



## 23 **Scaling the Bioeconomy**

24 The natural world synthesizes an impressive array of chemical structures that can be used as  
25 medicines, solvents, materials, and fuels [1–3]. Advances in synthetic biology have dramatically  
26 reduced the cost of DNA synthesis and expanded capability to construct vectors in high  
27 throughput. These tools allow researchers to rapidly test metabolic engineering strategies [4] — to  
28 the point that, with a concerted effort, almost any molecule can be synthesized in small  
29 concentrations at laboratory scale [5]. To tackle global challenges such as climate change, focus  
30 must now turn to translating this synthesis potential to industrial scales such that sustainable  
31 alternatives to the modern petrochemical industry can be established [6]. Biomanufacturing, the  
32 use of biotechnology to synthesize chemical products, has been demonstrated on industrial scales  
33 for high-value products (e.g., therapeutic proteins, enzymes, and antibiotics) and low-value  
34 commodities (e.g., ethanol, lactic acid, amino acids, and sweeteners) [7], but efforts to  
35 commercialize many other attractive products have failed [8]. Scale-up, the process of translating  
36 laboratory processes (<10 l) to commercial-sized volumes (>100,000 l), is commonly regarded as  
37 a major risk for new bioprocesses in development [9]. To succeed at scale, a biocatalyst must  
38 support conversions that approach theoretical yields (to minimize feedstock costs), thrive in the  
39 presence of toxins (including desired products), reliably progress through the seed train without  
40 losing productivity, and overcome environmental heterogeneities that are easily avoided in  
41 laboratory-scale reactors. Notably, scale-up failure does not necessarily consider the entire  
42 breakdown of microbial performance in large-scale. Instead, economic failures often occur that  
43 render original break-even points non-realistic. Some examples include smaller production  
44 volumes required to prevent intensive foaming, reduced final product titers, and increased by-  
45 product concentrations that challenge downstream processing. Each of these issues can be studied

46 on an industrial scale, but large-scale experiments are rare due to costs and access to appropriate  
47 equipment. Instead, researchers use specialized equipment and/or combinations of experiments to  
48 simulate, as closely as possible, industrial scale conditions in a laboratory environment. These  
49 experiments help reduce the risks associated with developing novel processes. This review  
50 summarizes what has been learned from studying and engineering industrial microbes grown under  
51 conditions that simulate large-scale environments. We review methodologies for industrial strain  
52 design considering how cells respond to simulated industrial stresses and strategies to minimize  
53 culture heterogeneity. For simplicity, this review will focus on examples from the Gram-negative  
54 bacteria *Escherichia coli*, however the methodologies discussed here can be used to study other  
55 biomanufacturing hosts.

## 56 **The Challenges of Bioprocess Engineering at Scale**

57 The challenges facing industrial cultivation can be separated into three types of limitations,  
58 physical, chemical, and biological. Physical limitations stem from the inability to match the  
59 characteristic mixing times of laboratory reactors without enormous power inputs at industrial  
60 scale. Chemical limitations result from changes in nutrient sources such as water or carbon that  
61 may differ when a process is scaled. Finally biological limitations signify the cellular response to  
62 both physical and chemical limitations but also the effects of industrially specific cultivations such  
63 as seed train growth, which increase the number of cell generations before production. Due to its  
64 accessibility, academic research has primarily focused on the physical limitations of increased  
65 mixing time. The effects of which are microenvironmental gradients in chemical composition  
66 (nutrients, media sterilization, pH, aeration) and physical properties (temperature variation, shear),  
67 which impact the biological performance (growth rate, uptake rate, productivity, viability, stress  
68 response) of the cells being cultured [10]. It should be noted that this review focuses on bioreactor

69 specific challenges, but chemical and physical challenges can additionally affect processes before  
70 and after the bioreactor culture (e.g., media sterilization and product isolation). The  
71 inhomogeneous environment in bioreactor cultivations leads to heterogeneity in strain behavior,  
72 often with poorer performance from cells transiently passing through non-ideal conditions. The  
73 reduced performance can be attributed to direct impacts of the stressful environments and to  
74 inefficient resource allocation, i.e., where cells use feedstock for unwanted activities instead of  
75 maximizing product synthesis [9]. It is the task of the metabolic engineer to design cells that  
76 optimize resource allocation while maintaining stable cultures that grow reliably.

77 Resource allocation has been modeled at genome scale [11,12], but perhaps is more easily  
78 illustrated by simple equations such as the Pirt model of substrate utilization. Here, substrate  
79 consumption is separated into three categories: substrate for making new cells, substrate for  
80 making product, and substrate for driving cellular systems and functions (i.e., maintenance energy)  
81 (**Figure 1**) [13]. Metabolic engineers spend most of their time designing pathways that maximize  
82 specific productivity,  $q_p$ , and specific yield of product ( $Y_{P/S}^{max}$ ). They do so by identifying the  
83 highest yielding biochemical pathways [14], bioprospecting or engineering enzymes to have high  
84 specific activity [15], balancing enzyme activity across a pathway [16,17], and providing the  
85 correct supply of ATP and required cofactors [18]. Biochemical engineers then design cultivation  
86 strategies (e.g., fed-batch) that often consist of a “growth” phase to build up biomass followed by  
87 a secondary “production” phase for product accumulation. Maximum specific yields of product  
88 are achieved when the fractions of substrate used for biomass  $\left(\frac{1}{Y_{XS}^{max}} \mu\right)$  and maintenance ( $m_s$ ) are  
89 minimized. To maintain cell growth within the physical limitations of mass and heat transfer at  
90 scale, often cell growth is restricted via nutrient (e.g., phosphate, carbon, or nitrogen) limitations  
91 [19]. However intentionally limiting a nutrient to halt or reduce cell growth can also alter core

92 metabolism and flux through heterologous pathways, possibly benefiting and/or harming a  
93 bioprocess. A good example is a phosphate limited environment, where native ATP synthase  
94 enzymes will import protons even without sufficient phosphate to regenerate ATP. As a result,  
95 ATP production is decoupled from proton import leading to reduced energy efficiency and reduced  
96 substrate yield. To regenerate the needed ATP, the cell will maintain an increased substrate uptake  
97 rate and high glycolytic activity independent of cell growth. This higher glycolytic flux may  
98 benefit a bioprocess if the goal is maximizing productivity. However, this increased metabolic  
99 activity also leads to accumulation of NADH. To then maintain a redox balance, the cell will  
100 compensate with overflow metabolites and utilize less efficient redox regeneration pathways,  
101 further reducing yield [20]. While these feeding strategies are developed in laboratory scale  
102 reactors where mixing is efficient, cells in large-scale reactors experience zones of nutrient  
103 depletion, pH extremes, hypoxia, and temperature swings. The naturally evolved stress responses  
104 cells used to survive these conditions expend additional resources (elevating maintenance energy)  
105 unnecessarily as stresses are often short-lived (relative to the timescale of true cellular impact).

## 106 **Scale Down Reactors: A View into Industrial Fermentations and the Stress Response**

107 Despite the maturity of the biomanufacturing industry, local measurements of all relevant culture  
108 variables (pH, dissolved oxygen, glucose, etc.) can only be achieved with traditional wall mounted  
109 sampling ports and probes. While necessary, these static sampling locations can only view a  
110 fraction of the whole culture volume, offering a glimpse of the culture along only the wall of the  
111 reactor. Instead, **Computational Fluid Dynamics (CFD)** (see Glossary) are used to model the  
112 complex mixing regimes within reactors. These simulations provide insight into both the timescale  
113 and the intensity of the nutrient gradients and can be validated with flow-following *in situ* probes  
114 [21–23].

