
Article:
Webb, J.P., Paiva, A.C., Rossoni, L. et al. (17 more authors) (2022) Multi-omic based production strain improvement (MOBpsi) for bio-manufacturing of toxic chemicals. Metabolic Engineering, 72. pp. 133-149. ISSN 1096-7176

https://doi.org/10.1016/j.ymben.2022.03.004

For Metabolic Engineering

Multi-omic Based Production Strain Improvement (MOB*psi*) for Bio-manufacturing of Toxic Chemicals

Joseph P Webb[a], Ana Carolina Paiva[b], Luca Rossoni[b], Amias Alstrom-Moore[a], Vicki Springthorpe[c], Sophie Vaud[b], Vivien Yeh[d], David-Paul Minde[e], Sven Langer[c], Heather Walker[f], Andrea Hounslow[a], David R Nielsen[g], Tony Larson[c], Kathryn Lilley[e], Gill Stephens[b], Gavin H Thomas[c], Boyan B Bonev[d], David J Kelly[a*], Alex Conradie[b**] & Jeffrey Green[a*]**

[a]*School of Biosciences, University of Sheffield, Sheffield S10 2TN, UK.*

[b]*Sustainable Process Technologies, Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, UK.*

[c]*Department of Biology, University of York, York YO10 5DD, UK.*

[d]*School of Life Sciences, University of Nottingham, QMC, Nottingham NG7 2UH, UK.*

[e]*Cambridge Centre for Proteomics, Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK.*

[f]*biOMICS Mass Spectrometry Facility, Department of Animal and Plant Sciences, University of Sheffield, Sheffield, S10 2TN, UK.*

[g]*Chemical Engineering, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, AZ 85287-6106, USA.*

* Corresponding author. Tel.: +44 (0)114 222 4403; E-mail: jeff.green@sheffield.ac.uk

** Corresponding author. Tel.: +44 (0)115 748 6451; E-mail: alex.conradie@nottingham.ac.uk

*** Corresponding author. Tel.: +44 (0)114 222 4414; E-mail: d.kelly@sheffield.ac.uk

Keywords: biotechnology/cell factory/lipidomics/proteomics/styrene/transcriptomics

ABSTRACT

Robust systematic approaches for the metabolic engineering of cell factories remain elusive. The available models for predicting phenotypical responses and mechanisms are incomplete, particularly within the context of compound toxicity that can be a significant impediment to achieving high yields of a target product. This study describes a Multi-Omic Based Production Strain Improvement (MOB*psi*) strategy that is distinguished by integrated time-resolved systems analyses of fed-batch fermentations. As a case study, MOB*psi* was applied to improve the performance of an *Escherichia coli* cell factory producing the commodity chemical styrene. Styrene can be bio-manufactured from phenylalanine *via* an engineered pathway composed of the enzymes phenylalanine ammonia lyase and ferulic acid decarboxylase. The toxicity, hydrophobicity, and volatility of styrene combine to make bio-production challenging. Previous attempts to create styrene tolerant *E. coli* strains by targeted genetic interventions have met with modest success. Application of MOB*psi* identified new potential targets for improving performance, resulting in two host strains (*E. coli* NST74Δ*aaeA* and NST74Δ*aaeA* $cpxP_o$) with increased styrene production. The best performing re-engineered chassis, NST74Δ*aaeA* $cpxP_o$, produced ~3× more styrene and exhibited increased viability in fed-batch fermentations. Thus, this case study demonstrates the utility of MOB*psi* as a systematic tool for improving the bio-manufacturing of toxic chemicals.

1. Introduction

The climate emergency and the changing attitudes of consumers to products deemed environmentally damaging necessitate a paradigm shift in manufacturing from fossil reserve feedstocks to circular economies utilizing renewable feedstocks. Biological production

platforms (cell factories) are a promising means to convert renewable feedstocks to chemical products with high selectivity. A major barrier to the adoption of such fermentation processes is product toxicity, resulting in productivities and yields that are well below the values required for commercial application. Many desirable chemicals poison the biological production platform, which is often a bacterium or yeast. One approach to improve cell factory performance is to add the toxic product to cultures and isolate/select more tolerant strains (reviewed by Zingaro et al., 2013). However, this approach ignores a fundamental property of many bioprocesses, i.e. the product is synthesized internally (within the cell factory), not externally. Hence, the conventional approach risks missing interventions to enhance cell factory performance, by failing to consider the effects of the toxic product and pathway intermediates that are synthesized within the cell. Moreover, the flask cultures that are often used in laboratory studies are not always scalable.

Here we introduce a different approach, Multi-Omic Based Production Strain Improvement (MOB*psi*). MOB*psi* involves the acquisition of time-resolved, multi-omic and physiological measurements from scalable fermentations. This approach requires comparisons between controls (expressing an inactivated production pathway in the presence and absence of externally added chemical) and the strain producing the chemical. The data obtained are used to populate an integrated analytical interface (Multi-Omics Research Factory, MORF; Springthorpe et al., 2020; https://morf-db.org/projects/DETOX/styrene), where all the time-resolved fermentation and multi-omic (transcriptomic, proteomic, metabolomic and lipidomic) data sets can be accessed, permitting objective analysis of the cell factory's response to chemical biosynthesis. This information is used to identify potential interventions to improve performance within the context of the chemical's toxicity profile. Cell factory re-engineering (gene deletion and/or over-expression) followed by performance testing in small-scale cultures, identifies the most promising candidates, which can be

76 combined to further enhance productivity. The best performing strains are then characterized
77 in scalable fed-batch fermentations to close the MOB*psi* methodology.

78 As a case study of the MOB*psi* approach, we describe its application to enhance the
79 performance of *E. coli* as a cell factory for styrene production. Styrene is a commodity
80 chemical, used in the manufacture of polystyrene, with an annual global consumption of ~25
81 million tons (Lian et al., 2016). Current styrene production relies on chemo-catalytic
82 dehydrogenation of fossil-derived ethylbenzene in an energy intensive process (Wu et al.,
83 1981; Lian et al., 2016). Alternatively, bio-production of styrene from biomass-derived
84 sugars has been demonstrated using an *E. coli* cell factory. In this case, strain NST74, which
85 over-produces phenylalanine, was transformed by introducing two non-native enzymes,
86 phenylalanine ammonia lyase (Pal2 from *Arabidopsis thaliana*) and ferulic acid
87 decarboxylase (Fdc1 from *Saccharomyces cerevisiae*) (Fig. S1; McKenna and Nielsen,
88 2011). Conversion of phenylalanine to *trans*-cinnamic acid (Pal2) and subsequent
89 decarboxylation (Fdc1) produces styrene.

90 Styrene bio-production represents a suitable MOB*psi* case study, because although
91 previous attempts to increase styrene bio-production by targeted genetic engineering of *E.*
92 *coli* have been made, further improvements are necessary to reach commercially feasible
93 values. The original flask cultures of NST74 transformed with *pal2-fdc1* produced 0.26 g L^{-1}
94 styrene (McKenna and Nielsen, 2011). Later, an *E. coli* BL21(DE3) cell factory expressing a
95 plasmid encoded, codon optimized, Pal2 produced ~0.06 g L^{-1} styrene and plasmid
96 optimization increased this value to ~0.10 g L^{-1} (Liu et al., 2018). Further genetic
97 interventions to promote carbon flow to L-phenylalanine (over-expression of *aroF* and *pheA*)
98 increased styrene production to ~0.21 g L^{-1}, and when combined with over-expression of
99 *tktA*, and *ppsA*, ~0.28 g L^{-1} styrene was produced; a ~5-fold increase compared to the starting
100 strain (Liu et al., 2018). Another strain of *E. coli* that over-produces phenylalanine (YHP05)

was also genetically modified to produce styrene (Lee et al., 2019). In this case optimizing expression of *pal2* resulted in a 2-fold increase in styrene production (to ~0.25 g L^{-1}). Heterologous expression of the *Pseudomonas aeruginosa cti* gene to increase membrane rigidity by increasing the amount *trans*-unsaturated fatty acyl chains incorporated into *E. coli* MG1655 cell membranes improved styrene production by ~10% (Tan et al., 2016). Furthermore, expression of *pssA*, (coding for phosphatidylserine synthase) increased the amount of phosphoethanolamine headgroups in *E. coli* cell membranes, as well as other changes, which increased styrene tolerance by 16% (Tan et al., 2017). Liang et al. (2020) constructed a library of >85,000 variants of 54 *E. coli* transcription factors, which was then screened for increased styrene production and tolerance. The best performing strain increased styrene production ~3.5-fold compared to the starting strain by expressing a variant of LexA, a repressor of the DNA damage SOS response. When combined with process engineering improvements, such as medium optimization, gas stripping and *in situ* product removal by addition of solvents to cultures (McKenna et al., 2015; Lee et al., 2019), further genetic engineering of *E. coli* has the potential to create cell factories that move styrene bio-manufacturing a step closer to reality.

We recently showed that the transcriptional responses of an *E. coli* cell factory producing styrene differed from those observed when styrene was added externally (Machas et al., 2021). These single time point, flask culture studies, alongside the previous attempts to improve styrene yields described above, suggested that styrene bio-manufacturing was an appropriate case study to test the potential power of MOB*psi*'s time-resolved systems analyses of scaled-down fed-batch fermentations to identify new genetic interventions that improve bio-manufacturing of toxic chemicals.

2. Results and discussion

126 Figure 1 shows an overview of MOB*psi* in the form of a decision tree to illustrate the
127 systematic nature and underpinning logic of the process. As noted above styrene bio-
128 production was deemed to be a suitable test case for MOB*psi*, and therefore the sub-sections
129 below relate to the steps shown in the decision tree aimed at identifying new genetic
130 interventions that enhance the production of this toxic commodity chemical.
131
132 *2.1. Styrene toxicity - partitioning of styrene into the membrane of E. coli cells*
133 The maximum inhibitory concentration of styrene against *E. coli* is ~250 mg L^{-1}
134 (McKenna and Nielsen, 2011) and exposure to ~100 mg L^{-1} styrene for 2 h resulted in a 2-
135 fold decrease in growth (Mingardon et al., 2015). Cell membrane dysfunction has been
136 strongly linked to styrene toxicity (Lian et al., 2016) and as a hydrophobic molecule (*logP* =
137 2.95), styrene would be expected to partition into the hydrophobic interior of biological
138 membranes. To better understand the potential effects of styrene on *E. coli* cell membranes a
139 200 ns, all-atom molecular dynamics (MD) simulation, consisting of a 100 × 100 Å model
140 lipid bilayer patch with 100 randomly distributed styrene molecules in the aqueous phase was
141 constructed. The MD simulation revealed a rapid phase separation of styrene (within 50 ns)
142 into a solvent sub-phase, which was in fast exchange with the aqueous phase (Fig. 2a; Movie
143 S1). This was accompanied by slower segregation of styrene into the hydrophobic interior of
144 the bilayer, which was complete by the end of the 200 ns evolution trajectory. Within the
145 hydrophobic region of the bilayer styrene dynamics was slower and no styrene molecules
146 were observed to return to the aqueous phase.
147 The ordering effects of the fatty acyl chains within lipid membranes severely restrict
148 styrene dynamics compared to free styrene in solution. This permits selective excitation and
149 observation of molecules constrained within the bilayer by high resolution ^{13}C cross-
150 polarization magic angle spinning (CP-MAS) NMR, as previously demonstrated for butyl

151 methacrylate membrane interactions by Yeh et al. (2020). Natural abundance ^{13}C CP-MAS
152 NMR of *E. coli* NST74 lipid extracts after addition of styrene showed its presence in a
153 phospholipid bilayer (Fig. 2b). Thus, having demonstrated that styrene could be detected in
154 NST74 lipid extracts, 7,8-^{13}C- styrene was added to NST74 cells. Samples were collected
155 immediately after styrene addition, as well as 20 min, 2 h and 24 h post-addition. High
156 resolution ^{13}C solid state MAS NMR spectra showed rapid incorporation of 7,8-^{13}C-styrene
157 into the membranes of the *E. coli* NST74 cells immediately after styrene addition (Fig. 2c).
158 However, within 20 min of exposure, ^{13}C-styrene was no longer observed within the
159 membranes of the cells, suggesting that, unlike the lipid extracts, NST74 cells did not retain
160 styrene within the hydrophobic interior of their membranes. Thus, a better understanding of
161 the process of adaptation to styrene stress might permit the identification of genetic
162 interventions that augment these responses and enhance cell factory performance.
163
164 *2.2. Scaled-down fed-batch cultures for time-resolved multi-omic and phenotypic analyses*
165 The next MOB*psi* step was to establish fed-batch fermentations that would permit the
166 collection of time-resolved multi-omic datasets. To determine the how *E. coli* adapted during
167 the production of styrene, the phenylalanine over-producing strain *E. coli* NST74
168 (ATCC31884) transformed with pGS2596, which permits expression of *pal2-fdc1* (Fig. S1)
169 from an arabinose-inducible promoter was used (Fig. 3a). Control fermentations in the
170 absence (negative control, Fig. 3b) and presence (external addition control, Fig. 3c) of
171 externally added styrene were carried out using the same host background (NST74)
172 transformed by pGS2597, which coded for Pal2 and Fdc1 proteins that were inactivated by
173 the introduction of mis-sense mutations in their active sites (Pal2: Y109F, S204A; Fdc1:
174 Q192A, H193A). For the external addition control (Fig. 3c), styrene was added to the
175 fermentation vessel post-induction using a feed rate profile that mimicked the accumulation

of styrene determined experimentally during bio-production (Fig. 3a), thereby approximating the growth/viability phenotypes of the production strain. Thus, the robust cultivation strategy ensured that the three biological replicates yielded consistent, time-resolved, multi-omic datasets.

During the first stage of the fermentations (pre-induction of *pal2-fdc1*), phenylalanine accumulated in the culture medium (up to 1.7 g L^{-1}). After induction of *pal2-fdc1* in the production fermentations, net consumption of phenylalanine was observed and its concentration in the medium decreased to 0.4 g L^{-1} (Fig. 3a). Six hours post-induction, 0.12 g L^{-1} styrene and 0.14 g L^{-1} of the precursor, *trans*-cinnamic acid, were detected in the culture supernatant (Fig. 3a). Between 2 h and 6 h post-induction, the ability of these styrene-producing bacteria to fully utilize the glucose feed was impaired and this coincided with cessation of growth, a decrease in viability and excretion of acetate (up to 5.4 g L^{-1}). Thus, the test fermentation produced sufficient styrene to exert toxic effects that should invoke systems level adaptive responses. In contrast to the production fermentations, the negative control continued to grow after induction of the inactivated Pal2 and Fdc1 enzymes and phenylalanine continued to accumulate (up to 3.3 g L^{-1}) (Fig. 3b). The external styrene control stopped growing 4 h after exposure to externally added styrene, and, similar to the production strain, cell viability and glucose utilization decreased, phenylalanine production ceased, and acetate (up to 1.8 g L^{-1}) accumulated in the medium (Fig. 3c). These data indicated that viability of the NST74 host was compromised when styrene was produced intracellularly or added extracellularly and consequently uptake of phenylalanine synthesized prior to *pal2-fdc1* induction and further conversion of glucose to phenylalanine rapidly became sub-optimal.

