *Nat. Chem. Biol.* **16**, 912–919 (2020).
21. Vögeli, B. et al. Cell-free prototyping enables implementation of optimized reverse β-oxidation pathways in heterotrophic and autotrophic bacteria. *Nat. Commun.* **13**, 3058 (2022).
22. Yuan, Q. et al. Rapid prototyping enzyme homologs to improve titer of nicotinamide mononucleotide using a strategy combining cell-free protein synthesis with split GFP. *Biotechnol. Bioeng.* **120**, 1133–1146 (2023).
23. Silverman, A. D., Karim, A. S. & Jewett, M. C. Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* **21**, 151–170 (2020).
24. Raduly, F. M., Raditoiu, V., Raditoiu, A. & Purcar, V. Curcumin: modern applications for a versatile additive. *Coatings* **11**, 519 (2021).
25. Beganovic, S. & Wittmann, C. Medical properties, market potential, and microbial production of golden polyketide curcumin for food, biomedical, and cosmetic applications. *Curr. Opin. Biotechnol.* **87**, 103112 (2024).
26. Tan, Z., Clomburg, J. M., Cheong, S., Qian, S. & Gonzalez, R. A polyketoacyl-CoA thiolase-dependent pathway for the synthesis of polyketide backbones. *Nat. Catal.* **3**, 593–603 (2020).
27. Chen, H. & Zhang, Y. H. P. J. Enzymatic regeneration and conservation of ATP: challenges and opportunities. *Crit. Rev. Biotechnol.* **41**, 16–33 (2021).
28. Martău, G. A., Călinoiu, L. F. & Vodnar, D. C. Bio-vanillin: towards a sustainable industrial production. *Trends Food Sci. Tech.* **109**, 579–592 (2021).
29. Ni, J., Gao, Y. Y., Tao, F., Liu, H. Y. & Xu, P. Temperature-directed biocatalysis for the sustainable production of aromatic aldehydes or alcohols. *Angew. Chem. Int. Ed.* **57**, 1214–1217 (2018).
30. Tovilla-Coutiño, D. B., Momany, C. & Eiteman, M. A. Engineered citrate synthase alters acetate accumulation in *Escherichia coli*. *Metab. Eng.* **61**, 171–180 (2020).
31. Li, F. et al. Installing a green engine to drive an enzyme cascade: a light-powered in vitro biosystem for poly(3-hydroxybutyrate) synthesis. *Angew. Chem. Int. Ed.* **61**, e202111054 (2022).
32. Moore, S. J. et al. Refactoring of a synthetic raspberry ketone pathway with EcoFlex. *Microb. Cell Fact.* **20**, 116 (2021).
33. Sharma, V. K., Hutchison, J. M. & Allgeier, A. M. Redox biocatalysis: quantitative comparisons of nicotinamide cofactor regeneration methods. *ChemSusChem* **15**, e202200888 (2022).
34. Wu, W., Zhu, D. & Hua, L. Site-saturation mutagenesis of formate dehydrogenase from *Candida bodinii* creating effective $NADP^+$-dependent FDH enzymes. *J. Mol. Catal. B* **61**, 157–161 (2009).
35. Wang, J. et al. Rational multienzyme architecture design with iMARS. *Cell* **188**, 1349–1362 (2025).
36. Lee, J. et al. Production of Tyrian purple indigoid dye from tryptophan in *Escherichia coli*. *Nat. Chem. Biol.* **17**, 104–112 (2021).
37. Li, C., Yin, L., Wang, J., Zheng, H. & Ni, J. Light-driven biosynthesis of volatile, unstable and photosensitive chemicals from CO_2. *Nat. Synth.* **2**, 960–971 (2023).
38. Brooks, S. M. et al. A tripartite microbial co-culture system for de novo biosynthesis of diverse plant phenylpropanoids. *Nat. Commun.* **14**, 4448 (2023).
39. Ni, J., Tao, F., Wang, Y., Yao, F. & Xu, P. A photoautotrophic platform for the sustainable production of valuable plant natural products from CO_2. *Green Chem.* **18**, 3537–3548 (2016).
40. Subbotina, E. et al. Oxidative cleavage of C–C bonds in lignin. *Nat. Chem.* **13**, 1118–1125 (2021).
41. Wołos, A. et al. Computer-designed repurposing of chemical wastes into drugs. *Nature* **604**, 668–676 (2022).
42. Linger, J. G. et al. Lignin valorization through integrated biological funneling and chemical catalysis. *Proc. Natl Acad. Sci. USA* **111**, 12013–12018 (2014).
43. Lan, H. N. et al. Biological valorization of lignin to flavonoids. *Biotechnol. Adv.* **64**, 108107 (2023).
44. Meng, X. et al. Rational identification of a high catalytic efficiency leucine dehydrogenase and process development for efficient synthesis of L-phenylglycine. *Biotechnol. J.* **18**, 2200465 (2023).
45. Gao, J., Li, Y., Yu, W. & Zhou, Y. J. Rescuing yeast from cell death enables overproduction of fatty acids from sole methanol. *Nat. Metab.* **4**, 932–943 (2022).
46. Zhang, H. et al. Engineering an efficient whole-cell catalyst for D-allulose production from glycerol. *Biotechnol. J.* **18**, 2200600 (2023).
47. Liao, X., Ma, H. & Tang, Y. J. Artificial intelligence: a solution to involution of design–build–test–learn cycle. *Curr. Opin. Biotechnol.* **75**, 102712 (2022).
48. Meng, F. & Ellis, T. The second decade of synthetic biology: 2010–2020. *Nat. Commun.* **11**, 5174 (2020).
49. Radivojević, T., Costello, Z., Workman, K. & Garcia Martin, H. A machine learning automated recommendation tool for synthetic biology. *Nat. Commun.* **11**, 4879 (2020).