115 With the combined knowledge of large-scale mixing and modelling results, laboratory-scale  
116 reactors can then be designed to mimic poor mixing, providing an industrial like environment to  
117 evaluate strains before scaling up (**Figure 2**). These systems, referred to as **Scale Down Reactors**,  
118 are commonly based on stirred tank reactors (STRs) and can be operated in either a chemostat or  
119 fed-batch mode. To simulate poor mixing, a baseline STR is connected via pumping to additional  
120 STRs or plug flow reactors (PFRs). This allows a fraction of the total culture to experience a  
121 different environment (varying mixing, nutrient starvation, aeration, pH, temperature, etc.) before  
122 returning to the main reactor. Each perturbation is intended to model transient stresses observed in  
123 large-scale reactors. Therefore, the residence time in the “stressed” zone is set to match the typical  
124 timescale experienced in the industrial reactor, often on the order of minutes. Sampling ports set  
125 along the PFR allow for collection of functional genomics samples to explore the timing of gene  
126 expression changes in response to the environmental challenge, in addition to the average response  
127 in the base STR. Stresses within industrial reactor are often a distribution of both intensity and  
128 time exposed [24]. To explore shorter timescales, dynamic microfluidic systems have been

129 developed to mimic the macroscale systems described previously. These tools allow studying  
130 cellular dynamics on the order of seconds, but future applications are contingent on improved  
131 quantification sensitivity for low metabolite concentrations at low volumes [22,25–27].

### 132 **Environmental Challenges and Cellular Response**

133 Nutrient starvation is a common stress experienced by cells in zones where fresh nutrient supply  
134 is substantially slower than local specific consumption rates. Nutrient depletion triggers the  
135 **stringent response**, a coordinated metabolic and regulatory program that arrests growth, slows  
136 nutrient utilization, and induces changes in gene expression that allows the microorganism to  
137 survive in the now depleted environment [28]. During a stringent response, *E. coli* translation rates  
138 slow leading to elevated synthesis of the alarmones guanosine tetraphosphate (ppGpp) and  
139 guanosine pentaphosphate (pppGpp) [29], referred to hereafter as (p)ppGpp. These alarmones  
140 alter the transcription of a wide regulon including ribosomal proteins and enzymes involved in  
141 DNA replication, nucleotide synthesis, transcription, ribosome maturation and function, and lipid  
142 metabolism [28,29]. The stringent response is of particular importance for both carbon (**BOX 1**)  
143 and nitrogen (**BOX 2**) limitations, ultimately controlling how the cell throttles its metabolism to  
144 prevent the build-up of unnecessary metabolites and wasteful consumption of ATP [29]. When the  
145 stringent response is triggered in cells entering transient starvation zones, the altered regulation  
146 can slow nutrient uptake, downregulate desired pathway enzymes, and expend energy expressing  
147 unneeded proteins – all of which reduce specific productivity and/or product yield [10]. Although  
148 not only relevant for industrial scale reactors, the metabolic cost of heterologous gene expression  
149 can also lead to stress responses. Additional costs to the cell will also depend on the solubility of  
150 the heterologous protein (**BOX 3**).

151

152 Other industrially relevant stresses potentially stem from increased gas solubility, pH variability  
153 and impeller sheer, with the significance of each dependent on the product and/or host of interest.  
154 Increased gas solubility is a result of increased hydrostatic pressure in industrial reactors, leading  
155 to increased gas transfer rates for both oxygen and carbon dioxide [30,31]. To be clear, increased  
156 hydrostatic pressure is not likely to affect *E. coli*, but instead higher concentrations of oxygen and  
157 carbon dioxide can lead to stress responses (**BOX 4**) [32]. Though pH is thought to not  
158 significantly change within large-scale reactors, acidic product generation, and concentrated  
159 alkaline addition ports can lead to local pH fluctuations [33,34]. Scale down experiments suggest,  
160 however, that there is little effect of pH shifts when considering *E. coli* as a host for plasmid  
161 production [34,35]. Finally, the action of impellers in industrial stirred tanks can create zones of  
162 high shear. Most microbes have robust cell walls and are not impacted significantly by shear. For  
163 example, high-shear bioreactor environments have been shown to strip the outer polysaccharide  
164 layer from *E. coli* but have had no further effect on metabolism [36]. Organisms which impeller  
165 sheer can lead to morphological changes, like filamentous fungi, may lead to decreased  
166 productivity but direct correlations have not been noted in every case [37]. All considered, the  
167 negative effects of increased pressure, pH variability, and sheer can vary greatly on the organism  
168 and product of interest and should be evaluated for each scenario.

## 169 **Optimization of the Stress Response**

170 The cellular stress responses induced during large scale fermentations leads to increased  
171 maintenance energy. While some responses are required to prevent irreversible damage (i.e., ROS  
172 and DNA damage), other responses can be overly costly and could be optimized to promote cell  
173 survival rather than growth [38]. However, the cellular response is likely organism, stress, and  
174 time dependent, as cells exposed to repeated short term stress can optimize their stress response  
175 for growth [39]. Adjusting the stress response provides an opportunity to increase available  
176 resources, potentially leading to improved growth and improved heterologous protein production  
177 in transiently stressed environments. Understanding which response can be tuned or removed is a  
178 matter of rational engineering, random screening, or both.

179 When considering major regulatory nodes in response to stress, (p)ppGpp is a clear target for  
180 controlling a wide range of metabolically expensive reactions. If the level of (p)ppGpp never  
181 increases in response to stress, then the stringent response will be avoided. (p)ppGpp synthesis is  
182 primarily controlled by the (p)ppGpp synthase *RelA* and the (p)ppGpp synthase/hydrolase *SpoT*,  
183 which in response to nutrient stress regulate (p)ppGpp levels [40]. However complete removal of  
184 *RelA* and *SpoT* creates a (p)ppGpp null strain reported to have amino acid auxotrophies and  
185 increased sensitivity to antibiotics [41,42]. This report suggests that a baseline concentration of  
186 (p)ppGpp is required for robust cellular growth and may be even more important in stressed  
187 conditions. Modulating (p)ppGpp levels however has also led to significant growth repression in  
188 minimal media with either artificially low or high ppGpp. High (p)ppGpp was found to limit  
189 ribosome synthesis and low (p)ppGpp reduced expression of catabolic and anabolic proteins,  
190 suggesting an intermediate (p)ppGpp concentration is ideal for optimal growth [43]. In one  
191 example *E. coli* was engineered for increased glucose uptake in nitrogen limited cultivations by

192 maintaining low levels of (p)ppGpp [44]. By adjusting the (p)ppGpp synthase activity of *spoT*  
193 through two-point mutations, and the deletion of *relA*, (p)ppGpp synthesis was maintained at a  
194 stable level throughout a nitrogen limited STR cultivation. An additional point mutation in *aceE*  
195 reduced its activity, leading to pyruvate as the terminal product [44]. The resulting strain  
196 consumed glucose three times as fast as the parent on a per-cell basis, leading to its name (high  
197 glucose throughput - HGT), but had a 40% lower maximum specific growth rate under nitrogen  
198 limited batch condition. This decreased growth rate speaks to the careful engineering required to  
199 adjust the stress response pathways without significantly affecting cell growth [44]. *E. coli* SR, a  
200 derivative of the HGT strain lacking the *aceE* mutation, displayed a muted stress response to  
201 repeated nitrogen limitation in an STR-PFR cultivation. Along the PFR, (p)ppGpp levels remained  
202 relatively low, however *E. coli* SR maintained nitrogen assimilation pathways during the  
203 cultivation. The diminished nitrogen stress response led to a 45% energy savings with the same  
204 biomass yield, suggesting that SR has additional energy which could be utilized for heterologous  
205 pathways and products [45].

206 With knowledge of the metabolic and regulatory nodes critical for glucose and nitrogen  
207 limitations, engineers have been able to modify nutrient uptake to maintain desired process  
208 parameters. For example, during nitrogen limitation the glucose transporter PtsI is normally  
209 inhibited by  $\alpha$ -ketoglutarate accumulation. When engineers overexpressed PtsI, the resulting strain  
210 had a four-fold increase in metabolic rates in nitrogen limited cultivations. As a result, the nitrogen  
211 limited, and non-growing culture had increased fatty alcohol yield [46]. In another example, strain  
212 engineers artificially limited glucose uptake and catabolic flux by removing glucose transporters  
213 and components of the phosphoenolpyruvate sugar phosphotransferase system (downstream  
214 glucose metabolism) to minimize overflow metabolism. These mutations allowed a high glucose

215 concentration in the batch media to mimic a fed batch culture [47]. A similar strategy could be  
216 used at scale to avoid nutrient heterogeneity due to the typical high concentration glucose feeds.