2.3. Time-resolved multi-omic datasets

Following the establishment of appropriate control and test fed-batch fermentations, the next MOB*psi* step was to collect time-resolved multi-omic datasets. For the control with externally added styrene, global gene expression was analyzed 2, 4 and 6 h after styrene addition and induction of the inactivated Pal2 and Fdc1 proteins and compared with the pre-induction samples (0 h). After 2 h no genes met the significance criteria (≥2-fold; false discovery rate (FDR) adjusted $p \leq 0.05$) and only one (*azuC*) was was differentially regulated 4 h post-induction (Table S1). However, 6 h post-induction there were 31 up-regulated and 53 down-regulated genes, coinciding with cessation of growth and decreased cell viability (Table S1). In contrast, the production strain (NST74 pGS2596 producing styrene internally) exhibited an earlier, more progressive response to styrene. Compared to the pre-induction sample 9, 16 and 24 genes were differentially regulated in the 2 h, 4 h and 6 h post-induction samples, respectively (Table S1). Six hours post-induction only three of the up-regulated genes (*azuC*, *norW* and *yjgX*) and two of the down-regulated genes (*seqA* and *srlB*) exhibited common responses when comparing intracellular styrene synthesis to extracellular styrene addition. Although not phenotypically evident at the macro-scale of fermentation (Fig. 3), the transcriptome comparison suggested that NST74 experiences very different challenges when producing styrene within the cell compared to external addition of styrene to the culture (Table S1).

Having established that internally produced and externally applied styrene invoked different transcriptional responses in our scaled down fed-batch fermentations, further 'omic analyses of the production and control fermentations were undertaken, i.e. proteomics, metabolomics and lipidomics. Comparison of the transcriptomes and proteomes indicated that 18 operons (13 of which did not respond to external styrene) and 114 proteins were differentially regulated in response to styrene production (Tables S2 and S3).

Principal component analysis of the lipidomes of control and production strains indicated that separation between the production and control fermentations arose from relatively small changes (<4% contribution) in several lipids (Fig. S2; Table S4). In positive ion and negative ion modes 151 and 101 masstags were obtained, respectively. In both modes the majority of the most abundant lipids did not change in response to styrene production (Table S4). However, the abundances of phosphatidylglycerol (PG) lipids, especially PG(16:0_16:1), PG(16:0_18:1) and PG(16:1:_18:1) (16:1 indicates a 16 carbon fatty acyl chain with one double bond) were greater in the styrene producing cultures compared to the controls. Notably, this occurred before induction of *pal2-fdc1*, suggesting that low level styrene production, arising from leaky expression of *pal2-fdc1*, was sufficient to invoke these changes. Total fatty acid analysis (Table S5) revealed a small increase in the degree of fatty acyl chain saturation in the styrene producing strain, suggesting a shift towards lower membrane fluidity.

Examination of metabolome datasets for negative and positive mode operation indicated that the abundances of 11 masses (metabolites) were increased in response to styrene production (Table S6). The products of the breakdown of cyclic-di-GMP (GMP, 363.06 Da and pGpG, 704.8 Da) increased in abundance during styrene production, suggesting increased phosphodiesterase activity, with possible implications for cyclic-di-GMP signaling pathways, which control the sessile-planktonic lifestyle switch (Hengge et al., 2016). The increased abundance of *S*-lactoylglutathione (378.9 Da) was consistent with decreased abundance of glutathione (GSH; 307.0 Da, ~10-fold lower in the 6 h post-induction sample for both ion modes compared to the pre-induction sample), which together suggest that methylglyoxal levels increase during styrene production. Four differentially abundant species did not match metabolites in the *E. coli* metabolite database (Table S6; ECMDB; Guo et al., 2012; Sajed et al., 2016).

250

251 *2.4. Selection of potential genetic interventions*

252 The next step in the MOB*psi* process (Fig. 1) was to input the multi-omic datasets into
253 an integrated database, MORF (Springthorpe et al., 2020; https://morf-
254 db.org/projects/DETOX/styrene) to facilitate identification of genetic interventions to
255 enhance cell factory performance. It should be noted that the MOB*psi* approach considers
256 responses within, as well as between, test and control fermentations. Thus, transcriptional
257 changes associated with generic fermentation stresses, e.g. gene expression 2 h, 4 h, and 6 h
258 post-induction of *pal2-fdc1* were compared to the pre-induction sample (0 h) (Table S1),
259 along with changes specifically linked to styrene production by comparing samples taken 2 h,
260 4 h and 6 h post-induction of production fermentations with the corresponding control
261 samples (Table S2). Styrene is well known to impair membrane function (see above; Lian et
262 al., 2016); therefore, genes associated with cell envelope function and others previously
263 linked with increased tolerance of toxic chemicals were prioritized in the selection process.

264 Genetic intervention candidates were categorized as follows: Class 1 candidates were
265 genes that exhibited significant differential transcriptional regulation (≥2-fold change
266 adjusted FDR p<0.05 in at least one sample) in response to styrene (Table S2); Class 2
267 candidates were proteins exhibiting differential abundance in response to styrene production
268 compared to the non-producing control (≥2-fold change adjusted FDR p<0.05 in at least one
269 sample) (Table S3); Class 3 candidates were those that exhibited a positively correlated
270 significant regulation in both transcriptome and proteome in response to styrene, i.e. a subset
271 of Class 1 and Class 2 candidates; Class 4 candidates were genes/proteins that did not exhibit
272 differential regulation, but potentially contributed to the observed changes in membrane lipid
273 composition or metabolite concentrations; Class 5 candidates were transcription factors with

altered activity in response to styrene as predicted from the transcriptomic data by TFInfer (Asif et al., 2010).

There were eight Class 1 candidates: *aaeXAB*, *azuC*, *cpxP*, *dtpD*, *pspABCDE*, *rmf*, *spy, yjgX* (Fig. 4a). The up-regulated *aaeXAB* operon codes for an aromatic carboxylic acid efflux pump that is known to be induced in response to *trans*-cinnamic acid (Van Dyk et al., 2004) an intermediate of the styrene production pathway (Fig. S1), which was observed to accumulate during styrene production (Fig. 3a). IsrB (*azuC*) is a sRNA of unknown function that contains an open reading frame coding for the small inner membrane-associated protein AzuC and was significantly up-regulated in response to external and internal styrene (Table S1). AzuC exhibits increased abundance when *E. coli* cultures are exposed to various stresses including low pH, heat, peroxide or diamide (Hemm et al., 2010). The *cpxP* and *spy* genes code for related periplasmic chaperones. CpxP is involved in resistance to extra-cytoplasmic stresses, while the expression of *spy* is induced by externally added protein denaturants such as butanol, ethanol and tannic acid, (Quan et al., 2011; Vogt et al., 2011). The *dtpD* gene codes for a peptide transporter and was up-regulated when *E. coli* was exposed to the volatile solvents N-cyclohexyl-pyrrolidone or cyclopentanone (Yung et al., 2016). The products of the *psp* operon are involved in maintaining cell membrane function in response to various stresses and over-expression previously enhanced the survival of *E. coli* in the presence of externally added *n*-hexane (Joly et al., 2010; Kobayashi et al., 1998). The inactivation of 70S ribosomes by ribosome modulation factor (Rmf) -mediated dimerization is required for long-term survival under stress and hence expression of *rmf* is generally negatively correlated with expression of core ribosomal proteins (Yamagishi et al., 1993), and so Rmf could play a role in styrene tolerance, although not specifically associated with cell envelope function. The *yjgX* transcript is a likely pseudogene coding for the N-terminal fragment of a putative

298 phosphoethanolamine transferase because of the insertion of a KpLE2 phage-like element in
299 the 3´ region of the gene (Keseler et al., 2021).

300 Seventeen Class 2 candidates were selected from the proteomic data set (Fig. 4b). Six
301 chaperone proteins (CpxP, CspC, HslO, IbpA/B and Spy) exhibited increased abundance
302 indicating that the cell factory was stressed. Of these the *cpxP* and *spy* genes were
303 significantly up-regulated during styrene production (see above). CspC is a component of a
304 network of proteins that facilitate stress-induced mutagenesis in *E. coli* (Al Mamun et al.,
305 2012); HslO binds unfolded proteins and is activated by intramolecular disulfide bond
306 formation under oxidative stress (Jakob et al., 1999); and IbpA/B function alongside ClpB
307 and DnaK to resolve protein aggregates, such as inclusion bodies (Mogk et al., 2003). The
308 outer membrane porin, OmpF, increased in abundance (as did OmpC). Solvent tolerant
309 strains of *E. coli* exhibited lower levels of OmpF, but similar amounts of OmpC, compared to
310 the parent strain (Aono and Kobayashi, 1997). Thus, it was reasoned that the increased
311 abundance of OmpF might contribute to styrene sensitivity. Several proteins encoded by the
312 up-regulated *psp* operon (see above) also exhibited increased abundance in response to
313 styrene production. A progressive decrease in the abundances of 15 ribosomal proteins was
314 consistent with a switch to 'survival mode' after induction of styrene production, but the
315 surprising down-regulation of *rmf* gene expression (see above) was confirmed by the lower
316 abundance of Rmf protein. Eight proteins of unknown function involved in stress
317 responses/membrane function were differentially regulated during styrene production. YbfA
318 possesses a DUF2517 domain, and *ybfA* deletion mutants are more sensitive to radiation
319 damage (Sargentini et al., 2016). YbrL is a protein kinase domain protein that facilitates
320 cross-talk between metal-responsive two-component systems. YciG is a stress-induced
321 acidophilic repeat motifs-containing protein. The YdcH and YhcN proteins are implicated in
322 the response to oxidative stress, with *ydcH* being up-regulated in response to peroxide (Zheng

et al., 2001) and YhcN being involved in resistance to peroxides and in tolerating epoxides (Lee et al., 2010). The expression of the gene coding for the inner membrane protein YebE is controlled by the extracytoplasmic stress regulator CpxR (Bury-Moné et al., 2009) and YgdR is a lipoprotein that is induced in response to cell wall damage (Laubacher and Ades, 2008). The YghW protein has been linked to tolerance of externally added butanol (Jarboe, 2018; Si et al., 2016). It is noted that the *yhcN* gene and the corresponding protein were up-regulated, although the former did not meet the significance criteria, and so comparison of the Class 2 candidates with those of Class 1 produced four Class 3 candidates (*cpxP*, *pspABCDE, rmf*, *spy*).

Class 4 genes were selected from the lipidomic and metabolic datasets. The lipidomics indicated small changes in lipid saturation (Tables S4 and S5), suggesting a fluidity change to control and regulate membrane permeability (Murzyn et al., 2005). This suggested that under- or over-expression of the essential β-hydroxyacyl-ACP dehydratases, FabA and FabZ, responsible for the synthesis of unsaturated fatty acids, might affect styrene productivity (Zhang and Rock, 2008). Moreover, heterologous expression of *cti*, which increases the proportion of *trans*-unsaturated fatty acyl chains to increase membrane rigidity enhanced styrene tolerance (Tan et al., 2016). The increased abundance of some unsaturated PG lipids (Table S4) was unexpected because Tan et al. (2017) reported tolerance to externally applied styrene was enhanced by over-expression of *pssA*, which increased the ratio of phosphatidylethanolamine:PG (PE:PG) lipids >2-fold. Thus, *cti* and *pssA* were also included in the list of possible genetic interventions.

Class 5 candidates sought to include the role of regulatory proteins. Whilst the expression of regulatory genes might not change, the corresponding proteins might respond to the stresses associated with styrene production to regulate other genes/proteins. Therefore, TFInfer software (Asif et al., 2010) was used to analyze changes in the activities of

transcription factors in response to styrene production. This analysis supported the involvement of the *pspA-E* and *spy* operons in the response to styrene production as the activities of the regulators PspF and BaeR increased (Fig. S3). Furthermore, the predicted changes in the activity of EnvY could contribute to the up-regulation of OmpC and OmpF (outer membrane porins) observed in the proteome (Table S3; Lundrigan and Earhart, 1984). Since members of the PspF, BaeR and EnvY regulons were represented in other classes in this test case further investigation of these potential Class 5 candidates was deferred according to the MOB*psi* protocol (Fig. 1). Further investigation of the predicted changes in UvrY activity (Fig. S3a) were also deferred because, although the *uvrY* variants were identified under styrene selection, this locus was not significantly enriched after two rounds of selection (Liang et al., 2020). In contrast, to the findings of Liang et al (2020) TFInfer analysis did not implicate LexA in styrene tolerance.

2.5. Small-scale screening of NST74 derivatives for improved styrene production

Having identified 25 candidate genes (including *cti* and *pssA*, see above) for improving styrene production (Fig. 5; Table S7) the next stage of MOB*psi* was to test deletion mutants, and over-expression derivatives of *E. coli* NST74 pGS2596 (production strains) (Fig. 6) and *E. coli* NST74 pGS2597 (control strains) in a flask culture-based small-scale screen. Both over-expression and deletion strains transformed by the control plasmid pGS2597 produced no styrene. Over-expression and deletion strains for each candidate gene were tested because the response to styrene production stress could include up-regulation of genes/proteins that limit toxicity, e.g. by efflux of a pathway intermediate (in which case gene deletion might be beneficial) or by efflux of the product (in which case over-expression might be beneficial). Furthermore, the response might include the down-regulation of genes/proteins that constrain styrene production, e.g. by lowering precursor production (in which case over-expression

might be beneficial), or promote survival by altering the properties of, for example, the cell membrane (in which case deletion might be beneficial). Strains were considered for further investigation if the styrene concentration normalized to culture OD_{600} was significantly higher than the corresponding control strain for one intervention with decreased or no improvement for the opposite intervention (e.g. styrene production increased for gene over-expression and decreased when the same gene was deleted). The *aaeA* mutant and those transformed with plasmids expressing *cpxP*, *rmf* or *yebE* met these criteria (Fig. 6). It was noted that both over-expression and deletion of the *hslO* gene increased specific styrene production and therefore did not match the criteria for further investigation (Fig. 6).

2.6. Small scale screening of styrene production by strains with combinations of genetic interventions

The next step in the MOBpsi process was to combine the best performing deletion mutant (*aaeA*) with the best over-expression plasmids (i.e. NST74*ΔaaeA cpxP*$_o$, *ΔaaeA rmf*$_o$, or *ΔaaeA yebE*$_o$ strains). In addition, a NST74 derivative with an heterologous *cti* gene and an additional *pssA* gene integrated into the chromosome at the *ldhA* locus (NST74*ΔaaeA ΔldhA::cti-pssA*) was constructed to complete the Class 4 tests. In the flask-based assay, of the NST74*ΔaaeA rmf*$_o$, NST74*ΔaaeA yebE*$_o$ and NST74*ΔaaeA ΔldhA::cti-pssA* strains produced less styrene than the NST74 *ΔaaeA* strain, and thus these strains were not taken forward to the next stage of MOB*psi* (Fig. 7a). Specific styrene production for flask cultures of NST74*ΔaaeA* and NST74*ΔaaeA cpxP*$_o$ were similar (Fig. 7a), but the lower concentration of phenylalanine in cultures of the latter strain suggested better conversion of the precursor to product (Fig. 7b), and therefore both these strains were taken forward to the next stage of MOB*psi*.

398 *2.7. Fed-batch fermentation of re-engineered strains*

399 The systematic MOB*psi* approach for cell factory improvement (Fig. 1) was closed by
400 subjecting four strains (NST74 pGS2597, negative control; NST74 pGS2596, starting strain;
401 and the two re-engineered strains, NST74*ΔaaeA* pGS2596 and NST74*ΔaaeA* $cpxP_o$
402 pGS2596) to single fed-batch fermentations using the same cultivation strategy (see
403 *Materials and methods*). To assess whether the re-engineered strains were more productive,
404 styrene was measured in the culture samples and in two dodecane traps, connected in series to
405 the fermenter off-gas outlet, to absorb styrene stripped from the broth. Styrene was not
406 detected in the negative control fermentations. The statistical significance of the differences
407 in total styrene synthesis by the production strains was assessed by creating radial basis
408 function models for each fermentation time course of the cumulative total styrene. These
409 radial basis function models include lower and upper 95% confidence limits (Fig. 8a and b),
410 i.e. for a particular calculated total styrene value serving as the observed measurement, any
411 value within the lower and upper confidence limit bound could have been produced by the
412 process treatment or mechanistic phenomenon. The analysis showed that the lower 95%
413 confidence limit for NST74*ΔaaeA* pGS2596 overlapped the upper 95% confidence limit for
414 the starting strain (NST74 pGS2596) throughout the fermentations (Fig. 8a). Therefore, even
415 though the measured total styrene produced was greater in all samples of NST74*ΔaaeA*
416 pGS2596 (1.21 g, 8 h post-induction) compared to NST74 pGS2596 (0.63 g, 8 h post-
417 induction), the two outcomes could have been produced by the same process treatment or
418 mechanistic phenomenon (Fig. 8a). However, this was not the case for NST74*ΔaaeA* $cpxP_o$
419 pGS2596, which out-performed NST74 pGS2596 with statistical significance in the last four
420 hours of the fermentation, resulting in production of 1.85 g of styrene after 8 h (Fig. 8b). The
421 phenotype of NST74*ΔaaeA* $cpxP_o$ pGS2596 thus represents a mechanistic phenomenon that
422 is distinct from the phenotype of NST74 pGS2596.