50. Dudley, Q. M., Karim, A. S. & Jewett, M. C. Cell-free metabolic engineering: biomanufacturing beyond the cell. *Biotechnol. J.* **10**, 69–82 (2015).
51. Liew, F. E. et al. Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale. *Nat. Biotechnol.* **40**, 335–344 (2022).
52. Peng, F. et al. Improvement of an enzymatic cascade synthesis of nicotinamide mononucleotide *via* protein engineering and reaction-process reinforcement. *Biotechnol. J.* **19**, 2300748 (2024).
53. Lu, L. et al. A bacterial platform for producing aromatic esters from glycerol. *Nat. Chem. Eng.* **1**, 751–764 (2024).
54. Zawada, J. F. et al. Microscale to manufacturing scale-up of cell-free cytokine production-a new approach for shortening protein production development timelines. *Biotechnol. Bioeng.* **108**, 1570–1578 (2011).
55. Smith, S. A. et al. 'Just add small molecules' cell-free protein synthesis: combining DNA template and cell extract preparation into a single fermentation. *Biotechnol. Prog.* **39**, e3332 (2023).
56. Guzman-Chavez, F. et al. Constructing cell-free expression systems for low-cost access. *ACS Synth. Biol.* **11**, 1114–1128 (2022).
57. Heide, C., Ces, O., Polizzi, K. & Kontoravdi, C. Creating cell-free protein synthesis factories. *Pharm. Bioprocess* **6**, 3–6 (2018).
58. Hunt, A. C. et al. Cell-free gene expression: methods and applications. *Chem. Rev.* https://doi.org/10.1021/acs.chemrev.4c00116 (2024).
59. Xie, W. et al. Metal-free reduction of CO_2 to formate using a photochemical organohydride-catalyst recycling strategy. *Nat. Chem.* **15**, 794–802 (2023).
60. Lee, Y. S., Gerulskis, R. & Minteer, S. D. Advances in electrochemical cofactor regeneration: enzymatic and non-enzymatic approaches. *Curr. Opin. Biotechnol.* **73**, 14–21 (2022).
61. Schoborg, J. A., Hodgman, C. E., Anderson, M. J. & Jewett, M. C. Substrate replenishment and byproduct removal improve yeast cell-free protein synthesis. *Biotechnol. J.* **9**, 630–640 (2014).
62. Monck, C., Elani, Y. & Ceroni, F. Cell-free protein synthesis: biomedical applications and future perspectives. *Chem. Eng. Res. Des.* **177**, 653–658 (2022).
63. Warfel, K. F. et al. A low-cost, thermostable, cell-free protein synthesis platform for on-demand production of conjugate vaccines. *ACS Synth. Biol.* **12**, 95–107 (2022).

Acknowledgements

This work was supported by National Natural Science Foundation of China (32071418 to J.N.), National Key Research and Development Program of China (2019YFA0904603 and 2018YFA0903600 to J.N.) and Shanghai Rising-Star Program (20QA1404800 to J.N.). We are grateful to the Y-BioFoundry of LumyBio. We also thank C. Liao at the Shanghai Jiao Tong University for polishing this article.

Author contributions

J.N. conceived the project and supervised the research. L.L., X.W., Y.W., H.L., C.L. and H.Z. performed the experiments and analysed the data. J.W. performed the in silico analysis. J.N. and L.L. wrote the paper with contributions and discussions from all the authors.

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/s44160-025-00819-2.

Correspondence and requests for materials should be addressed to Jun Ni.

Peer review information *Nature Synthesis* thanks Ludmilla Aristilde, Yongjin Zhou and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Thomas West, in collaboration with the *Nature Synthesis* team.

nature portfolio

Corresponding author(s): Jun Ni

Last updated by author(s): Apr 29, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Agilent OpenLab CDS 2.6; Tecan FluentControl; SparkControl 3.2
Data analysis	Agilent OpenLab CDS 2.6; OriginPro 2018C; Microsoft Excel 2208; ParaFold; Pymol 2.5.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The authors declare that all data supporting the findings of this study are available within this paper and its supplementary files.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. Sample size and experimental replicates were chosen based on previous experience and literature reports in the field.
Data exclusions	No data were excluded from the analyses.
Replication	Each experiment was reproduced at least three times independently. All attempts at replication were successful.
Randomization	N/A
Blinding	N/A

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A