## 217 **Minimizing Heterogeneity and Cellular Adaption**

218 Titer and yield evaluation often require a sufficient volume and analyte concentration to be  
219 measurable, demanding larger volumes of culture that contain millions of individual cells.  
220 Therefore, the final titers reported represents an average output from what is likely a distribution  
221 of heterogenous cellular production phenotypes [48,49]. This **Cell Culture Heterogeneity** stems  
222 from natural or stress induced cellular adaptation where cells attempt to improve their fitness  
223 through genetic, transcriptional, or translational modifications and errors. Adaptive laboratory  
224 evolution (ALE), for example, takes advantage of DNA replication errors to isolate strains with  
225 improved growth and/or fitness. Conversely, within a bioprocess, it is preferred that **Cell Culture**  
226 **Homogeneity** is maintained such that strains do not adapt and perform consistently throughout the  
227 cultivation. Cellular adaptation is a larger concern at scale due to the large number of generations  
228 needed in the seed train to grow sufficient biomass for the final reactor. Industrial strains are grown  
229 far past the number of generations typically considered in laboratory cultivations, and therefore  
230 have greater potential to accumulate genetic errors and/or adapt [50].

231 During growth, *E. coli* is thought to modulate the diversity (and thus adaptability) of the whole  
232 culture by balancing the error rate of the DNA replication with the potential for both beneficial  
233 and harmful genetic errors [51,52]. This adaptation strategy is accelerated during the **SOS**  
234 **response**, where *E. coli* is known to express the more mutation prone DNA polymerase IV and V.  
235 This expression leads to increased SNPs or insertion sequences (IS) to find a favorable mutation,  
236 even at the risks of altering an essential gene [53,54]. As a result, there is an increase in  
237 heterogeneity, and thus overall survivability, of the culture. Industrial strains, however, may  
238 consequently lose heterologous genes when adapting to stressed conditions. One study found that  
239 when these error prone polymerases were removed, *E. coli* retained expression of a toxic

240 methyltransferase with less SNPs over time compared to a wildtype control [55]. When studying  
241 plasmid-based production of 2,3-butanediol and mevalonic acid in multiple background strains,  
242 genetically mobile insertion sequences and SNPs were found to disrupt plasmid expression of the  
243 heterologous enzymes over 90+ generations [50]. Removal of both error prone polymerases and  
244 native IS in the genome reduced MDS42 strain resulted in a longer production half-life over 89  
245 generations. However, MDS42 had lower final titer relative to wildtype, suggesting one of the  
246 modifications was not beneficial to mevalonate production [56]. While genetic mutations offer a  
247 view of long-term adaptation through replication, the transcriptional error rate also varies based  
248 on gene expression level and cell fitness. More highly expressed genes with strong fitness selection  
249 are transcribed with lower error rates in *E. coli* [57]. Recent research highlighted that the  
250 translational error rate of prokaryotes is 3-4 times greater than transcription error rates [58]. In  
251 unstressed conditions, translational errors are typically controlled based on the frequency of gene  
252 expression with more highly expressed genes having less frequent errors [59]. However  
253 translational error rates, like replication error, may even be intentionally increased in stressed  
254 environments to trigger an SOS response [60,61]. When constructing strains for use in industrial  
255 processes, it is important to consider the potentially unnecessary effects of errors in replication,  
256 transcription, and translation. These errors could be further reduced by improving error prone  
257 enzymes, for example modifications to DNA pol III HE [53].

258 Considering long term effects of adaptation, the more burdensome a heterologous gene is, the more  
259 likely an inactivating mutation will lead to a growth benefit. Therefore, if burden is minimized,  
260 deleterious mutations may be less likely. This line of thinking has led to methods to improve  
261 heterogeneity by modulating the stress of heterologous expression through “host aware” strategies  
262 which adjust expression based on available cellular resources [62]. In one experiment, a protein

263 stress promoter drove expression of dCas9, creating a feedback loop for controlled expression of  
264 mCherry. The result was a strain which produced mCherry based on a cell's protein synthesis  
265 capacity; leading to more biomass and higher mCherry titer [63]. This system requires further tests  
266 in fluctuating conditions or long-term cultures to prove general robustness. Other process driven  
267 strategies to reduce cellular burden such as oscillating heterologous expression with external  
268 inducers and separation of growth/production phases have been reviewed in detail elsewhere [64].

269 An alternative method to increase homogeneity of cultures takes advantage of expression noise  
270 through metabolic coupling. Also known as synthetic addition, this strategy couples a product  
271 detecting biosensor or native product sensing transcriptional regulator with expression of an  
272 essential or beneficial gene for cell growth. As a result, product synthesis confers a fitness benefit.

273 The first demonstrations of this methodology explored producing free fatty acids and tyrosine. In  
274 the free fatty acid producing strains, the authors coupled production to a tetracycline resistance  
275 marker; creating a selection pressure for fatty acids by supplementing the antibiotic to the growth  
276 media. Using this approach, a greater proportion of cells were classified as high fatty acid  
277 producers. Coupling the fatty acid production pathway to leucine biosynthesis in an auxotrophic  
278 *E. coli* (instead of an antibiotic resistance marker), led to 21.5g/L FFA in a fed batch cultivation,  
279 a greater than four times improvement compared to a strain without metabolic coupling [49].

280 Iterative flask cultivations suggested improvements were due to non-genetic variations. When high  
281 producing cells were isolated and returned to fresh media without selection, the enhanced product  
282 titer was not reproduced, and genome sequencing did not suggest any beneficial genetic mutations  
283 [49]. Analogously, coupling essential gene expressions to a mevalonate biosensor a synthetically  
284 addicted strain could maintain mevalonate production over 85 generations. The un-addicted strain  
285 by comparison had reduced mevalonate production due to the accumulation of insertion sequences

286 (IS) and single nucleotide polymorphisms (SNP) in the heterologous pathway [65]. These studies  
287 establish that coupling a fitness benefit to flux through heterologous pathways may both take  
288 advantage of non-genetic gene expression heterogeneity machinery and reduce deleterious genetic  
289 modifications in the heterologous pathway. While the fitness benefit offers a clear reason for  
290 genetic stability of heterologous pathway, it is less clear what the non-genetic improvements were.  
291 Additionally, it is unknown how these addiction strategies function in oscillating conditions, thus  
292 warranting further research.

293 **Minimize basal maintenance energy:**

294 *E. coli*'s robust adaptability is a result of a proteome that anticipates many possible limitations  
295 [66]. However, if the anticipated environment never occurs, then the additional unnecessary  
296 maintenance cost could potentially divert substrate flux away from the intended production  
297 pathway (**Figure 1**). Furthermore, given the total protein concentration is relatively constant across  
298 conditions [29], preventing unnecessary protein production may increase resource availability for  
299 expression of other proteins, including heterologous enzymes used in industrial strains. One  
300 method to create efficient strains is genome reduction, which aims to create a strain with the  
301 absolute minimum number of genes. However, this methodology often results in a strain with  
302 significantly reduced growth rates, which is unfit for industrial applications [67]. More industrially  
303 relevant strategies have instead focused on gene removal without significantly affecting growth.  
304 However, only two strains (one *E. coli* and one *P. putida*) have shown improved growth and  
305 production parameters in scale down reactors when compared to wildtype [67]. In general,  
306 industrial strain design is centered around strategies to create and sort a diverse library of  
307 mutations. These libraries can either be created with more rational approaches to remove both  
308 costly and potentially unnecessary genes informed by functional genomics or randomly through  
309 strategies like random bar code transposon-site sequencing (RB-TnSeq) or adaptive laboratory  
310 evolution (ALE).