To provide statistical significance for the calculated cumulative specific productivities (Figs 8c, d and e), Monte Carlo simulations were undertaken to construct the lower and upper 90% confidence limit bounds for the technical repeats of the analyses (Fig. S4). The enhanced performance of the re-engineered strains was supported by continued glucose consumption and sustained synthesis of the precursor, phenylalanine, throughout the fermentation (Figs 8d and e). Although the cumulative glucose consumption rate was higher after induction of styrene synthesis in the original strain, compared to the re-engineered strains (Fig. 8d), synthesis of the styrene precursor, phenylalanine, declined rapidly (Fig. 8e).

The two re-engineered cell factories exhibited higher viability compared to the starting strain, and the cell counts increased post-induction, consistent with greater styrene tolerance, which permitted extension of the styrene production phase to 8 h (Fig. 8e). Notably, the lower cell counts for NST74 pGS2596 even at the point of inducing styrene production was plausibly associated with low level leaky expression of *pal2-fdc1*, consistent with the observed pre-induction changes in phospholipids (see above). For the original strain, the number of cells possessing the pGS2596 plasmid declined post-induction, (Fig. 8f) presumably through cell division that, overall, maintained viability (Fig. 8e). In contrast, the pGS2596 plasmid was markedly unstable post-induction in both re-engineered cell factories (Fig. 8f), probably due to continued cell division that, overall, increased viability (Fig. 8e). Notably, the *cpxP* over-expression plasmid was 100% retained by NST74*ΔaaeA cpxP*$_o$ pGS2596, suggesting that CpxP was beneficial during styrene production.

After one round of MOB*psi*, the best performing strain combined a Class 1 intervention, *aaeA* (differential transcription regulation, but undetected in the proteome), with a Class 3 intervention, *cpxP* (regulated in both the transcriptome and proteome). Therefore, it is possible that in this particular case both interventions could have emerged from analysis of transcriptomic data alone. However, given the many reported examples of poor correlation

between mRNA and protein copy numbers (e.g. Taniguchi et al., 2010) this is unlikely to be true in every case, and here further combinatorial re-engineering of the styrene cell factory based on the enhanced performances of the Class 2 (differentially regulated in the proteome only) candidates *hslO* and *yebE* in the small scale screening (Fig. 6) might prove beneficial. Therefore, based on the experience gained from the test case reported here a minimum set of time-resolved transcriptomic and proteomic data sets are needed to apply MOB*psi* to improve cell factory performance.

3. Conclusions

To test the potential power of MOB*psi* we constructed and characterized an *E. coli* cell factory producing styrene. The importance of considering the responses of the cell factory during intracellular styrene production was evident from the differences in the transcriptional responses of the host strain (NST74) to intracellular styrene synthesis and external styrene addition (Table S1). Machas et al. (2021) also showed that endogenous and exogenous exposure to styrene invoked different transcriptional responses in sealed bottle cultures when similarly employing NST74 as the host strain. As such, Machas et al. (2021) reported transcriptional profiling at a single time point, 27 h post-inoculation, where 1075 genes were differently expressed in cultures subject to external styrene addition, 256 genes were uniquely differentially expressed in cultures producing styrene, and 499 genes responded in both (Machas et al., 2021). This represents a far broader transcriptional response than we observed in fed-batch cultures exposed to styrene for up to 6 h. The difference in the transcriptional response between this study and Machas et al. (2021) to external styrene addition suggests that shorter exposure to external styrene (6 h compared to 27 h) has a lesser effect on *E. coli*, but as styrene accumulates and partitions into cell membranes there is a catastrophic systems failure resulting from membrane disruption with minimal opportunity to adapt, as judged by

473 the decrease in viability. In contrast, the production strain is exposed to a gradient of styrene
474 across the inner membrane, with high cytoplasmic concentrations invoking a more
475 progressive response to styrene exposure, indicative of adaptation to stress. These
476 observations emphasize the importance of using strains producing toxic chemicals and of
477 obtaining time-resolved datasets with carefully designed controls, which are embedded
478 features of MOB*psi*. In this way greater clarity on the dynamic stresses experienced by a cell
479 factory producing a toxic compound is provided, thereby elucidating opportunities for
480 interventions

481 The observation that ^{13}C-labeled styrene was rapidly incorporated into the membranes
482 of *E. coli* cells, but became undetectable within 20 min of exposure (Fig. 2) and the analyses
483 of time-resolved multi-omic data obtained from scaled-down fed-batch fermentations, led to
484 the selection of 25 target genes for cell factory re-engineering. Deletion and over-expression
485 strains were constructed to permit the effects on styrene production to be assessed in a small-
486 scale screen. The MOB*psi* methodology identified two modification strategies for NST74
487 (*ΔaaeA* and *ΔaaeA cpxP*$_o$), which were then tested in fed-batch fermentation. The most
488 probable explanation for the increased styrene productivity of the *ΔaaeA* strains is that
489 disruption of the efflux pump, AaeXAB, increases the intracellular concentration of the
490 substrate for Fdc1 (K_m for *trans*-cinnamic acid 0.18 mM; Lin et al., 2015), since AaeXAB
491 appears to be effective with cinnamic acid as a substrate (Van Dyk et al., 2004; Vargas-Tah
492 and Gosset, 2015). The exo-metabolic profiles of styrene production fermentations with *E.*
493 *coli* NST74 pGS2596 indicated that up to 0.14 g L^{-1} of *trans*-cinnamic acid was present in the
494 culture medium (Fig. 3a). However, even if this *trans*-cinnamic acid was retained by the re-
495 engineered *ΔaaeA* cell factories and converted to styrene it would not account for the
496 observed increase in styrene production. Furthermore, in flask cultures of re-engineered
497 strains, *trans*-cinnamic acid was still detected in the medium, suggesting that its efflux can

498 occur in the absence of AaeA. In this context it is noted that the AcrAB-TolC efflux pump
499 could possibly account for *trans*-cinnamic acid export in the absence of AaeXAB. However,
500 AcrAB-TolC is likely responsible for clearing styrene from the inner membrane and therefore
501 it would probably be detrimental to consider deleting *acrAB* (Mingardon et al., 2015; Van
502 Dyk et al., 2004; Vargas-Tah and Gosset, 2015). Nevertheless, the presence of *trans*-
503 cinnamic acid in the culture medium indicates that the final step of the manufacturing
504 pathway was sub-optimal in the initial strains, and tuning of Pal2 and Fdc1 activities could be
505 crucial in further improving styrene production as reported by Lee et al. (2019).

506 A desirable property of the *ΔaaeA* strains was their increased viability and sustained
507 phenylalanine synthesis during styrene production (Fig. 8e and f). It appears that even prior to
508 induction leaky expression of the styrene operon by the unmodified NST74 pGS2596 strain
509 reduces its viability, and after induction the capacity to produce phenylalanine, is impaired.
510 By contrast, the re-engineered cell factories exhibited higher viability during production and
511 maintained the flux of carbon to phenylalanine, thereby sustaining styrene production.
512 Although the reasons for these observations are unclear at present, it is noted that the *ΔaaeA*
513 strains are likely to express *aaeX*, which codes for a 90 amino acid protein of unknown
514 function with two trans-membrane helices. An *aaeX* mutant exhibits impaired biofilm
515 formation and thus it could be inferred that AaeX plays a role in cell envelope function (Kvist
516 et al., 2003). An undesirable property of the re-engineered strains was the post-induction
517 selection pressure that led to the segregational loss of the *pal2-fdc1* expression plasmid
518 (pGS2596), and it is likely that this plasmid loss limited styrene production in these
519 fermentations (Fig. 8g).

520 To build a cell factory that is capable of commercial styrene production, and other
521 similar chemicals causing widespread cellular (membrane) damage, will require more
522 extensive engineering than has been attempted thus far. The gene expression changes that

occur during a 6 h post-induction period in fed-batch fermentation were less extensive than those measured after 27 h of flask culture (Machas et al., 2021) and yielded different intervention targets (AaeXAB and CpxP versus PlsX). Changes in gene expression and protein abundances as fermentations progressed suggested that different interventions might be appropriate at different times (Tables S1-3). For example, whereas there was an early and sustained up-regulation of *cpxP* expression upon induction of styrene production, another related chaperone gene, *spy*, was only up-regulated 6 h after induction. This suggests that sophisticated control strategies will be needed to allow beneficial changes in gene expression to be implemented dynamically, either through native regulatory networks or by external control, as the fermentation proceeds to sustain productivity.

Direct comparison of cell factory performance in different studies is compromised by the different culture conditions, styrene capture and measurement techniques and the general sparsity of data. Lee et al. (2019) compared productivity data from four *E. coli* studies, including their own, and showed flask-based values ranging from 7.3 to 35.4 mg L^{-1} h^{-1}. Best productivity (~80 mg L^{-1} h^{-1}) was achieved in a 60 h fed-batch culture with *n*-dodecane solvent extraction (Lee et al., 2019), which is comparable to that of NST74Δ*aaeA* $cpxP_o$ (~70 mg L^{-1} h^{-1}). A techno-economic analysis of bio-based styrene production suggested that productivity values of 1.4 to 2.6 g^{-1} L^{-1} h^{-1} could be commercially competitive and thus even the best cell factories still perform below the level required for industrial application (Claypool et al., 2014). Nevertheless, rapid improvements are being made through increased understanding of the stresses that occur during the production of toxic chemicals and the responses of the cell factory to these challenges. The observation of permanent saturation of model membranes by styrene in MD simulations and in solid state NMR experiments with lipid extracts, but loss of styrene from NST74 cell membranes exposed to styrene, suggests that strategies to protect membranes from styrene that has entered the culture medium and

548 intensifying the efflux of styrene across and from the inner membrane could be fruitful
549 strategies for further improving cell factory performance. This might be achieved through
550 recycling re-engineered cell factories through the MOB*psi* process and combinatorial
551 investigation of a full range of potentially beneficial genes, alongside improved media
552 formulations and culture conditions.

553 In conclusion, there are three hallmarks of MOB*psi* (Fig. 1) that combine to represent
554 a conceptual advance over conventional approaches to enhancing the performance of cell
555 factories. Firstly, cell factories producing the toxic chemical of interest, rather than applying
556 it externally, are used as the basis for the analysis. Secondly, time-resolved multi-omic data
557 sets and quantitative analytics are used to populate an integrated database to facilitate the
558 identification of genetic interventions to enhance cell factory performance. Thirdly, scalable
559 fermentations, rather than plate or flask cultures, are used to generate representative data
560 thereby enhancing the likelihood of the improved cell factories having commercial relevance.
561 Although fed-batch fermentation is the most common mode of operation in industry, and was
562 thus the most appropriate choice for exemplifying the MOB*psi* approach, MOB*psi* should be
563 a widely applicable strategy for improving cell factory performance in other scalable
564 fermentations, whether batch, fed-batch or continuous. Future developments might include
565 implementing systems to down-regulate gene expression in addition to mutant analysis and
566 creating an automated pipeline for selecting candidate genetic interventions in MORF.

567

568 **4. Materials and Methods**

569 *4.1. Growth media*

570 *Escherichia coli* strain NST74 engineered for L-phenylalanine overproduction was
571 the initial cell factory used in this study (Tribe, 1987). This and derivative strains were
572 cultured in either LB (Melford, UK) or ML media at 37°C (Webb et al., 2019). For small-

573 scale screening studies M9 glycerol phenylalanine medium (pH 7.0) was used: Na_2HPO_4
574 (12.8 g L^{-1}), KH_2PO_4 (3 g L^{-1}), NaCl (0.5 g L^{-1}), NH_4Cl (1 g L^{-1}), $MgSO_4$ (0.24 g L^{-1}), $CaCl_2$
575 (0.011 g L^{-1}), glucose (4 g L^{-1}), phenylalanine (1 g L^{-1}) and arabinose (2 g L^{-1}). When
576 required, carbenicillin (50 mg L^{-1}) and/or kanamycin (50 mg L^{-1}) were added to media.
577
578 *4.2. Preparation of E. coli strains*
579 The plasmid pTpalfdc1 (obtained from Addgene) was used as a PCR template to
580 amplify the *pal2-fdc1* fragment using the primers 5´ PAL2 HIFI and 3´ fdc1 HIFI. The PCR
581 product was cloned into linearized pBAD24 using NEB HIFI assembly (creating pGS2596).
582 Primers for amplifying the linear pBAD fragment were 5´ pBAD HIFI with 3´ pBAD HIFI
583 (Table S8). A control plasmid, pBAD24-Pal2fdcmut, containing an inactivated Pal2
584 (Ser204Ala and Tyr109Phe) and inactivated Fdc1 (Gln192Ala and His193Ala) was generated
585 using Quickchange II XL site-directed mutagenesis kit (Agilent Technologies) following the
586 manufacturer's protocol with primer pairs 5´ PAL2 mut1 with 3´ PAL2 mut1, 5´ PAL2 mut2
587 with 3´ PAL2 mut2 and 5´ fdc1 mut with 3´ fdc1 mut (creating pGS2597). The pGS2596 and
588 pGS2957 plasmids were used to transform *E. coli* NST74 by electroporation (creating
589 JRG7006 and JRG7007 respectively). Transformants were selected on LB agar supplemented
590 with carbenicillin (50 mg L^{-1}).
591 *Escherichia coli* NST74 *ΔaaeAΔldhA*::*pssA-cti* was constructed using gene doctoring
592 (Lee et al., 2000) to delete the *ldhA* locus and simultaneously insert *E. coli pssA* under the
593 control of M1-93 promoter and *P. aeruginosa cti* under control of the M1-12 promoter and a
594 kanamycin resistance cassette. The *cti* gene was amplified from the genomic DNA of *P.*
595 *aeruginosa* using the primers pacti M1-12 and pacti b0015 (Table S8). The *pssA* gene was
596 amplified from *E. coli* genomic DNA using primers pssA M1-93 and pssA ldhA del (Table
597 S8). The kanamycin resistance cassette was amplified from pkd3 (Datsenko and Wanner,

2000) using kanR M1-12 link and kanR ldhA del. The M1-12 promoter was encoded in the primers for amplifying cti and the kanamycin resistance cassette. A synthetic DNA gBlock (Eurofins genomics) contained a dual transcriptional terminator and M1-93 promoter. pDOC-C (Lee et al., 2000) was linearized using the primers pDOC-C-ldhA and pDOC-C-ldhA. Amplified DNA fragments and the plasmid backbone were assembled using NEB HIFI assembly. The assembled pDOC-C-ldhApssActi plasmid and the helper plasmid pACBSCE were used to transform *E. coli* NST74 by electroporation and transformants were selected on LB agar supplemented with carbenicillin (50 mg L^{-1}) and chloramphenicol (34 mg L^{-1}). A single, successfully transformed colony was picked and transferred to LB medium (1 mL) supplemented with carbenicillin (50 mg L^{-1}), chloramphenicol (34 mg L^{-1}) and glucose (5 g L^{-1}) and the culture incubated (37°C, 200 rpm, 2 h). Cells were harvested by centrifugation and re-suspended in LB (1 mL) supplemented with L-arabinose (5 g L^{-1}). The culture was incubated (37°C, 200 rpm, 5 h) and then gene doctored *E. coli* were selected by plating on LB agar supplemented with kanamycin (50 mg L^{-1}) and sucrose (50 g L^{-1}). Successful construction of the desired strain was confirmed by colony PCR using the primers 5´ ldhA flank and 3´ ldhA flank.