311 In an example of the rational engineering strategy, unessential gene targets were removed if they  
312 had high maintenance cost under glucose limited conditions, were also orthogonal to central  
313 metabolism, and did not affect major regulatory elements [68]. The resulting strain (RM214) had  
314 a reduced maintenance coefficient in carbon limited STR-PFR chemostats but retained similar  
315 growth parameters in STR chemostats. In carbon limited STR-PFR cultivations RM214

316 additionally reached a 44% higher eGFP yield, likely due to excess energy available from the lower  
317 cellular maintenance. Upon PFR connection, the fraction of high eGFP producing cells decreased  
318 for both MG1655 and RM214, however RM214 maintained a higher proportion of producing cells  
319 over time. So, an additional benefit of reduced maintenance was also improved culture  
320 heterogeneity in carbon limited environments (See **Minimizing Heterogeneity and Cellular**  
321 **Adaption**) [68]. Similar studies have suggested that relief of heterologous production load can  
322 reduce escape rate and could improve stability of industrial strains over many generations [56].  
323 However, the connection between a reduction in maintenance and culture heterogeneity remains  
324 unclear.

325 Other gene knockout selection strategies combine transcriptional networks and resource allocation  
326 to identify the minimal number of non-essential transcription factor knock outs to save the  
327 maximal amount cellular resources [69,70]. The work culminated in the removal of three  
328 transcription factors related to phosphate scavenging, flagella synthesis, and copper efflux. The  
329 resulting strain had no growth defect and showed an improved production of violacein in minimal  
330 media and increased yield of plasmid DNA in both rich and minimal media [69,70]. In another  
331 rational engineering example, genes encoding effector proteins were removed if they were  
332 upregulated in response to a stress, but had little or no regulatory effect on other genes. In this way,  
333 genes which are at the end of a signaling pathway can be removed without significant regulatory  
334 ramifications like RelA or SpoT. By observing heterologous protein induction stress, one study  
335 selectively removed one or two of the terminal stress response genes (whose function may or may  
336 not have been known). These knockouts showed an improved protein yield per unit biomass  
337 depending on the protein expressed [71]. However, these last two strategies were not evaluated in  
338 scale down reactors, so it is unclear how these genetic edits may affect the greater stress response.

339 Future rational gene selection strategies will likely continue to improve by use of *in vitro* tools  
340 [72], functional genomics [68], transcriptional network analysis [69], resource allocation models  
341 [73], and knowledge of the optimal enzyme substrate concentrations within cells [74].

342 While the previously discussed genome minimization strategies rely on knowledge of gene  
343 function and its estimated cost, non-obvious beneficial deletions can also be screened with random  
344 gene deletion techniques. For example, RB-TnSeq which replaces a sequence within the genome  
345 with a barcoded transposon, creates a library of identifiable knockouts.[75] RB-TnSeq has been  
346 used to evaluate beneficial mutations at multiple scales and has even found mutations that are  
347 unique to bioreactor cultures [76]. Furthermore, single gene deletions, although beneficial, likely  
348 will have a minor effect on cell growth parameters individually and therefore are often combined  
349 for more significant gains [77]. Random multigene knockout libraries offer further genotypic  
350 diversity and can be generated with iterative plasmid-based methods. Applications of these  
351 multigene knockout libraries show improved growth rate through genome reduction but offer  
352 diminishing returns with each cycle [78]. Another common random mutagenesis strategy is ALE,  
353 which takes advantage of *E. coli*'s natural accumulation of genetic errors to search for non-obvious  
354 beneficial mutations [79]. ALE-derived mutations are selected purely based on growth rate and  
355 often include genetic modifications by mobile insertion sequences (IS), gene specific single  
356 nucleotide polymorphisms (SNP), deletions, insertions, copy number variation, or multiple base  
357 pair substitutions [79]. ALE is frequently applied using static or increasing stress levels to improve  
358 tolerance, growth rate, substrate utilization, titer, and yield. ALE can also be used in fluctuating  
359 conditions like temperature shifts, demonstrating the potential for evolving strains exposed to  
360 transient stresses [79,80]. Combined use of ALE and genome reduction has potential to overcome  
361 some of the phenotypic issues associated with greatly reduced genomes [81]. However, over

362 modification of the genome through ALE or more targeted strategies, while potentially beneficial,  
363 is likely to have a diminishing or even negative effect if overused [78,82]. Unfortunately,  
364 applications of RB-TnSeq and ALE select for improved strains based solely on growth rate, and  
365 improved growth rate does not always result in improved process parameters. More efficient  
366 strategies may combine random or rational library generation with biosensors to instead select for  
367 both improved growth and improved product production [83]. Clearly the ability to generate strain  
368 diversity is no longer the bottleneck, but instead our ability to screen beneficial mutations and  
369 verify improvement in product production [84]. Regardless of the strategy used, it is vital to screen  
370 final production strains at industrial conditions to evaluate any unintentional negative effects.

371 **Concluding Remarks:**

372 Transitioning laboratory success to the industrial scale relies on predictable and consistent  
373 performance parameters at many volumes. Scale down technology offers an excellent  
374 methodology to study cells in a controlled environment that mimics industrial scale bioreactors  
375 without significant capital and operational investment. With design guided by computational fluid  
376 dynamics and industrial data, scale down reactors can accurately represent both the timescale and  
377 severity of industrial nutrient gradients. With this information, novel strains can be designed and  
378 tested for industrial scale challenges to evaluate their performance, and de-risk scale up. These  
379 industrial strains must be metabolically efficient, expressing only genes which are required for  
380 maintaining high productivity, titer, and yield. Strategies for selecting genes to remove can be  
381 either rational or random but must be evaluated in production strains at industrial conditions to  
382 understand their effect. Further, strains must be both capable of surviving and metabolically  
383 impartial to transient nutrient limitations (i.e., carbon, oxygen, etc.) and stresses (i.e., high CO<sub>2</sub>,  
384 O<sub>2</sub>, etc.) at scale. To optimize the cellular response to these environments, engineers may derive  
385 information from functional genomics and transcriptional network analysis to select either  
386 wasteful genes or transcriptional regulators for removal or modification. Additionally, a growing  
387 subset of microfluidics and cell sorting strategies are expanding our understanding of culture  
388 heterogeneity with single cell resolution. Methods like synthetic addition take advantage of this  
389 heterogeneity to improve culture performance, reduce escape mutants, and reduce heterogenous  
390 populations. Future studies evaluating how heterologous burden and maintenance energy interact  
391 with transcriptional, translational, and replication error will likely offer additional opportunities  
392 for engineering. However, adjusting the cellular response to multiple different stress has not yet  
393 been fully evaluated in the field. Therefore, strain engineers now have the challenge to map and

394 understand how the cell responds to multiple stresses and evaluate potential genetic interventions  
395 at the DNA, RNA, and protein level. By focusing on these **outstanding questions**, future industrial  
396 strains will spend their resources efficiently, limit their genetic adaptation and most of all maintain  
397 production at scale.

398 **Glossary**

399 **Cell Culture Heterogeneity:** Refers to the variation in cellular metabolism at a single cell level.

400 Naturally cells evolve over time, utilizing errors in DNA replication to diversify the population of  
401 cells and thus increase its survivability. On shorter timescales, cells are likely to achieve additional  
402 survivability through errors and variations in both translation and transcription. Often this can lead  
403 to changes in process parameters (e.g., titer, rate, and yield) as the culture is scaled to larger  
404 volumes. This may be due to longer seed trains requiring more doublings for larger volumes, or  
405 stresses induced at large scales which is often not captured in laboratory cultivations.

406 **Cell Culture Homogeneity:** In general, this term represents a population of cells which produce  
407 a product more similarly on a single cell level. Potential areas that can improve culture  
408 homogeneity include reducing cellular maintenance, reducing heterologous burden, removing  
409 error prone polymerases, and utilizing burden aware or synthetically coupled heterologous gene  
410 expression strategies. Increased culture homogeneity may improve process parameters like titer,  
411 rates, and yield at various scales.

412 **Cell Free:** A methodology utilizing cytosol isolated from cells to express proteins in an *in vitro*  
413 setting, typically used for balancing enzyme expression of heterologous pathways.

414 **Computational Fluid Dynamics:** A computational method to model the fluid flow in various  
415 geometries to estimate nonideal mixing conditions.

416 **Microfluidics:** A culturing technique relying on small channels with a characteristic height or  
417 width on the scale of micrometers. Constructed typically with soft polymers (like  
418 polydimethylsiloxane), these lab-on-a-chip devices can rapidly flow small volumes of media to  
419 shift conditions within the channels. Each channel can grow individual groups of cells for

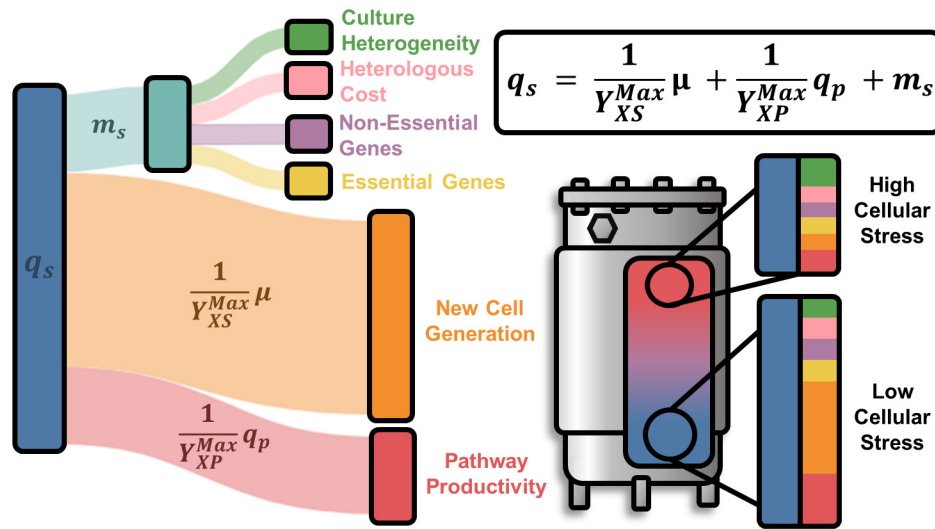
420 experimentation. Cultures are often visualized with an external microscope to evaluate growth  
421 parameters.

422 **RpoS Mediated Stress Response:** Controlled by the levels of the RpoS sigma factor, this stress  
423 response is typically activated upon reaching stationary phase and various other stressed  
424 conditions. RpoS is an adaptive response regulator known to regulate expression of over 1000  
425 genes, allowing cells to respond to starvation or unfavorable external conditions.

426 **Scale Down:** Culturing techniques which are modeled after large scale reactors intended to mimic  
427 non ideal mixing conditions at a laboratory scale, including stirred tank reactors and microfluidics.

428 **SOS response:** Induced upon sudden DNA Damage (like from reactive oxygen species), the  
429 regulator RecA binds to single stranded DNA, and is activated. The activated RecA then  
430 encourages self-cleavage of the LexA protein repressor lowering its affinity for DNA and allowing  
431 expression of SOS genes. Depending on the severity of the DNA damage, various DNA damage  
432 repair enzymes may be expressed until the DNA damage is repaired.

433 **Stringent response:** In response to a nutrient limitation, the signaling molecule (p)ppGpp  
434 accumulates leading to a reduced growth rate and metabolism in accordance with the severity of  
435 the limitation. Some examples of starvation events include lacking amino acids, carbon, iron,  
436 phosphate, nitrogen, and fatty acids.

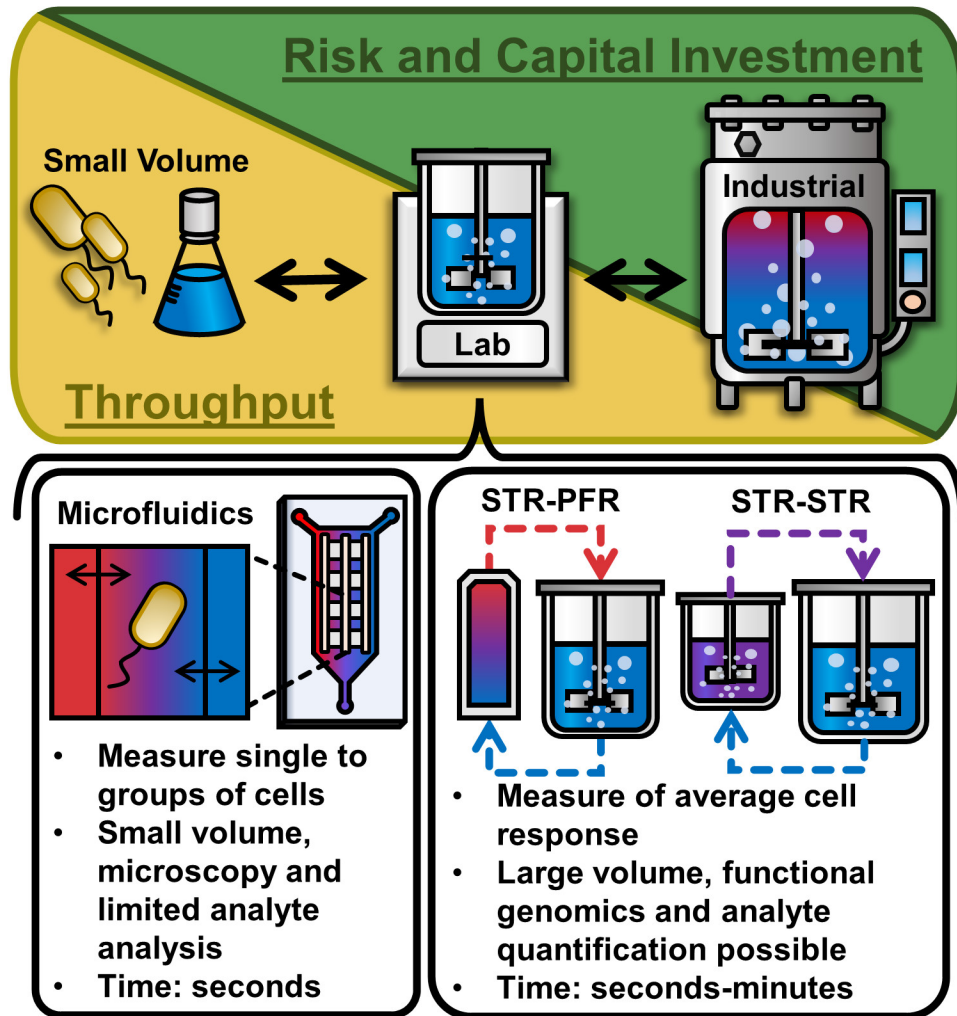


437

438 **FIGURE 1: Pirt visualization of cellular expenditures during cultivation.**

439 Represented as a Sankey diagram, substrate uptake rate ( $q_s$ ) can be dividing into new cell  
 440 generation  $\left(\frac{1}{Y_{XS}^{Max}} \mu\right)$ , the cost of the heterologous production pathway  $\left(\frac{1}{Y_{XP}^{Max}} q_p\right)$ , and cellular  
 441 maintenance ( $m_s$ ) [13]. Note that this is a simplified steady state representation to illustrate  
 442 cellular expenditures. Substrate uptake rate, new cell generation, pathway cost, and maintenance  
 443 will vary based on the growth rate and environmental stresses. Illustrated with a gradient of stress  
 444 typically seen in industrial fermentations, with red as high stress and blue as low stress, the total  
 445 substrate uptake as well as the distribution of expenditures can shift depending on the culture  
 446 conditions.