Golden gate assembly was used to construct plasmids for over-expression of genes identified as potential intervention targets. Genes were amplified from *E. coli* BW25113 genomic DNA using primers pairs listed in Table S8 and KOD polymerase (Novagen) following the manufacturer's guidelines. Each primer introduced flanking BsaI sites. The optimized assembly method from the CIDAR moclo kit (Iverson et al., 2015; Addgene) was used for parts assembly with BsaI from New England Biolabs.

4.3. Small-scale screening cultures

622 Inocula for flask cultures were developed by inoculating the relevant strain into LB
623 medium (10 mL) supplemented with the relevant antibiotic and incubating overnight (200
624 rpm, 37°C). The cultures were diluted to OD_{600} 0.1 in M9 glycerol phenylalanine medium
625 containing the appropriate antibiotic in 250 mL Erlenmeyer flask with a ground glass stopper
626 to seal the flask. Cultures were incubated (200 rpm, 37°C, 72 h) and samples taken for NMR
627 analysis and to measure the OD_{600nm}.
628
629 *4.4. Fed-batch fermentations*
630 Inocula for fed-batch fermentations were developed by inoculating JRG7006 or
631 JRG7007 (*E. coli* NST74 pGS2596 or *E. coli* NST74 pGS2597) into ML medium (50 mL)
632 supplemented with glucose (10 g L^1) and carbenicillin (50 mg L^{-1}) and incubating overnight
633 (200 rpm, 37°C) (Webb et al., 2019). The cultures were diluted to OD_{600} 0.1 in sterile H_2O
634 (50 mL) and used to inoculate ML medium (1 L) supplemented with glucose (11.9 g L^{-1}) and
635 carbenicillin in a 3 L BioFlo®/CelliGen® 115 Bioreactor (New Brunswick Scientific).
636 Cultures were grown at 37°C and the pH was maintained at 7.0 ± 0.1 by the addition of 28-
637 30% NH_4OH and 2M H_2SO_4. For the batch phase, the air flow rate was set at 1 L min^{-1} and
638 the dO_2 was maintained at 30% of saturation by automatic control of stirrer speed between
639 400-1200 rpm. For the initial fermentations, the air flow rate was decreased to 0.1 L min^{-1}
640 post induction of Pal2 and Fdc1 to decrease styrene stripping such that cell factory responses
641 to styrene could be measured. Once the glucose in the batch medium had been consumed,
642 indicated by a sharp increase in dO_2 and confirmed using Glucose Test Strips (Merck
643 Millipore - 117866), a feed of glucose (650 g L^{-1}), trace elements (11.8 mL) and thiamine
644 (13.5 mg L^{-1}) was started. For the fermentations associated with time-resolved 'omics data
645 collection, the feed rate was 0.046 mL min^{-1} (1.79 g L^{-1} h^{-1} D-glucose). This was maintained
646 for 1 h, before increasing to 0.091 mL min^{-1} (3.55 g L^{-1} h^{-1} D-glucose) for 1 h and then to

0.14 mL min^{-1} (5.46 g L^{-1} h^{-1} D-glucose) for the remainder of the fermentation. For the fermentations testing the performance of the re-engineered strains, an air-flow cascade of 0.1-0.3 L min^{-1} was applied after induction to maintain dO_2 at 30% with stripped styrene being collected in two dodecane traps (in series) attached to the gas outlet. The initial D-glucose feed rate was 0.033 mL min^{-1} (1.28 g L^{-1} h^{-1} D-glucose) for 1 h, before increasing to 0.084 mL min^{-1} (3.23 g L^{-1} h^{-1} D-glucose) for 1 h and then to 0.13 mL min^{-1} (5.06 g L^{-1} h^{-1} D-glucose) for the remainder of the fermentation. In all cases, Pal2 and Fdc1 expression was induced by the addition of L-arabinose (0.02 g L^{-1}) at OD_{600} ~30 and polypropylene glycol was manually added to the medium to avoid foaming as required.

For experiments in which cultures were exposed to external styrene, the same conditions applied except the fermenter was connected to a single-syringe infusion pump (Cole Parmer - 78-74900-05) fitted with a 50 mL Gastight™ syringe (Hamilton – 85020) using neoprene tubing (Cole Parmer) to feed styrene into the vessel at 2 mL h^{-1} for the first 2 h after induction and then at 3 mL h^{-1} thereafter.

For all fermentations samples (25 mL) were taken at the indicated times after purging the sampler line.

4.5. Radial basis function modeling

Enabling statistically significant comparison between fed-batch fermentations, radial basis function models were constructed following the approach advocated by Leonard et al. (1992). The observed fed-batch fermentation data sets associated with the time course of total styrene production were modeled. Notably, the use of radial basis functions enables the calculation of 95% confidence limits for the respective models, bounding the statistical significance of where the measured observations may be attributed to the process treatment or process mechanistic phenomenon.

672

673 *4.6. Molecular dynamics simulations*

674 All MD simulations were carried out on a Supermicro server with NVIDIA K80 and
675 K40 general purpose graphical processing units, GPGPUs. Molecular system assembly was
676 done using CHARMM (Brooks et al., 1983) through the Multicomponent Builder in
677 CHARMM-GUI (Jo et al., 2008) and MD simulations were done using NAMD 2.12 (Phillips
678 et al., 2005). The system was equilibrated for 750 ps at 1 fs per step and further 750 ps at 2
679 fs/step as an isothermal, constant volume canonical ensemble (NPT). Production runs were
680 carried out at 37°C at 2 fs/step under isobaric canonical ensemble using Langevin piston
681 control. Trajectory analysis and visualization was done using UCSF Chimera (Pettersen et al.,
682 2004).

683

684 *4.7. Solid state NMR*

685 All solid state NMR experiments were carried out on a Varian VNMRS400 direct
686 drive solid state NMR spectrometer equipped with a balanced heated/vortex tube-cooled
687 airflow temperature controlled (Ciesielski et al., 2009) 4mm T4 MAS probe. All ^{13}C spectra
688 were referenced externally using adamantane CH_2 at 37.54 ppm. Cross polarization magic
689 angle spinning (CP MAS) (Hartmann and Hahn,1962; Pines et al., 1973) excitation was done
690 using a tangent ramp (Hediger et al., 1994) on ^{1}H with constant amplitude on ^{13}C and free
691 induction decays were acquired at 5 kHz MAS and 8°C under SPINAL-64 heteronuclear
692 proton decoupling (Fung et al., 2000). Proton excitation of 120 kHz was followed by 3.5 ms
693 45 kHz Hartmann-Hahn contact and 125 ms acquisition under 67 kHz proton decoupling over
694 a 50 kHz spectral width with 3.5 s repeat delay. Each spectrum was obtained by averaging
695 1024 FIDs and processed with 2 Hz exponential line broadening using ACD-labs. Unlabelled
696 styrene was purchased from Merck at >99% purity and used without further purification.

Lipid extracts were collected using chloroform-methanol method (Bligh and Dyer 1959). Samples for CP experiments were prepared by hydrating 10 mg of lipid extract with 30 µl HPLC grade water, and for sample containing styrene, 5 mg of unlabelled styrene was added to the dried lipid extract prior to hydration.

4.8. Whole cell solid state NMR

Doubly $^{13}C(\alpha,\beta)$-labeled styrene containing 4-*t*-butylcatechol stabilizer was purchased from Merck at >99% purity and used without further purification. Lipid extracts were collected as described in *Lipidomics* section. Fed-batch *E. coli* NST74 pGS2597 culture, stabilized at OD_{600} = 4, was inoculated with 300 µl of ^{13}C labeled styrene and samples were collected immediately, 20 min, 2 h and 24 h post-inoculation. NMR experiments were done using cells harvested by centrifugation from the fed-batch fermentation. The cells were washed with PBS and the centrifuged, hydrated pellets were kept on ice until loaded into 4 mm zirconia MAS rotors *via* gentle centrifugation. All samples bar the one collected 20 min post addition of ^{13}C-labeled styrene were measured immediately after cell harvesting. The sample collected 20 min after styrene addition was stored at -20°C until previous NMR experiments were completed. The ^{13}C CP-MAS NMR spectra of whole cells, harvested from fermentation, were acquired by averaging 1024 transient at a recycle delay of 3.5 s, ~1 h of experimental time.

4.9. Analytical methods

Growth was monitored by measuring optical densities (OD_{600nm}). Samples were diluted in deionized water when the OD_{600nm} was >0.8. Dry cell weight was measured by centrifuging 1 mL samples in pre-weighed polypropylene tubes, removing the supernatant and drying the pellets to a constant weight. D-Glucose, styrene, phenylalanine, tyrosine,

722 *trans*-cinnamic acid, acetate, pyruvate and lactate were measured by High Pressure Liquid
723 Chromatography (HPLC) using an Agilent 1200 series HPLC system equipped with both UV
724 (215 nm) and refractive index detectors. Samples were resolved using a Rezex ROA Organic
725 Acid H+ column (Phenomenex) at 55°C with 0.01 N H_2SO_4 (0.5 mL min^{-1}) as the mobile
726 phase. Samples were prepared for HPLC analysis by centrifuging (12,000×*g*, 5 min) and
727 filtering the supernatants (0.2 μm filter). Data analysis was performed with ChemStation
728 software, using calibration curves prepared using authentic standards of each compound (0.1-
729 200 mM).

730 Chemical composition of flask cultures (whole culture samples consisting of cells and
731 medium) was determined by high resolution solution-state NMR. Samples (540 μL) were
732 mixed with D_2O (60 μL) containing trimethylsilylpropanoic acid (TSP, 2 mM) and loaded
733 into 5 mm (o.d.) nmr tubes. Each tube was hand-centrifuged for 10 s before loading into an
734 automatic sample changer, SampleJet (Bruker), for acquisition. 1D Nuclear Overhauser
735 Effect (pulse program, noesygppr1d, Bruker) ^{1}H NMR spectra were acquired at 800.34 MHz
736 using a Bruker Neo 800 MHz NMR spectrometer fitted with a 5 mm TCI cryoprobe. Samples
737 were referenced to TSP at 0 ppm. 1D spectra had spectral widths of 20 ppm, 6 s relaxation
738 delay, 32 scans collected with 128k data points (3.93 s acquisition time). Temperature was set
739 to 25°C (± 0.1°C) prior to acquisition and receiver gain was constant for all acquisition. All
740 spectra were processed to 128k data points with exponential line broadening of 0.1 Hz
741 applied before Fourier transformation, spectra were automatically phased, referenced and
742 baseline corrected in Topspin 4.0.5 (Bruker, UK). Chemical composition was determined by
743 comparison with chemical shifts from in-house reference spectra and concentrations of
744 styrene, *trans*-cinnamic acid and phenylalanine were determined by comparison with the TSP
745 signal (0.2 mM).
746

747 *4.10. Transcriptomics*

748 Fermentation samples (3 x 0.1 mL) were obtained and processed as previously
749 described to obtain Cy3-labeled cDNA (Webb et al., 2019). Cy5-labeled *E. coli* genomic
750 DNA was produced using BioPrime DNA Labeling Kit (Invitrogen) with Cy5-dCTP in the
751 dNTP mixture. Labeled genomic DNA and cDNA were combined and hybridized to
752 oligonucleotide microarrays (Agilent Technologies). Quantification of cDNA samples,
753 hybridization to microarrays, microarray processing and scanning were carried out as
754 described in the Fairplay III labeling kits (Agilent Technologies, 252009, Version 1.1) and
755 scanned with a high-resolution microarray scanner (Agilent Technologies). Features with
756 background intensities exceeding 10 times the array median, or with a signal to background
757 ratio below 3 were excluded from further analysis. Background correction (Ritchie et al.,
758 2007), within-array Loess normalization (Smyth and Speed, 2003), and between-array
759 quantile normalization were applied to the remaining features using the R statistical package
760 LIMMA from Bioconductor (Gentleman et al., 2004). Moderated t-statistics were calculated
761 using gene-wise linear models with an empirical Bayes approach (Ritchie et al., 2015;
762 Phipson et al., 2016). *p*-Values were adjusted for multiple testing using the Benjamini-
763 Hochberg method (Benjamini and Hochberg, 1995).

764 Changes in the activities of transcription factors in response to styrene were inferred
765 from the transcriptomic data combined with knowledge of the regulon structure of *E. coli*
766 using the TFInfer software package as previously described (Asif et al., 2010). Changes in
767 transcription factor activity were deemed significant if the difference between the maximum
768 and minimum activity divided by the confidence interval was ≥3.

769

770 *4.11. Metabolomics*

771 Quenching of fermentation samples (0.5 mL) was performed as previously described (Webb
772 et al., 2019; Spura et al., 2009). Sample analysis was performed by automated direct
773 injection analysis using an Acquity UPLC system into a Synapt G2 mass spectrometer
774 (Waters, UK) in both positive and negative electrospray modes. Each sample was analyzed 3
775 times to obtain technical replication. Data acquisition and processing was performed on
776 MassLynx (version 4.1) to create centroid peak lists (m/z accurate to 4 decimal places vs. ion
777 counts). Samples were analyzed in a randomized order to minimize effects of day-to-day
778 machine variation. Data processing and downstream analysis was performed in R, using
779 Bioconductor package XCMS (Smith et al., 2006). Peaks were aligned across analytical
780 replicates and grouped into 0.2 m/z width bins (Kazmi et al., 2006). Peaks were rejected if all
781 three replicates were not present, or the mass variance fell outside an acceptable range
782 defined as a function of the m/z (formula modified from Overy et al., 2004).
783
784 *4.12. Proteomics*
785 Fermentation samples (500 µL) were harvested by centrifugation (16,000×*g*, 2.5 min,
786 4°C) the supernatant was discarded and the pellet stored (-80°C). Cell pellets corresponding
787 to 0.3 mg dry protein were suspended in Tris buffer (20 mM, pH 8.0, 20 µL), centrifuged
788 (12,000×*g*, 15 min) and the supernatant was removed. Denaturation, reduction and alkylation
789 were induced by addition of urea (6 M), triethylammonium bicarbonate (TEAB; 100 mM),
790 chloracetamide (10 mM), Tris(2-carboxyethyl)phosphine hydrochloride (TCEP; 5 mM) and
791 incubation (30 min, room temp.). Formation of gel plugs was induced by addition of
792 ammonium persulfate (2.5 µL) and TEMED (2.5 µL). Samples were mixed by vortexing. Gel
793 plugs were shredded into gel pieces by centrifugation through the mesh of a filter holder from
794 an empty Spin-X column. Gel pieces were washed and dried as described previously Fischer
795 and Kessler (2015). Overnight digestion was induced by addition of modified porcine trypsin

(6 µg) to reach a 1:50 enzyme:substrate ratio. Two additional protease additions were performed on the next morning and 4 h later.

Peptides (50 µg) were labelled with TMT (100 µg) using 10-plex (Thermo Fisher Scientific) in HEPES buffer as previously described (Zecha et al., 2019). Pooled samples were acidified using trifluoroacetic acid (0.2% v/v) and desalted using C18 cartridges (SolaQ, Thermo).