447



448

449 **FIGURE 2: Transition from small scale cultivations to industrial scale.**

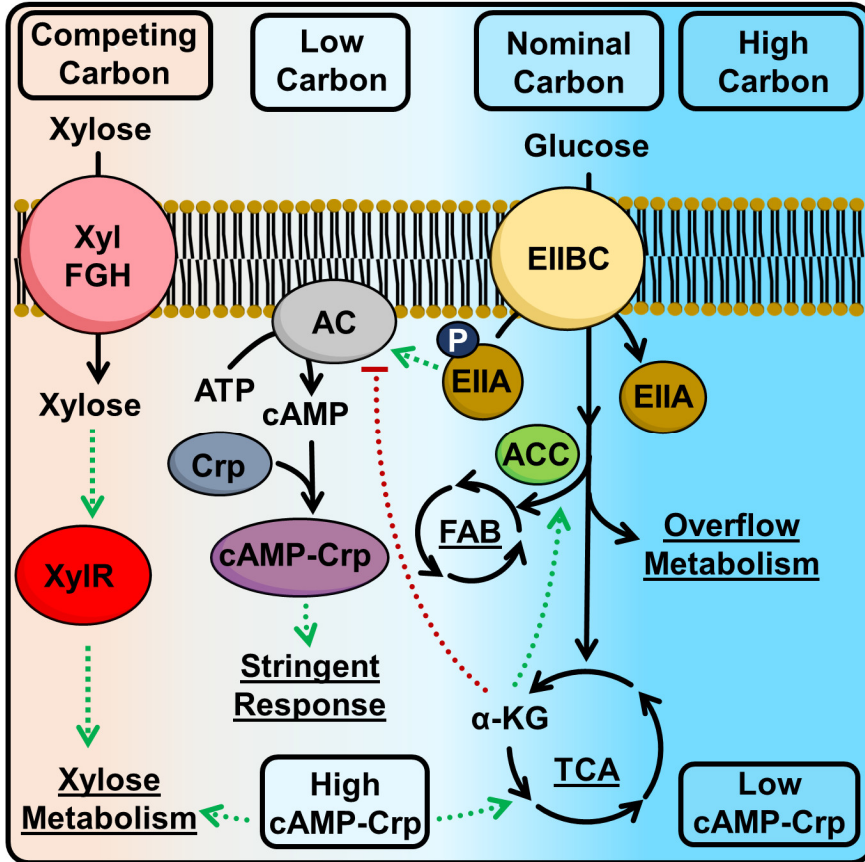
450 From single cell to lab scale bioreactors, titer, rate, and yield is often quantified in an ideal or well  
 451 mixed environment in which nutrients are readily available. Gradients modeled in large scale  
 452 reactors however show that nutrient availability can vary from excess to starvation at a timescale  
 453 from seconds to minutes [21–23]. With their increased risk and capital cost along with their low  
 454 throughput, large-scale reactors are often excluded from strain evaluation. Scale down reactors  
 455 offer an alternative with a relatively higher throughput and lower cost. These reactors can mimic  
 456 individual and combined stresses with the added benefit of product quantification and functional

457 genomics on a minute timescale [85]. Samples, however, will show the average cellular response  
458 and may not be able to capture individual cell variations unless cell sorting is utilized. Scale down  
459 microfluidics, however, can allow observations of single cells with fluorescent microscopy-based  
460 measurements on a timescale of seconds [85]. The frequency of the stress fluctuations is limited  
461 only by the speed of media flushing in the microfluidic design. However quantitative analysis is  
462 primarily limited by low analytic concentrations and low volumes making product quantification  
463 and functional genomics difficult, if not impossible, with current methods [22,27].

464

465 **BOX 1: Carbon dependent response**

466



467

468 **Figure I: Simplification of carbon metabolism and signaling in *E. coli*.** Note that larger

469 chemical or signaling pathways (TCA, FAB, stringent response, etc.) are underlined for clarity.

470 Enzymes and proteins are represented by ovals and individual chemicals are represented as text.

471 Positive regulation is represented with a dotted green arrow and negative regulation is represented

472 with a dotted red line. Columns represented by color and boxed header (High, Nominal, Low, and

473 Competing carbon) show the signaling pathway for each condition. The levels of core signaling

474 molecules are boxed at the bottom of each condition, in this case cAMP-Crp levels.

475

476 Transporters and catabolic enzymes responsible for carbon assimilation are regulated by both  
477 substrate-specific (e.g., XylR sensing of xylose) and global (e.g., cyclic-AMP-CRP) regulation  
478 mechanisms (**Figure I**). Glucose as a preferred carbon source is imported into the cell from the  
479 periplasm via the EIIABC transporter at the cost of 1 ATP. The EIIA component is not membrane  
480 bound and allows for sensing glucose availability. When cells sense a lack of preferred carbon  
481 sources, the activity of adenylate cyclase (AC) is upregulated, increasing cyclic-AMP (cAMP)  
482 synthesis. Conversely, cAMP synthesis is negatively regulated by accumulation of  $\alpha$ -keto acids,  
483 (e.g.,  $\alpha$ -ketoglutarate, pyruvate, and oxaloacetate) hallmarks of abundant catabolic flux or reduced  
484 synthesis of amino acids and other biomass precursors. The latter mechanism connects regulation  
485 of catabolic flux in fatty acid biosynthesis (FAB) via acetyl-CoA carboxylase (ACC) and the  
486 tricarboxylic acid (TCA) cycle to the stringent response. A complex (cAMP-CRP) of cAMP with  
487 its cognate receptor protein (CRP) activates transcription of promoters associated with secondary  
488 carbon uptake pathways. These pathways are typically positively regulated by the presence of  
489 cognate carbon sources (e.g., XylR) to ensure that pathways are expressed only when the carbon-  
490 source is present [29,86].

491 The absence of cAMP-CRP returns transcription of regulons to basal, often very low, levels. In  
492 addition to direct regulation, increased expression of the cAMP-CRP regulon decreases the pool  
493 of available cellular resources and indirectly downregulate genes outside the regulon [29].  
494 Conversely, downregulating the cAMP-CRP regulon makes resources available for translating  
495 other proteins. These global effects are because a cell's translation capacity is finite.

496 In scale down reactors, glucose starvation quickly (over about 2 hours) induces expression of the  
497 cAMP-CRP regulon including sugar transporters and the outer membrane porin OmpF to improve  
498 carbon uptake [87]. In continuous STR-PFR experiments, when exposed to a glucose excess along

499 the PFR, cells showed a rapid, short-term metabolic response leading to byproducts such as acetate  
500 and lactate. A function of overflow metabolism, these metabolic by-products negatively affect the  
501 substrate to biomass yield [88]. When instead cells are starved of glucose along a PFR, cells  
502 experience a decrease in the adenylate energy charge (a measure of energy availability in the cell)  
503 and an accumulation of ppGpp accumulation (signaling the stringent response) [77]. After 110s of  
504 glucose starvation, 400-600 genes were up- or downregulation by at least 1.5-fold. Carbohydrate  
505 and amino acid metabolism, glucose transport, post translational modification, protein turnover  
506 and folding were among the upregulated genes and translation, ribosomal structure, replication,  
507 recombination, and repair genes were on average downregulated. After 28 hours of intermittent  
508 glucose starvation, *E. coli* expressed genes associated with amino acids biosynthesis/metabolism,  
509 the general stress response and stationary phase growth [77,89].