Samples were pre-fractionated in 10 subsequent elution fractions using a high-pH reverse phase fractionation kit (Thermo Fisher Scientific) (Paulo et al., 2018). Mass spectra were acquired in positive ion mode applying data acquisition using synchronous precursor selection MS3 (SPS-MS3) acquisition mode (McAlister et al., 2014; Queiroz et al., 2019). Samples were analyzed in an Orbitrap Fusion Lumos (Thermo Fisher Scientific), coupled to a Dionex Ultimate 3000 UHPLC. Samples were resolved using a 50 cm long PepMap nanoLC column with a gradient (9%-45% over 2 h) of Buffer B (80% acetonitrile containing 0.1% formic acid) and SPS-MS3 acquisition.

MS spectra processing and peptide and protein identification: Raw data were processed using Proteome Discoverer v2.1 (Thermo Fisher Scientific). The raw files were submitted to a database search using Proteome Discoverer with Mascot and SequestHF algorithms against the *E. coli* database downloaded in early 2017, UniProt/TrEMBL. Common contaminant proteins (several types of human keratins, BSA and porcine trypsin) were added to the database, and all contaminant proteins identified were removed from the result lists before further analysis. The spectra identification was performed with the following parameters: MS accuracy, 10 p.p.m.; MS/MS accuracy of 0.05 Da for spectra acquired in Orbitrap analyzer and 0.5 Da for spectra acquired in Ion Trap analyzer; up to two missed cleavage sites allowed; carbamidomethylation of cysteine (as well as TMT6plex tagging of lysine and peptide N terminus for TMT labeled samples) as a fixed modification;

and oxidation of methionine and deamidated asparagine and glutamine as variable modifications. Percolator node was used for false discovery rate estimation and only rank 1 peptide identifications of high confidence (FDR<1%) were accepted. TMT reporter values were assessed through Proteome Discoverer v2.1 using the Most Confident Centroid method for peak integration and integration tolerance of 20 p.p.m. Reporter ion intensities were adjusted to correct for the isotopic impurities of the different TMT reagents (according to the manufacturer specifications for the respective batch number).

The data obtained from Proteome discoverer was abundance data at the protein level. Statistical analysis was performed using the LIMMA package from Bioconductor. Proteins exhibiting a ≥2-fold change in abundance and an FDR adjusted *p*-value <0.05 were deemed to be enriched.

4.13. Lipidomics

Bacteria from fermentation samples (500 µL) were collected by centrifugation (16,000×*g*, 2.5 min, 4°C) and processed as previously described (Webb et al., 2019). Technical triplicates were extracted from each fermenter at each time point, and extraction blanks were processed in parallel to correct for any background contamination in subsequent analyses.

For GC-MS and GC-FID analyses, FAMEs were generated and analyzed as previously described (Webb et al., 2019) For LC-MS analysis, samples were reconstituted in acetonitrile:isopropanol (200 µl, 70:30, v/v) and then processed and analyzed as previously described (Webb et al., 2019). Peak processing workflows were performed in R 3.5.0. MS1 peaks were extracted using the centWaveWithPredictedIsotopeROIs in the XCMS package (version 3.2.0; Smith et al., 2006; Tautenhahn et al., 2008). Peaks were grouped across samples, missing peaks imputed by re-integration within group boundaries, and group median

m/z values further processed using the CAMERA package (version 1.23.3; Kuhl et al., 2012). Candidate formulae were generated with adapted code from the rcdk package (version 3.4.9; Guha et al., 2007), with the following limits: positive mode C10-300, H20-500, O0-20, N0-3, P0-2, Na02, RDBE -0.5-18, 2 ppm error; negative mode C10-300, H20-500, O0-20, N0-3, P0-2, S0-2, RDBE -0.5-18, 20 ppm error. Consensus peak groups were then filtered with custom R scripts to: 1) exclude any peak groups where any peak in the group was present with an area less than the mean +3 standard deviations of the value from blank extracts; 2) only retain the most intense monoisotopic peak identified by CAMERA; 3) only retain peaks with valid molecular formulae. MS1 peaks were searched against downloaded local copies of the *E. coli* metabolome database (ECMDB; Guo et al., 2013) and the Lipid Maps Structure Database (LMSD; http://www.lipidmaps.org). Consensus HCD and CID MS2 spectra were extracted with custom scripts and searched against the *in silico* LipidBlast (Kind et al., 2013) and LipidMatch (Koelmel et al., 2017) databases. Peaks were annotated following manual examination of MS2 spectra in consensus with returned database hits. All retained peaks were normalized to the SPLASH deuterated PC(15:0_18:1(d7)) internal standard and then to sample DW. Statistical analysis were carried out on glog normalized data (lambda value in glog transform taken to be 1/10 of non-zero minimum), using time-series ANOVA2 models from the online MetaboAnalyst resource (https://www.metaboanalyst.ca).

Author statements

Joseph P Webb: Methodology, Formal analysis, Investigation, Writing - Original Draft, Writing - Review and Editing, Visualization.

Ana Carolina Paiva: Formal analysis, Investigation, Writing - Review and Editing.

Luca Rossoni: Investigation.

Amias Moore: Investigation.

Vicki Springthorpe: Methodology, Software, Formal analysis.

Sophie Vaud: Investigation.

Vivien Yeh: Investigation.

David-Paul Minde: Investigation.

Sven Langer: Investigation.

Heather Walker: Investigation.

Andrea Hounslow: Investigation.

David Nielsen: Resources, Writing - Review and Editing.

Tony Larson: Investigation, Supervision.

Kathryn Lilley: Conceptualization, Methodology, Supervision, Funding acquisition.

Gill Stephens: Conceptualization, Methodology, Writing - Review and Editing, Supervision, Funding acquisition.

Gavin H Thomas: Conceptualization, Methodology, Writing - Review and Editing, Supervision, Project administration, Funding acquisition.

Boyan Bonev: Conceptualization, Methodology, Software, Formal Analysis, Writing - Review and Editing, Visualization, Supervision.

David J Kelly: Conceptualization, Methodology, Writing - Review and Editing, Supervision, Funding acquisition.

Alex Conradie: Conceptualization, Methodology, Formal analysis, Writing - Review and Editing, Visualization, Supervision.

Jeffrey Green: Conceptualization, Methodology, Writing - Original Draft, Writing - Review and Editing, Supervision, Funding acquisition.

Declaration of competing interests

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the InnovateUK and the Biotechnology and Biological Sciences Research Council (grant Industrial Biotechnology Catalyst BB/N01040X/1). Solid state NMR instrument funding was provided by the Biotechnology and Biological Sciences Research Council (grant BB/C510924) to BB. The authors thank; NVIDIA Corporation for GPGPU hardware grant; Mike Deery (University of Cambridge) for technical assistance with mass spectrometry; Ben Willson, Reyme Herman, Joyce Bennett (University of York), Susan Molyneux-Hodgson, Sally Atkinson (University of Exeter), Chris Hills (Biocleave), Graham Eastham, David Johnson (Mitsubishi Chemical), Bob Holt (CPI), Preben Krabben (Deep Branch Biotechnology) and Reuben Carr (Ingenza) for helpful discussions.

Data Availability

All the time-resolved fermentation, transcriptomic, proteomic, metabolomic and lipidomic data sets are available in MORF (https://morf-db.org/projects/DETOX/styrene). Transcriptomic and proteomic data are also available at ArrayExpress E-MTAB-8703 (https://www.ebi.ac.uk/arrayexpress) and at PRIDE PXD028971 (https://www.ebi.ac.uk/pride).

Appendix A. Supplementary data

Movie S1

Supplementary data for this article can be found online

References

Al Mamun AA, Lombardo MJ, Shee C, Lisewski AM, Gonzalez C, Lin D, Nehring RB, Saint-Ruf C, Gibson JL, Frisch RL, et al (2012) Identity and function of a large gene network underlying mutagenic repair of DNA breaks. *Science* 338: 1344-1348

Aono R, Kobayashi H (1997) Cell surface properties of organic solvent-tolerant mutants of *Escherichia coli* K-12. *Appl Environ Microbiol* 63: 3637-3642

Asif HM, Rolfe MD, Green J, Lawrence ND, Rattray M, Sanguinetti G (2010) TFInfer: a tool for probabilistic inference of transcription factor activities. *Bioinformatics* 26: 2635-2636

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B* 57: 289-300

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 37: 911-917

Brooks BR, Bruccoleri RE, Olafson, BD, States D J, Swaminathan S, Karplus M (1983) CHARMM - a program for macromolecular energy minimization and dynamics calculations. *J Comp Chem* 4: 187-217

935 Bury-Moné S, Nomane Y, Reymond N, Barbet R, Jacquet E, Imbeaud S, Jacq A, Bouloc P
936 (2009) Global analysis of extracytoplasmic stress signaling in *Escherichia coli*. *PLoS Genet*
937 5: e1000651

938 Ciesielski F, Griffin DC, Rittig M, Bonev BB (2009) High-resolution J-coupled ^{13}C MAS
939 NMR spectroscopy of lipid membranes. *Chem Phys Lipids* 161: 77-85

940 Claypool JT, Raman DR, Jarboe LR, Nielsen DR (2014) Technoeconomic evaluation of bio-
941 based styrene production by engineered *Escherichia coli*. *J Ind Microb Biotechnol* 41: 1211-
942 1216

943 Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in
944 *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97: 6640–6645

945 Fischer R, Kessler BM (2015) Gel-aided sample preparation (GASP)—A simplified method
946 for gel-assisted proteomic sample generation from protein extracts and intact cells.
947 *Proteomics* 15: 1224-1229

948 Fung BM, Khitrin AK, Ermolaev K (2000) An improved broadband decoupling sequence for
949 liquid crystals and solids. *J Mag Res* 142: 97-101

950 Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge
951 Y, Gentry J, et al. (2004) Bioconductor: open software development for computational
952 biology and bioinformatics. *Genome Biol* 5: R80

953 Guha R (2007) Chemical informatics functionality in R. *J Stat Software* 6: 18

954 Guo AC, Jewison T, Wilson M, Liu Y, Knox C, Djoumbou Y, Lo P, Mandal R,
955 Krishnamurthy R, Wishart DS (2012) ECMDB: The *E. coli* Metabolome Database. *Nucleic*
956 *Acids Res* 41:D625-630

957 Hartmann SR, Hahn EL (1962) Nuclear double resonance in rotating frame. *Phys Rev* 128:
958 2042

959 Hediger S, Meier BH, Kurur ND, Bodenhausen G, Ernst RR (1994) NMR cross polarization
960 by adiabatic passage through the Hartmann—Hahn condition (APHH). *Chem Phys Lett* 223:
961 283-288

962 Hemm MR, Paul BJ, Miranda-Ríos J, Zhang A, Soltanzad N, Storz G (2010) Small stress
963 response proteins in *Escherichia coli*: proteins missed by classical proteomic studies. *J*
964 *Bacteriol* 192: 46-58

965 Hengge R, Galperin MY, Ghigo JM, Gomelsky M, Green J, Hughes KT, Jenal U, Landini P
966 (2016) Systematic nomenclature for GGDEF and EAL domain-containing cyclic-di-GMP
967 turnover proteins of *Escherichia coli. J Bacteriol* 198: 7-11

968 Iverson SV, Haddock TL, Beal J, Densmore DM (2015) CIDAR MoClo: Improved MoClo
969 assembly standard and new *E. coli* part library enable rapid combinatorial design for
970 synthetic and traditional biology. *ACS Synth Biol* 5: 99-103

971 Jakob U, Muse W, Eser M, Bardwell JC (1999) Chaperone activity with a redox switch. *Cell*
972 96: 341-352

973 Jarboe LJ (2018) Improving the success and impact of the metabolic engineering design build
974 test learn cycle by addressing proteins of unknown function. *Curr Opin Biotechnol* 53: 93-98

975 Jo S, Kim T, Iyer VG, Im W (2008) CHARNIM-GUI: A web-based grraphical user interface
976 for CHARMM. *J Comp Chem* 29: 1859-1865

977 Joly N, Engl C, Jovanovic G, Huvet M, Toni T, Sheng X, Stumpf MP, Buck M (2010)
978 Managing membrane stress: the phage shock protein (Psp) response, from molecular
979 mechanisms to physiology. *FEMS Microbiol Rev* 34: 797-827

980 Kazmi SA, Ghosh S, Shin DG, Hill DW, Grant DF (2006) Alignment of high resolution mass
981 spectra: development of a heuristic approach for metabolomics. *Metabolomics* 2: 75–83

982 Keseler IM, Gama-Castro S, Mackie A, Billington R, Billington R, Caspi R, Kothari A,
983 Krummenacker M, Midford P, Muñiz-Rascado L, Ong W, Paley S, Santos-Zavaleta A,
984 Subhraveti P, Tierrafria V, Wolfe A, Collado-Vides J, Paulsen I, Karp PD (2021) The
985 EcoCyc database in 2021 *Front Microbiol* 12: 711077

986 Kind T, Liu KH, Lee DY, DeFelice B, Meissen JK, Fiehn O (2013) LipidBlast in silico
987 tandem mass spectrometry database for lipid identification. *Nat Methods* 10: 755-758

988 Kobayashi H, Yamamoto M, Aono R (1998) Appearance of a stress-response protein phage-
989 shock protein A in *Escherichia coli* exposed to hydrophobic organic solvents. *Microbiology*
990 144: 353-359

991 Koelmel JP, Kroeger NM, Ulmer CZ, Bowden JA, Patterson RE, Cochran JA, Beecher
992 CWW, Garrett TJ, Yost RA (2017) LipidMatch: an automated workflow for rule-based lipid
993 identification using untargeted high-resolution tandem mass spectrometry data. *BMC*
994 *Bioinformatics* 18: 331

995 Kuhl C, Tautenhahn R, Böttcher C, Larson TR, Neumann S (2012) CAMERA: an integrated
996 strategy for compound spectra extraction and annotation of liquid chromatography/mass
997 spectrometry data sets. *Anal Chem* 84: 283-289

998 Kvist M, Hancock V. Klemm P (2008) Inactivation of efflux pumps abolishes bacterial
999 biofilm formation. *Appl Environ Microbiol* 74: 7376-7382

1000 Laubacher ME, Ades SE. (2008) The Rcs phosphorelay is a cell envelope stress response
1001 activated by peptidoglycan stress and contributes to intrinsic antibiotic resistance. *J Bacteriol*
1002 190: 2065-2074

1003 Lee DJ, Lewis EH, Heurlier K, Pallen MJ, Penn CJ, Busby SJW, Hobman JL (2009)Gene
1004 doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains.
1005 *BMC Microbiol* 9: 252

1006 Lee K, Bang HB, Lee YH, Jeong KJ (2019) Enhanced production of styrene by engineered
1007 *Escherichia coli* and *in situ* product recovery (ISPR) with an organic solvent. *Microb Cell*
1008 *Fact* 18: 79

1009 Lee J. Hiibel SR. Reardon KF. Wood TK (2010) Identification of stress-related proteins in
1010 *Escherichia coli* using the pollutant *cis*-dichloroethylene. *J Appl Microbiol* 108: 2088-2102

1011 Leonard JA, Kramer MA, Ungar LH (1992) A neural network architecture that computes its
1012 own reliability. *Comput Chem Eng* 16: 819-835

1013 Lian J, McKenna R, Rover MR, Nielsen DR, Wen Z, Jarboe LR (2016) Production of
1014 biorenewable styrene: utilization of biomass-derived sugars and insights into toxicity. *J Ind*
1015 *Microbiol Biotechnol* 43: 595-604

1016 Liang L, Liu R, Foster KEO, Choudhury A, Cook S, Cameron JC, Srubar WV, Gill RT
1017 (2020) Genome engineering of *E. coli* for improved styrene production. *Metab Eng* 57: 74–
1018 84

1019 Lin F, Ferguson KL, Boyer DR, Lin XN, Marsh EN (2015) Isofunctional enzymes PAD1 and
1020 UbiX catalyze formation of a novel cofactor required by ferulic acid decarboxylase and 4-
1021 hydroxy-3-polyprenylbenzoic acid decarboxylase. *ACS Chem Biol* 10: 1137-1144