521

522 Like carbon stress, nitrogen extremes can be caused by concentrated nitrogen feeds or poor mixing.  
523 Summarized in **Figure II**, nitrogen availability is primarily sensed by internal levels of glutamine  
524 and  $\alpha$ -ketoglutarate in *E. coli*. This is achieved by the uridylyl-transferase/uridylyl-removing  
525 enzyme (GlnD) which, depending on the glutamine concentration, determines the uridylation  
526 state of the PII homologs GlnK and GlnB. In nitrogen limited conditions or low glutamine levels,  
527 uridylated GlnK and GlnB accumulate. Uridylation of GlnK reduces its affinity for the AmtB  
528 transporter, allowing increased ammonia uptake. Whereas uridylated GlnB has two effects to  
529 increase nitrogen uptake. First, uridylated GlnB encourages de-adenylation of glutamine synthase  
530 (GS), increasing its activity and encouraging nitrogen assimilation. Second, uridylation of GlnB  
531 inhibits its interaction with the histidine kinase NtrB and maintaining the phosphorylated  
532 nitrogen response transcriptional regulator NtrC. While phosphorylated, NtrC increases expression  
533 of  $\sigma^{54}$  and RelA leading to a coordinated expression of nitrogen assimilation genes and the stringent  
534 response [90,91].  $\alpha$ -ketoglutarate additionally regulates nitrogen assimilation through direct  
535 binding to the GlnB protein. High concentrations of  $\alpha$ -ketoglutarate lead to direct binding of  $\alpha$ -  
536 ketoglutarate to GlnB, encouraging uridylation of GlnB, and deadenylation and thus activation of  
537 GS. High  $\alpha$ -ketoglutarate also maintains a phosphorylated NtrC by reducing binding of GlnB to  
538 NtrB, in general acting antagonistically to glutamine concentrations. In relation to carbon flux,  
539 GlnB binds to and inhibits the initiating enzyme of the fatty acid biosynthesis pathway, acetyl-  
540 CoA carboxylase (ACC). The Uridylation of GlnB reduces this inhibitory binding, and increases  
541 flux through the fatty acid biosynthesis pathway, allowing carbon flux in nitrogen limited  
542 conditions.  $\alpha$ -ketoglutarate binding to GlnB also encourages dissociation from ACC, increasing  
543 ACC activity. In this way GlnB and  $\alpha$ -ketoglutarate allow full ACC activity depending on both

544 carbon and nitrogen availability [92]. Additional modes of  $\alpha$ -ketoglutarate, ATP, and glutamine  
545 regulation have been reviewed elsewhere [93]. With excess glutamine (high nitrogen  
546 concentrations), GlnD will deuridylylate PII homologs GlnB and GlnK, preventing the nitrogen  
547 stress response summarized above. In nitrogen excess, diffusion of ammonia is thought to be the  
548 primary nitrogen transport method rather than any transporters. Under normal growth conditions  
549 glutamate dehydrogenase (GDH) acts as the primary assimilation pathway with glutamine  
550 synthase (GS) and glutamate synthase (GOGAT) is primarily used under nitrogen limitation  
551 requiring ATP [87].

552 When observing a STR shifting from nitrogen limitation to carbon limitation and finally dual  
553 nitrogen and carbon limitation,  $\alpha$ -ketoglutarate was again shown to balance carbon and nitrogen  
554 utilization. In this case,  $\alpha$ -ketoglutarate only remains high with carbon excess, but decreases under  
555 carbon limitations and dual carbon-nitrogen limitations. Thereby maintaining nitrogen  
556 assimilation pathways only if carbon was available [87]. In nitrogen limited STR-PFR cultivations,  
557 (p)ppGpp accumulated along the PFR (128 seconds), corresponding with the stringent response  
558 and NtrC mediated regulation of genes. After 5 minutes and 28 hours of cultivation, genes related  
559 to translation, nucleotide metabolism, and ribosomal structure and biogenesis were downregulated  
560 showing the expected nitrogen starvation response occurs even with transient limitations [45].

### 561 **BOX 3: Exploring Protein Expression Stress**

562 Expression of heterologous proteins at high levels can pull resources from native cellular  
563 processes. In flasks, expression of 45 different genes in *E. coli* showed a general trend balancing  
564 its proteome through the **RpoS mediated response** and translation related machinery [94]. In the  
565 same study, protein expression levels correlated with metal homeostasis genes likely due to a loss  
566 redox balance, the RpoH mediated heat shock response depending on the foldability of the protein  
567 expressed, and increased expression of nucleotide and amino acid biosynthetic pathways to support  
568 plasmid and protein production [94]. When observing a carbon limited fed batch at high growth  
569 rates, high recombinant protein expression leads to a significant reduction in CRP-cAMP and a  
570 significant increase in expression of the transcriptional regulator ArcA. This results in decreased  
571 glucose uptake, reduced TCA cycle activity, and lower cellular respiration likely influenced by  
572 accumulation of intermediate metabolites like pyruvate or  $\alpha$ -ketoglutarate [95]. In total, the  
573 response to sudden protein production was down regulation of glucose catabolism in an attempt to  
574 match the anabolism of plasmid and protein production while minimizing accumulation of  
575 byproducts, similar to an excess carbon response [96]. Industrial reactors are often used in a fed  
576 batch mode, so the observed glucose overfeeding response may be more relevant. From gene to  
577 protein, the majority of the energetic cost is in protein translation and amino acid generation [97].  
578 However additional cellular burden may depend also on either untranslatable RNA or the  
579 formation of inclusion bodies depending on the sequence and protein produced [98]. Finally, while  
580 the burden of heterologous gene expression can be estimated *a priori* [77,97], these strategies do  
581 not always consider variability in actual RNA and protein produced. An alternative *in vitro cell*  
582 **free** assay can estimate expression cost of an operon and separate out metabolic burden from  
583 expression burden in a single snapshot [72]. Advances in mass spectroscopy have also shown the

584 ability to accurately quantify the proteome in various stressed conditions, allowing more accurate  
585 predictions of protein production cost [99].

586 **BOX 4: The effect of Pressure, Oxygen, and Carbon Dioxide on Cells**

587 Elevated pressure in bioreactors is known to cause increased concentrations of gases at scale,  
588 affecting the maximum concentrations of oxygen and carbon dioxide cells experience. For  
589 example, mild hydrostatic pressure (1 MPa) is known to trigger an oxidative stress response due  
590 to increased oxygen concentrations and the formation of reactive oxygen species (ROS). These  
591 conditions have been shown to initiate the **SOS stress response** [100,101]. Oxidative stress  
592 additionally leads to rapid depletion of NADPH (a result of glutathione pathway defense for ROS)  
593 that must be recovered by the reserve flux capacity in the oxidative branch of the pentose phosphate  
594 pathway [102]. This decrease in reduction power may negatively influence cellular growth and  
595 product production, particularly if a heterologous pathway is dependent on NADPH as a co-factor.  
596 If oxygen is instead lacking, shifting to anaerobic conditions (over about ~13 seconds) reduces  
597 growth rate and recombinant protein production, and activates anerobic metabolism leading to  
598 organic acid production [103]. Cells could be engineered for better uptake of oxygen, such as  
599 expression of oxygen transporters [104], but may need to be balanced with the cells' ability to  
600 tolerate additional ROS.

601 For *E. coli*, carbon dioxide concentrations are not typically considered to be a critical variable in  
602 laboratory cultivations. However, industrial reactors can accumulate significantly more CO<sub>2</sub> due  
603 to the height of the reactors and high cell density cultures. Growth in consistently high CO<sub>2</sub>  
604 conditions has been shown to induce expression of amino acid pH resistance systems and down  
605 regulation of the tricarboxylic acid (TCA) cycle leading to a decrease in cell biomass, growth  
606 rate, and heterologous protein production [105,106]. Oscillatory dCO<sub>2</sub> concentrations, however,  
607 were not found to have a significant effect on cells in an STR-STR reactor mimicking industrial  
608 mixing times, suggesting high CO<sub>2</sub> environments on their own do not lead to significant

609 metabolic cost [105,106]. In a recent fed batch study, cells grown with CO<sub>2</sub> concentrations  
610 modeled after a 450m<sup>3</sup> tank only showed reduced productivity when increased CO<sub>2</sub>  
611 concentrations were combined with carbon and nitrogen limitations [107]. However, CO<sub>2</sub>  
612 enrichment was implemented by a constant feed of gas and is thus unable to completely mimic  
613 the fluctuating concentrations seen in large scale reactors [107]. Constant high CO<sub>2</sub> enrichment  
614 did however lead to premature termination of the culture due to acidification [107]. In total, the  
615 results suggest that carbon dioxide is unlikely to be the primary stress for industrial  
616 fermentations, but in combination with other limitations may affect core process parameters like  
617 biomass yield [107]. It is important to stress that while CO<sub>2</sub> may not be a significant factor for *E.*  
618 *coli*, the effect varies extremely by host [108,109].