1022 Liu C, Me, X, Chen H, Li M, Ding Z, Chen G, Wang F, Liu H, Wang Q, Zhu Y, Zhang H,
1023 Xian M (2018) A systematic optimization of styrene biosynthesis in *Escherichia coli* BL21
1024 (DE3). *Biotechnol. Biofuels* 11: 14

1025 Lundrigan MD, Earhart CF (1984) Gene *envY* of *Escherichia coli* K-12 affects
1026 thermoregulation of major porin expression. *J Bacteriol* 157: 262-268

1027 Machas M, Kurgan G, Abed OA, Shapiro A, Wang X, Nielsen D (2021) Characterizing
1028 *Escherichia coli*'s transcriptional response to different styrene exposure modes reveals novel
1029 toxicity and tolerance insights. *J Ind Microbiol Biotechnol* 48: kuab019

1030 McAlister GC, Nusinow DP, Jedrychowski MP, Wühr M, Huttlin EL, Erickson BK, Rad R,
1031 Haas W, Gygi SP (2014) MultiNotch MS3 enables accurate sensitive and multiplexed
1032 detection of differential expression across cancer cell line proteomes. *Anal Chem* 86: 7150-
1033 7158

1034 McKenna R, Moya L, McDaniel M, Nielsen DR (2015) Comparing *in situ* removal strategies
1035 for improving styrene bioproduction. *Bioprocess Biosyst Eng* 38: 165-174

1036 McKeena R, Nielsen DR (2011) Styrene biosynthesis from glucose by engineered *E. coli*.
1037 *Metab Eng* 13: 544-554

1038 Mingardon F, Clement C, Hirano K, Nhan M, Luning EG, Chanal A, Mukhopadhyay A
1039 (2015) Improving olefin tolerance and production in *E. coli* using native and evolved AcrB.
1040 *Biotechnol Bioeng* 112: 879-888

1041 Mogk A, Deuerling E, Vorderwülbecke S, Vierling E, Bukau B (2003) Small heat shock
1042 proteins ClpB and the DnaK system form a functional triade in reversing protein aggregation.
1043 *Mol Microbiol* 50: 585-595

1044 Murzyn, K, Rog T, Pasenkiewicz-Gierula M (2005) Phosphatidylethanolamine-
1045 phosphatidylglycerol bilayer as a model of the inner bacterial membrane. *Biophys J* 88: 1091-
1046 1103

1047 Overy DP, Blunt JW (2004) Corymbiferan lactones from *Penicillium hordei*: stimulation of
1048 novel phenolic metabolites using plant tissue media. *J Nat Prod* 67: 1850-1853

1049 Paulo JA, Jedrychowski MP, Chouchani ET, Kazak L, Gygi SP (2018) Multiplexed isobaric
1050 tag-based profiling of seven murine tissues following in vivo nicotine treatment using a
1051 minimalistic proteomics strategy. *Proteomics* 18: e1700326

1052 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE
1053 (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J*
1054 *Comp Chem* 25: 1605-1612

1055 Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale
1056 L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comp Chem* 26: 1781-
1057 1802

1058 Phipson B, Lee S, Majewski IJ, Alexander WS, Smyth GK (2016) Robust hyperparameter
1059 estimation protects against hypervariable genes and improves the power to detect differential
1060 expression. *Ann Appl Stat.* 10: 946-963

1061 Pines A, Gibby MG, Waugh JS (1973) Proton-enhanced NMR of dilute spins in solids. *J*
1062 *Chem Phys* 59: 569-590

1063 Quan S, Koldewey P, Tapley T, Kirsch N, Ruane KM, Pfizenmaier J, Shi R, Hofmann S, Foit
1064 L, Ren G, et al. (2011) Genetic selection designed to stabilize proteins uncovers a chaperone
1065 called Spy. *Nat Struct Mol Biol* 18: 262-269

1066 Queiroz RML, Smith T, Villanueva E, Marti-Solano M, Monti M, Pizzinga M, Mirea DM,
1067 Ramakrishna M, Harvey RF, Dezi V, et al. (2019) Comprehensive identification of RNA-
1068 protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat*
1069 *Biotechnol* 37: 169-178

1070 Ritchie M, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, Smyth G (2007) A
1071 comparison of background correction methods for two-colour microarrays. *Bioinformatics*
1072 23: 2700–2707

1073 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) Limma powers
1074 differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids*
1075 *Res* 43: e47

1076 Sajed T, Marcu A, Ramirez M, Pon A, Guo A, Knox C, Wilson M, Grant J, Djoumbou Y,
1077 Wishart D (2016) ECMDB 20: A richer resource for understanding the biochemistry of *E.*
1078 *coli*. *Nucleic Acids Res* 44:D495-501

1079 Sargentini NJ, Gularte NP, Hudman DA (2016) Screen for genes involved in radiation
1080 survival of *Escherichia coli* and construction of a reference database. *Mutat Res* 793-794: 1-
1081 14

1082 Si HM, Zhang F, Wu AN, Han RZ, Xu GC, Ni Y (2016) DNA microarray of global
1083 transcription factor mutant reveals membrane-related proteins involved in *n*-butanol tolerance
1084 in *Escherichia coli*. *Biotechnol Biofuels* 9: 114

1085 Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G (2006) XCMS: processing mass
1086 spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and
1087 identification. *Anal Chem* 78: 779-787

1088 Smyth GK, Speed T (2003) Normalization of cDNA microarray data. *Methods* 31: 265-273

1089 Springthorpe V, Leaman R, Sifouna D, Bennett J, Thomas G(2020) MORF: An online tool
1090 for exploring microbial cell responses using multi-omics analysis. *Access Microbiol* 2: 7A

1091 Spura J, Reimer LC, Wieloch P, Schreiber K, Buchinger S, Schomburg D (2009) A method
1092 for enzyme quenching in microbial metabolome analysis successfully applied to gram-
1093 positive and gram-negative bacteria and yeast. *Anal Biochem* 394: 192-201

1094 Tautenhahn R, Böttcher C, Neumann S (2008) Highly sensitive feature detection for high
1095 resolution LC/MS. *BMC Bioinformatics* 28: 504

1096 Tan Z, Khakbaz P, Chen Y, Lombardo J, Yoon JM, Shanks JV, Klauda JB, Jarboe LR (2017)
1097 Engineering *Escherichia coli* membrane phospholipid head distribution improves tolerance
1098 and production of biorenewables. *Metab Eng* 44: 1-12

1099 Tan Z, Yoon JM, Nielsen DR, Shanks JV, Jarboe LR (2016) Membrane engineering via *trans*
1100 unsaturated fatty acids production improves *Escherichia coli* robustness and production of
1101 biorenewables. *Metab Eng* 35: 105-113

1102 Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS (2010)
1103 Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single
1104 cells. *Science* 329: 533-538

1105 Tribe DE (1987) Novel microorganism and method. US Patent 4,681,852

1106 Van Dyk TK, Templeton LJ, Cantera KA, Sharpe PL, Sariaslani FS (2004) Characterization
1107 of the *Escherichia coli* AaeAB efflux pump: a metabolic relief valve? *J Bacteriol* 186: 7196-
1108 7204

1109 Vargas-Tah A, Gosset G (2015) Production of cinnamic and *p*-hydroxycinnamic acids in
1110 engineered microbes. *Front Bioeng Biotechnol* 3: 116

1111 Vogt SL, Raivio TL (2011) Just scratching the surface: an expanding view of the Cpx
1112 envelope stress response. *FEMS Microbiol Lett* 326: 2-11

1113 Webb J, Springthorpe V, Rossoni L, Minde DP, Langer S, Walker H, Alstrom-Moore A,
1114 Larson T, Lilley K, Eastham G, et al. (2019) Systems analyses reveal the resilience of
1115 *Escherichia coli* physiology during accumulation and export of the nonnative organic acid
1116 citramalate. *mSystems*. 4: e00187-19

1117 Wu C, Koylinski T, Bozik J (1981) Preparation of styrene from ethylbenzene. US Patent
1118 4255599

1119 Yamagishi M, Matsushima H, Wada A, Sakagami M, Fujita N, Ishihama A (1993)
1120 Regulation of the *Escherichia coli rmf* gene encoding the ribosome modulation factor: growth
1121 phase-and growth rate dependent control. *EMBO J* 12: 625–630

1122 Yeh V, Goode A, Eastham G, Rambo RP, Inoue K, Doutch J, Bonev BB (2020) Membrane
1123 stability in the presence of methacrylate esters. *Langmuir* 36: 9649-9657

1124 Yung PY, Grasso LL, Mohidin AF, Acerbi E, Hinks J, Seviour T, Marsili E, Lauro FM
1125 (2016) Global transcriptomic responses of *Escherichia coli* K-12 to volatile organic
1126 compounds. *Sci Rep* 6: 19899

1127 Zecha J, Satpathy S, Kanashova T, Avanessian SC, Kane MH, Clauser KR, Mertins P, Carr
1128 SA, Kuster B (2019) TMT labeling for the masses: a robust and cost-efficient in-solution
1129 labeling approach. *Mol Cell Proteomics* 18: 1468-1478

1130 Zheng M, Wang X, Templeton LJ, Smulski DR, LaRossa RA, Storz G. (2001) DNA
1131 microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen
1132 peroxide. *J Bacteriol* 183: 4562-4570

1133 Zhang YM, Rock CO (2008) Membrane lipid homeostasis in bacteria. *Nat Rev Microbiol* 6:
1134 222-233

1135 Zingaro KA, Nicolaou SA, Papoutsakis ET (2013) Dissecting the assays to assess microbial
1136 tolerance to toxic chemicals in bioprocessing. *Trends Biotechnol* 31: 643-653

1137

Figure legends

Fig. 1. Multi-Omic Based Production Strain Improvement (MOB*psi*): a systematic approach for improving the synthesis of toxic products by cell factories. The diagram shows the MOB*psi* strategy in the form of a decision tree. Standard flow chart symbols are used (rectangles represent processes/experiments; rounded rectangles represent alternative process/start/end points; rhombuses represent datasets; diamonds represent decision points). The details of each MOB*psi* stage are described in the main text. The five classes of candidate genetic interventions are color coded and the same coding is retained in the subsequent figures.

Fig. 2. Interaction of styrene with model lipid bilayers and *E. coli* NST74 membranes. (a) Evolution trajectory frames from a 200 ns all atom MD simulation of 100 styrene molecules in a hydrated membrane patch composed of 75% POPE (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine), 20% POPG (1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)) and 5% CL (cardiolipin) to mimic the composition of the *E. coli* inner membrane. Left to right: frames from Movie S1 1, 21, 63, 100 and 200 ns into the trajectory. Styrene (blue) rapidly phase separates and forms styrene droplets (pink), styrene molecules in the droplets exchange with the water phase, but irreversibly partition with slower kinetics into the membrane interior. The phospholipid P atoms are shown as orange spheres to indicate the location of the bilayer. (b) High resolution natural abundance ^{13}C CP-MAS NMR spectra from *E. coli* NST74 lipid extracts without (bottom) and with addition of natural abundance ^{13}C styrene (top) showing the stoichiometric incorporation of styrene. Inset: styrene carbon numbering and assignment. (c) ^{13}C solid state CP-MAS NMR spectra of whole *E. coli* NST74 cells before and after external addition of ^{13}C-labeled styrene (labels located at S7 and S8).

1162 The spectra were recorded at MAS frequency of 5 kHz and at 8ºC, with 1024 transients
1163 averaged.

1164

1165 **Fig. 3.** Time courses of fed-batch fermentations. (a) *E. coli* NST74 expressing active *pal2-*
1166 *fdc1* from pGS2596. (b) *E. coli* NST74 expressing inactivated *pal2-fdc1* from pGS2597. (c)
1167 *E. coli* NST74 expressing inactivated *pal2-fdc1* exposed to exogenous styrene, which was
1168 added to the cultures to mimic the profile of the styrene production strain shown in (a). The
1169 arrows indicate points (0, 2, 4 and 6 h post-induction) when samples were taken for multi-
1170 omic analyses. Upper panels: optical density (OD_{600}; orange squares); dry cell weight (DCW;
1171 blue circles); viable counts (CFU mL^{-1}; red triangles). Lower panels: styrene (black squares);
1172 glucose (blue diamonds); L-phenylalanine (open red squares); *trans*-cinnamic acid (purple
1173 circles); acetate (brown triangles). The data show the mean values and standard deviations
1174 obtained from biological replicates (n=3); for some data points the error bars are within the
1175 boundaries of the symbols.

1176

1177 **Fig. 4.** Differential regulation of Class 1, Class2 and Class 3 potential genetic interventions to
1178 enhance styrene production. Volcano plots showing changes in, (a) gene expression and (b)
1179 protein abundance, 2 h, 4 h and 6 h post-induction for NST74 pGS2596 (production strain
1180 expressing *pal2-fdc1*) compared to NST74 pGS2597 (negative control expressing inactivated
1181 *pal2-fdc1*). Significant changes in expression/abundance of Class 1 (light brown), Class 2
1182 (green) and Class 3 (blue) genes/proteins are highlighted.

1183

Fig. 5. Schematic showing potential genetic intervention targets in relation to the cell envelope. The following components are shown: outer membrane (OM); peptidoglycan (PG); inner membrane (IM); unsaturated phospholipids (red squiggle in IM). The candidate genetic interventions were grouped into five classes as described in the main text. Class 1 interventions are shown in light brown, Class 2 in green, Class 3 in blue, Class 4 in purple, and Class 5 in orange.

Fig. 6. Identification of cell factory modifications that increase styrene production. (a) Styrene concentration after 72 h growth normalized to culture OD for strains with the indicated gene over-expressed. (b) Styrene concentration after 72 h growth normalized to culture OD for strains with the indicated gene deleted. Styrene was measured for derivatives of NST74 pGS2596 either lacking or over-expressing the indicated genes in batch flask cultures as described in *Materials and methods.* Data shown are the mean values from three independent cultures; the error bars show the standard deviation from the mean. Strains exhibiting statistically significant greater styrene production for the test strains compared to the relevant control (dashed line; NST74 pGS2596 transformed with an empty DVK plasmid and an NST74 pGS2596 *lacZ* mutant for the knockout strains) were calculated using ANOVA **** $p<0.0001$, *** $p<0.001$. Bars are color coded for the different classes of genes tested (Class 1 light brown, Class 2, green, Class 3 blue, and Class 4 purple). ^ indicates gene cloning unsuccessful; [§] indicates genes expressed from the weak promoter J23103; [#] indicates essential genes.

Fig. 7. Effects of combining genetic interventions on the production of styrene. (a) Styrene concentration after 72 h growth normalized to culture OD for the indicated strains. (b)

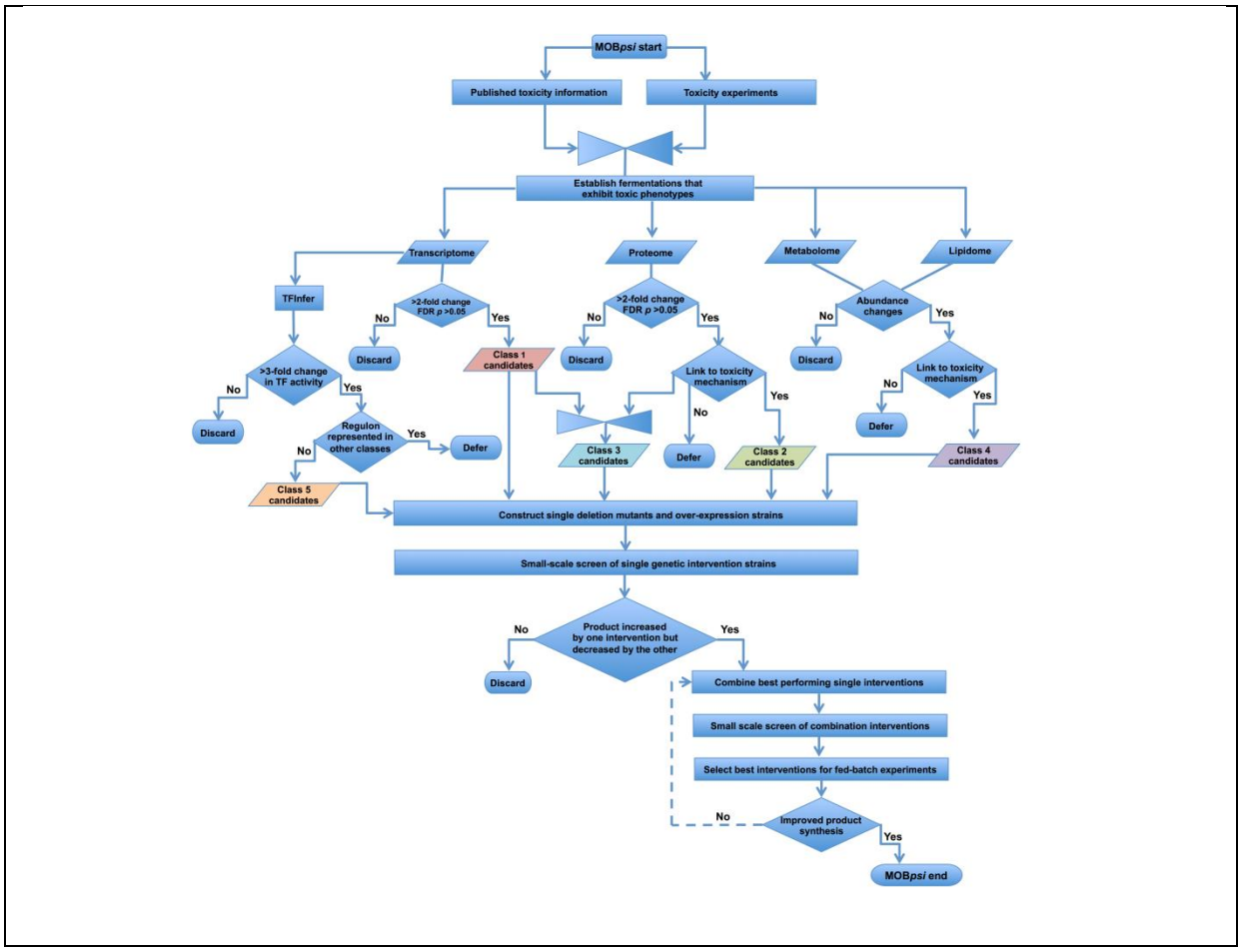
1208 Phenylalanine concentration after 72 h growth normalized to culture OD for the indicated
1209 strains. Proton NMR measurements were made on samples from flask cultures of the
1210 indicated strains as described in the *Materials and methods*. Data for over-expression of
1211 *cpxP*, *rmf* or *yebE* in the *E. coli* NST74 *ΔaaeA* pGS2596 background are mean values from
1212 triplicate cultures. Data for the gene-doctored NST74 *ΔaaeA* (*ΔaaeA Δldh::cti-pssA*) host
1213 strain and the corresponding control (*ΔaaeA Δldh::kan*) are mean values from six cultures. In
1214 both cases the error bars show the standard deviation from the mean.

1215

1216 **Fig. 8.** MOB*psi* re-engineered *E. coli* NST74 results in increased styrene production and cell
1217 factory viability. (a) Radial basis function model of styrene production for *E. coli*
1218 NST74Δ*aaeA* expressing active *pal2-fdc1* from pGS2596 compared to the starting strain. (b)
1219 Radial basis function model of styrene production for *E. coli* NST74Δ*aaeA* over-expressing
1220 *cpxP* and active *pal2-fdc1* from pGS2596 compared to the starting strain. (c) Post-induction
1221 cumulative specific styrene productivity profiles for re-engineered NST74 cell factories. (d)
1222 Post-induction cumulative specific glucose consumption for re-engineered NST74 cell
1223 factories. (e) Post-induction cumulative specific phenylalanine productivities for re-
1224 engineered NST74 cell factories. (f) Post-induction viability (colony forming units) for re-
1225 engineered NST74 cell factories. (g) Post-induction pGS2596 plasmid stability during re-
1226 engineered NST74 fermentations. The different cell factories are color coded as follows: *E.*
1227 *coli* NST74 expressing inactive *pal2-fdc1* from pGS2597 (negative control, black); *E. coli*
1228 NST74 expressing active *pal2-fdc1* from pGS2596 (positive control, blue), *E. coli*
1229 NST74Δ*aaeA* pGS2596 (green); and *E. coli* NST74Δ*aaeA* $cpxP_o$ pGS2596 (red).
1230 Measurement data are shown as symbols (squares). Radial basis function model fits are
1231 shown as solid lines (panels a and b, with upper (dotted lines) and lower (dashed lines) 95%

confidence limits. The error bars shown for the cumulative specific production data (panels c, d and e) represent 95% confidence limits estimated from Monte Carlo simulations (Fig. S4). For the viability data (panel f) the error bars are standard deviations from the mean of technical replicates (n=3). The plasmid stability data (panel g) are from a single experiment and the error bars represent the standard deviation of triplicate technical replicates. Some error bars sit within the boundaries of the symbols.

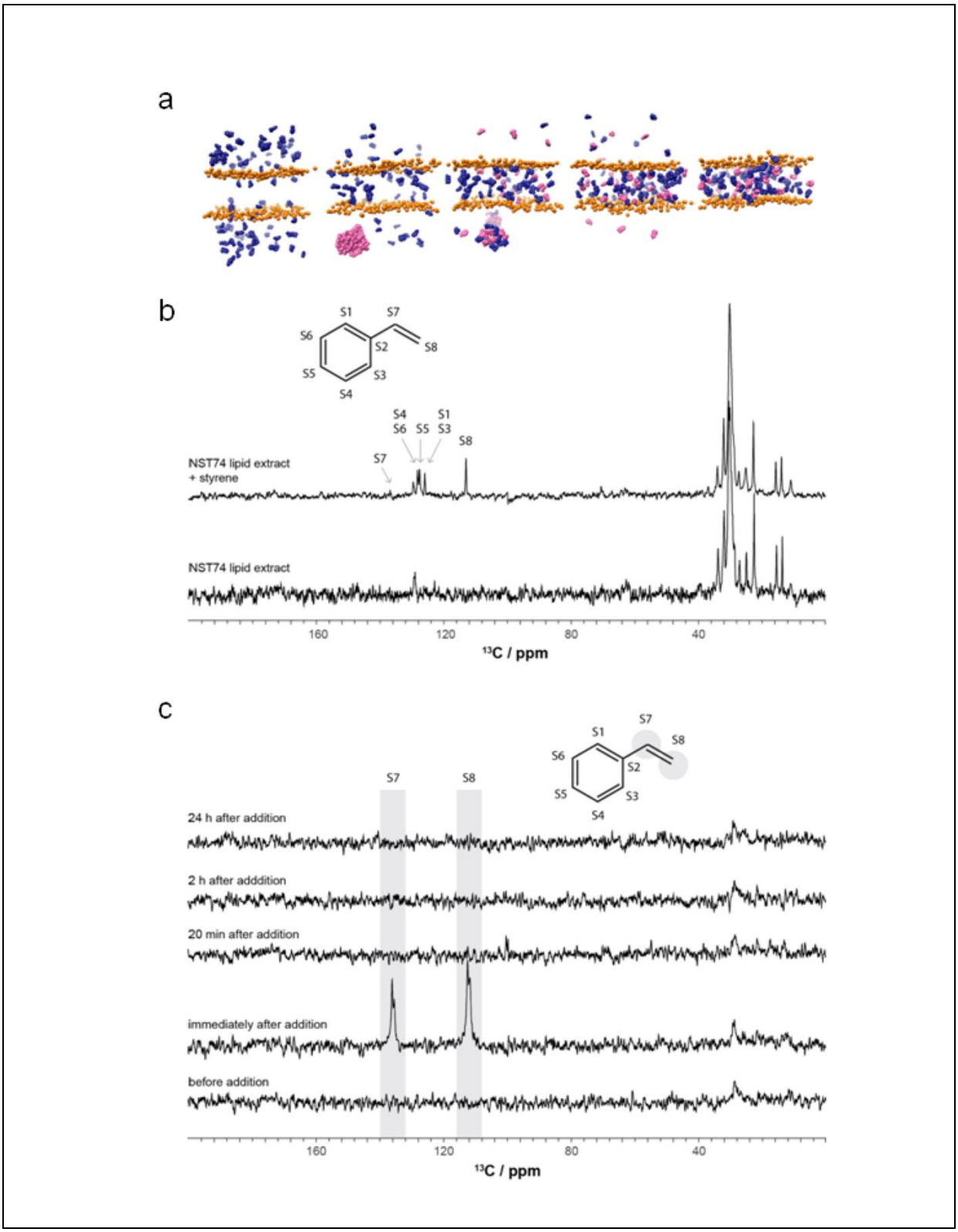
1239 Figure 1



1240

1241

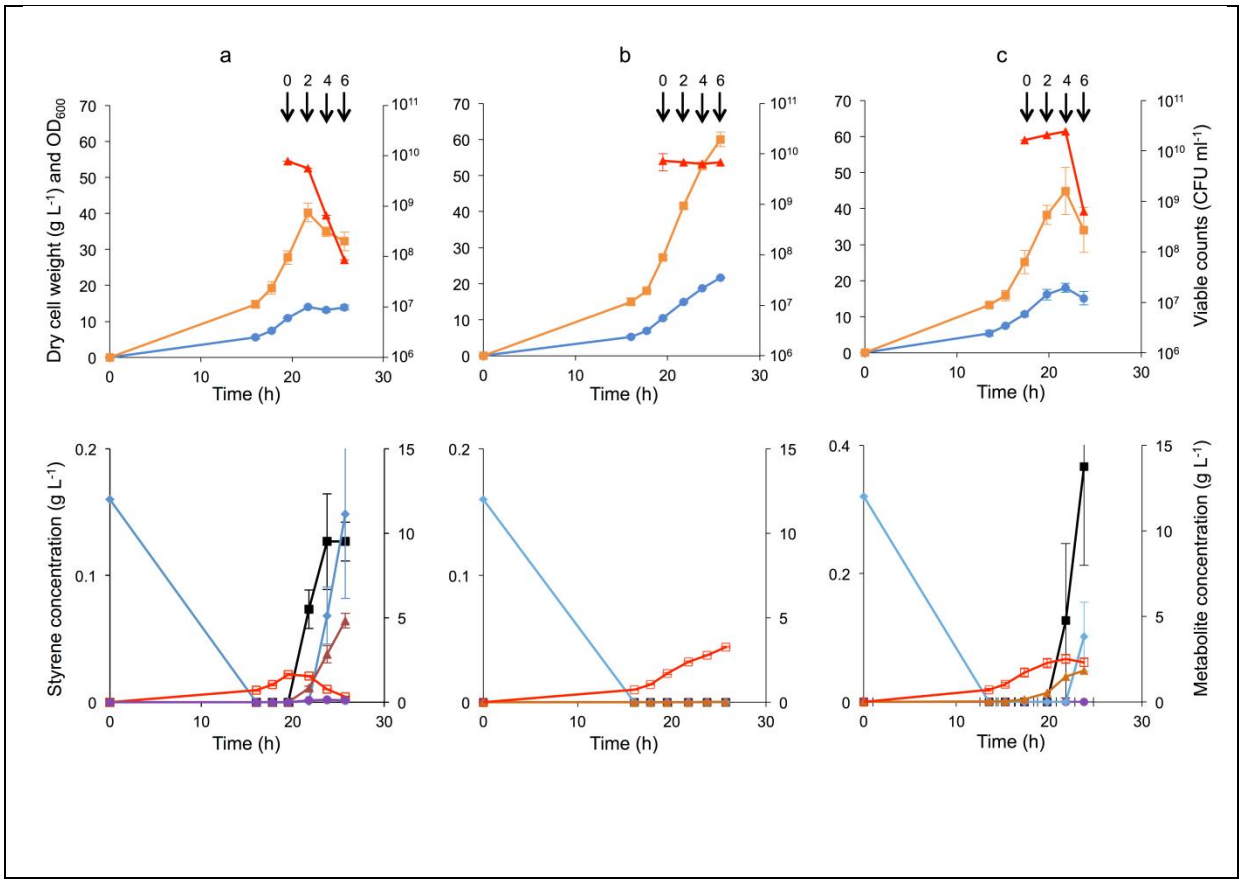
1242 Figure 2



1243

1244

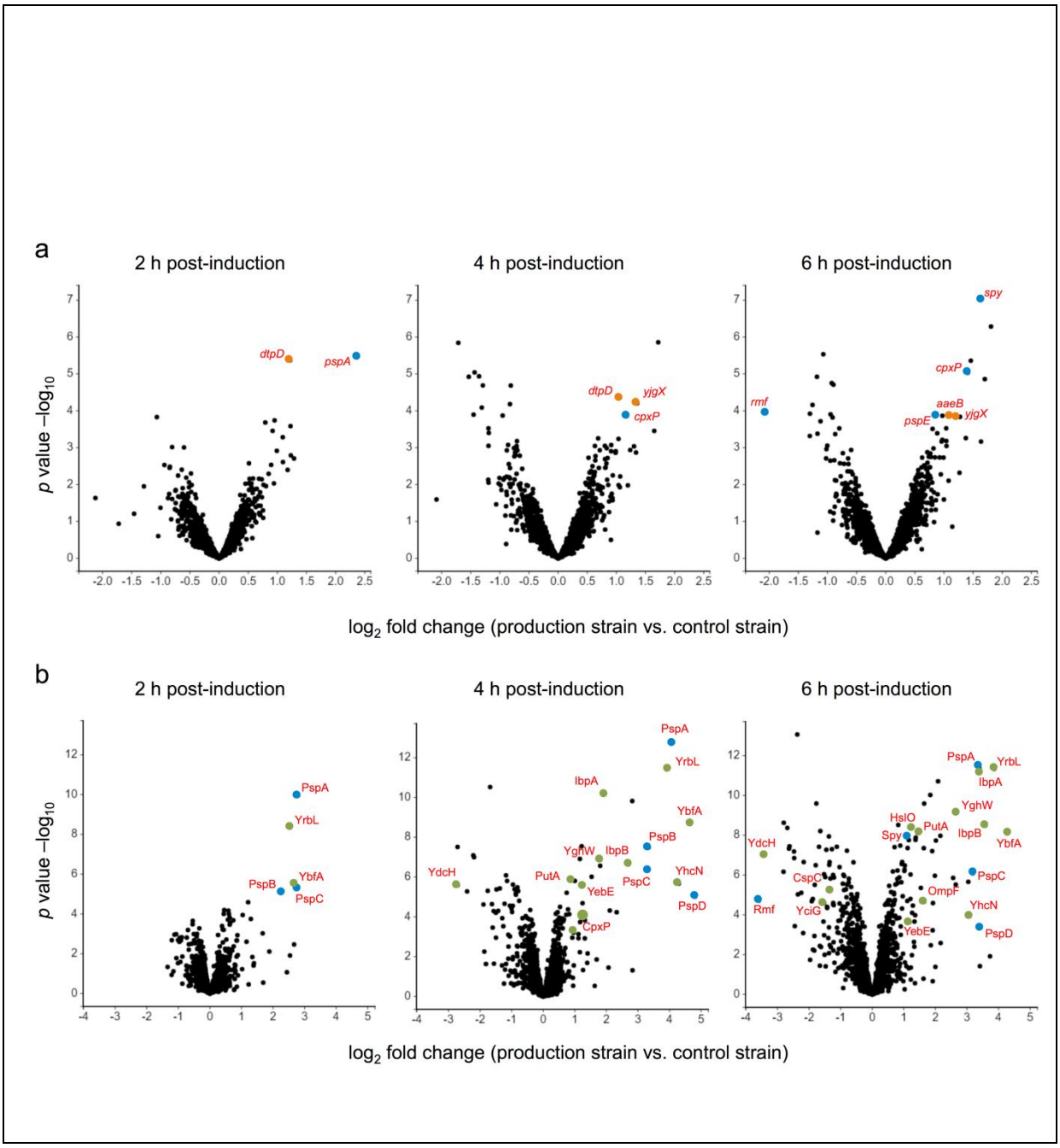
1245 Figure 3



1246

1247

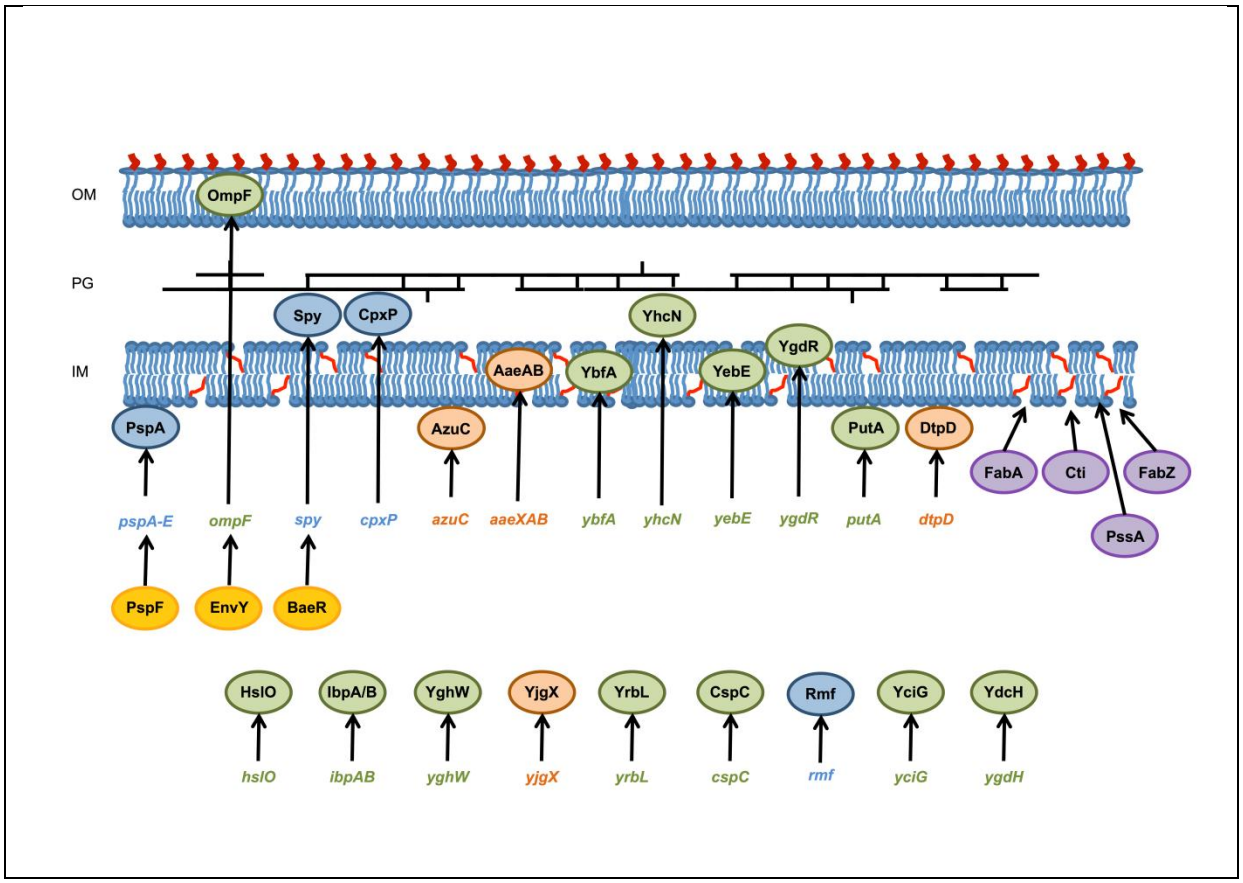
1248 Figure 4



1249

1250

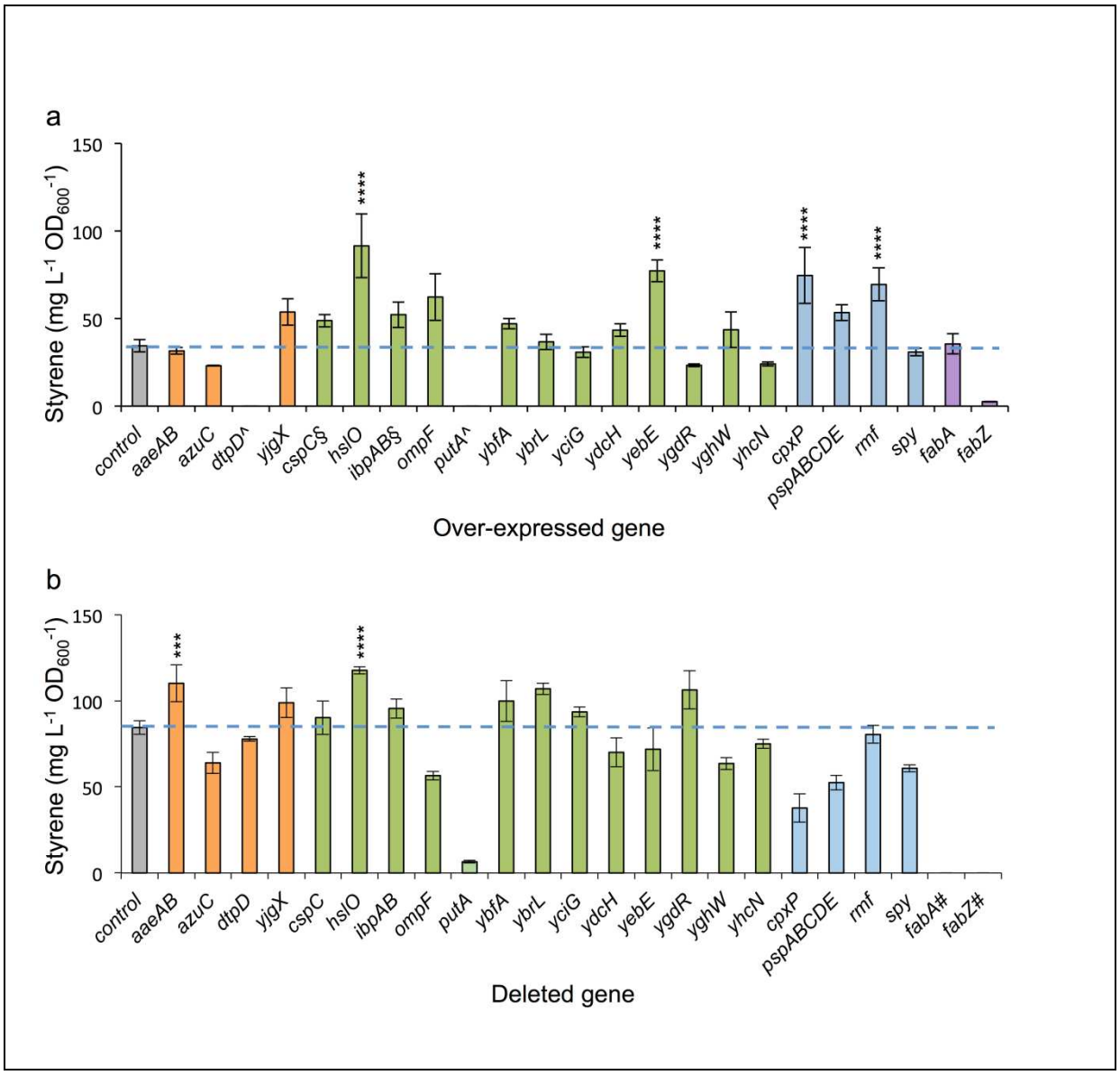
1251 Figure 5



1252

1253

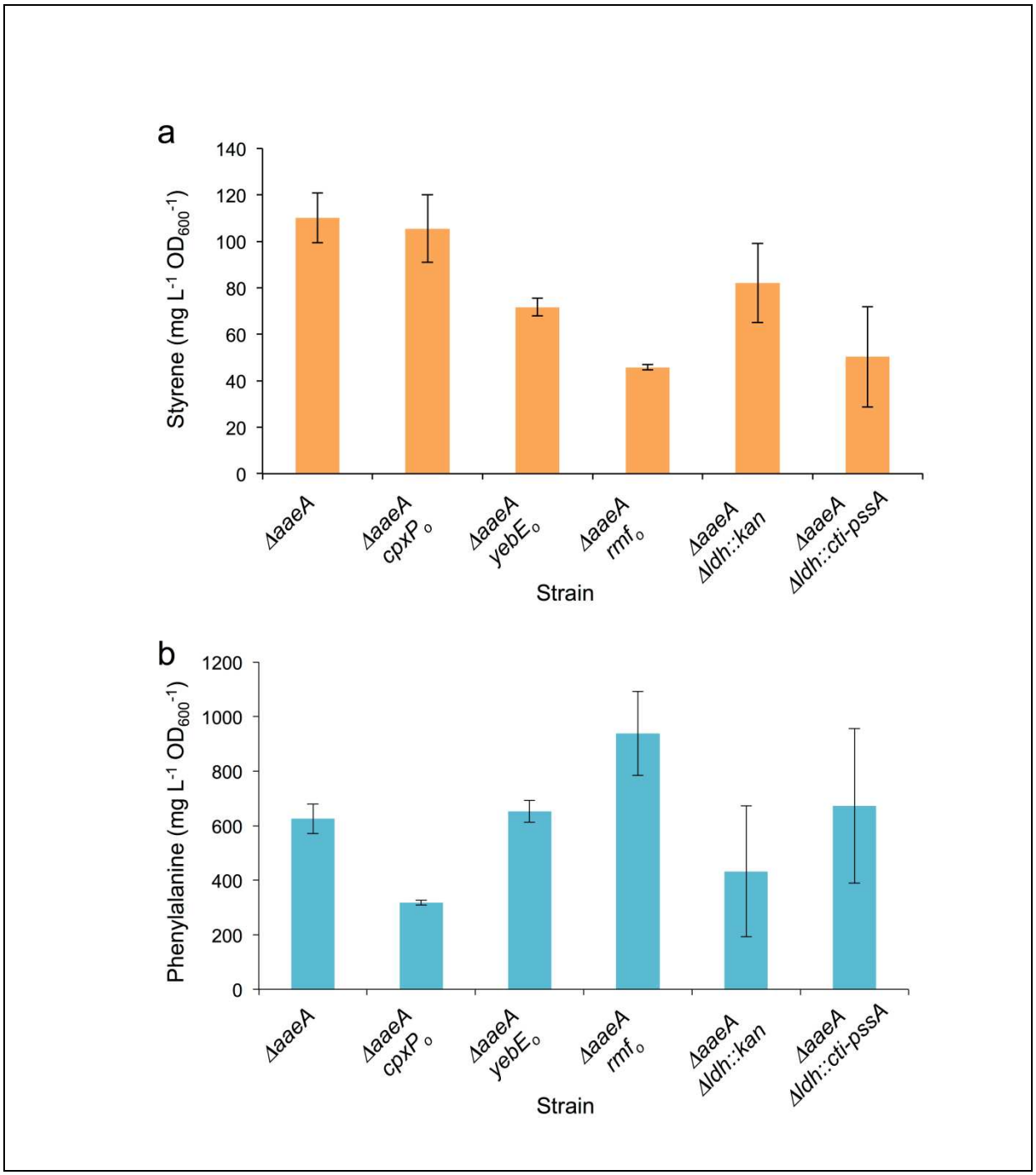
1254 Figure 6



1255

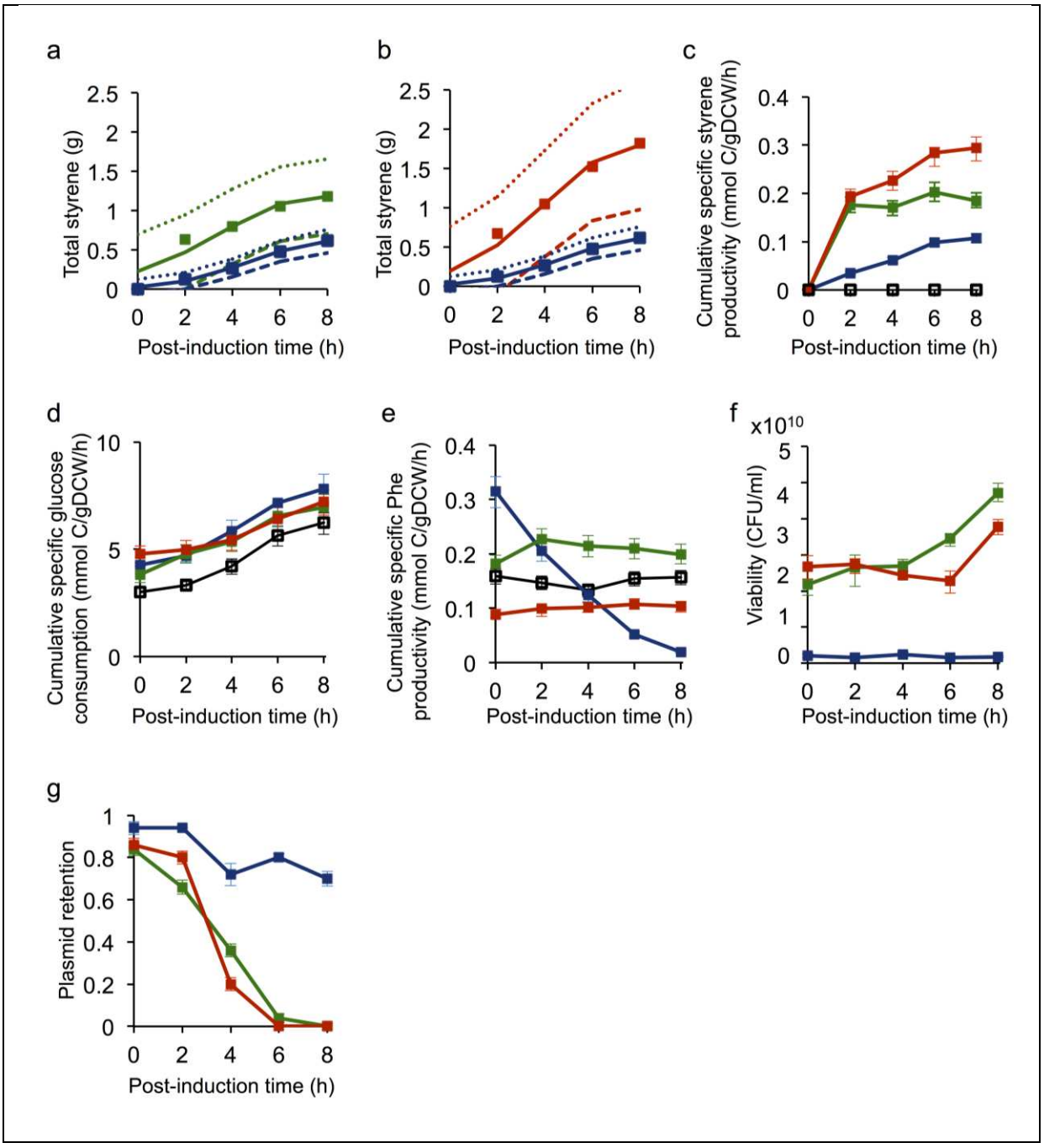
1256

1257 Figure 7



1258

1259 Figure 8



1260