619 **Acknowledgements:**

620 The viewpoints expressed in this review are solely those of the authors. At the time of writing,  
621 Prof. Pflieger was a visiting scholar at the IBVT at the University of Stuttgart supported by a  
622 Bessel Research Award from the Alexander von Humboldt Foundation. Prof. Pflieger's opinions  
623 are based on research conducted in his laboratory with support from the US National Science  
624 Foundation (EFRI-2132036), the US Department of Agriculture (NIFA-2020-67021-31140), and  
625 the US Department of Energy (DE-AR0001503, DE-SC0022207, DE-SC0018420, DE-  
626 SC0018409). William Cordell was supported by a Fenton May Fellowship from the Department  
627 of Chemical and Biological Engineering at the University of Wisconsin-Madison.

628 **Bibliography:**

- 629 1. MohammadiPeyhani, H. *et al.* (2022) Expanding biochemical knowledge and illuminating  
630 metabolic dark matter with ATLASx. *Nat Commun* 13, 1560
- 631 2. Voigt, C.A. (2020) Synthetic biology 2020–2030: six commercially-available products that  
632 are changing our world. *Nat Commun* 11, 6379
- 633 3. Lee, S.Y. *et al.* (2019) A comprehensive metabolic map for production of bio-based  
634 chemicals. *Nat Catal* 2, 18–33
- 635 4. Nielsen, J. and Keasling, J.D. (2016) Engineering Cellular Metabolism. *Cell* 164, 1185–  
636 1197
- 637 5. Casini, A. *et al.* (2018) A Pressure Test to Make 10 Molecules in 90 Days: External  
638 Evaluation of Methods to Engineer Biology. *J Am Chem Soc* 140, 4302–4316
- 639 6. Biggs, B.W. *et al.* (2021) Enabling commercial success of industrial biotechnology. *Science*  
640 (1979) 374, 1563–1565
- 641 7. Julleson, D. *et al.* (2015) Impact of synthetic biology and metabolic engineering on  
642 industrial production of fine chemicals. *Biotechnol Adv* 33, 1395–1402
- 643 8. Kampers, L.F.C. *et al.* (2021) From Innovation to Application: Bridging the Valley of Death  
644 in Industrial Biotechnology. *Trends Biotechnol* 39, 1240–1242
- 645 9. Crater, J.S. and Lievens, J.C. (2018) Scale-up of industrial microbial processes. *FEMS*  
646 *Microbiol Lett* 365, fny138
- 647 10. Zieringer, J. *et al.* (2021) Data-driven in silico prediction of regulation heterogeneity and  
648 ATP demands of *Escherichia coli* in large-scale bioreactors. *Biotechnol Bioeng* 118, 265–  
649 278

- 650 11. Yang, L. *et al.* (2018) Modeling the multi-scale mechanisms of macromolecular resource  
651 allocation. *Curr Opin Microbiol* 45, 8–15
- 652 12. Chen, K. *et al.* (2021) Bacterial fitness landscapes stratify based on proteome allocation  
653 associated with discrete aero-types. *PLoS Comput Biol* 17, e1008596
- 654 13. Pirt, S.J. (1965) The maintenance energy of bacteria in growing cultures. *Proceedings of*  
655 *the Royal Society of London Biology* 163, 224–231
- 656 14. Hafner, J. *et al.* (2020) Updated ATLAS of Biochemistry with New Metabolites and  
657 Improved Enzyme Prediction Power. *ACS Synth Biol* 9, 1479–1482
- 658 15. Hicks, M.A. and Prather, K.L.J. (2014) Bioprospecting in the Genomic Age. In *Advances*  
659 *in Applied Microbiology* 87, pp. 111–146, Academic Press Inc.
- 660 16. Dueber, J.E. *et al.* (2009) Synthetic protein scaffolds provide modular control over  
661 metabolic flux. *Nat Biotechnol* 27, 753–759
- 662 17. Pfleger, B.F. *et al.* (2006) Combinatorial engineering of intergenic regions in operons tunes  
663 expression of multiple genes. *Nat Biotechnol* 24, 1027–1032
- 664 18. Montaña López, J. *et al.* (2022) Physiological limitations and opportunities in microbial  
665 metabolic engineering. *Nat Rev Microbiol* 20, 35–48
- 666 19. Rajpurohit, H. and Eiteman, M.A. (2022) Nutrient-Limited Operational Strategies for the  
667 Microbial Production of Biochemicals. *Microorganisms* 10, 2226
- 668 20. Menacho-Melgar, R. *et al.* (2021) Optimization of phosphate-limited autoinduction broth  
669 for two-stage heterologous protein expression in *Escherichia coli*. *Biotechniques* 71, 566–  
670 572

- 671 21. Bisgaard, J. *et al.* (2021) Automated compartment model development based on data from  
672 flow-following sensor devices. *Processes* 9, 1651
- 673 22. Haringa, C. *et al.* (2018) From industrial fermentor to CFD-guided downscaling: what have  
674 we learned? *Biochem Eng J* 140, 57–71
- 675 23. Nadal-Rey, G. *et al.* (2022) Computational fluid dynamics modelling of hydrodynamics,  
676 mixing and oxygen transfer in industrial bioreactors with Newtonian broths. *Biochem Eng*  
677 *J* 177, 108265
- 678 24. Olughu, W. *et al.* (2019) Insight into the large-scale upstream fermentation environment  
679 using scaled-down models. *Journal of Chemical Technology and Biotechnology* 94, 647–  
680 657
- 681 25. Ho, P. *et al.* (2019) Reproduction of large-scale bioreactor conditions on microfluidic chips.  
682 *Microorganisms* 7, 105
- 683 26. Täuber, S. *et al.* (2020) DMSCC: A microfluidic platform for microbial single-cell  
684 cultivation of: *Corynebacterium glutamicum* under dynamic environmental medium  
685 conditions. *Lab Chip* 20, 4442–4455
- 686 27. Ho, P. *et al.* (2022) Microfluidic Reproduction of Dynamic Bioreactor Environment Based  
687 on Computational Lifelines. *Frontiers in Chemical Engineering* 4, 826485
- 688 28. Irving, S.E. *et al.* (2021) The stringent response and physiological roles of (pp)pGpp in  
689 bacteria. *Nat Rev Microbiol* 19, 256–271
- 690 29. Scott, M. and Hwa, T. (2022) Shaping bacterial gene expression by physiological and  
691 proteome allocation constraints. *Nat Rev Microbiol* 21, 327–342
- 692 30. Doran, P. (2013) *Bioprocess Engineering Principles*. In (Second.), pp. 428, Elsevier

- 693 31. Wilkinson, P.M. *et al.* (1994) MASS TRANSFER AND BUBBLE SIZE IN A BUBBLE  
694 COLUMN UNDER PRESSURE. *Chem Eng Sci* 49, 1417–1427
- 695 32. Follonier, S. *et al.* (2012) Pressure to kill or pressure to boost: A review on the various  
696 effects and applications of hydrostatic pressure in bacterial biotechnology. *Appl Microbiol*  
697 *Biotechnol* 93, 1805–1815
- 698 33. Spann, R. *et al.* (2019) CFD predicted pH gradients in lactic acid bacteria cultivations.  
699 *Biotechnol Bioeng* 116, 769–780
- 700 34. Cortes, J.T. *et al.* (2016) Physiological effects of pH gradients on Escherichia coli during  
701 plasmid DNA production. *Biotechnol. Bioeng* 113, 598–611
- 702 35. Paul, K. *et al.* (2020) Investigation of cell line specific responses to pH inhomogeneity and  
703 consequences for process design. *Eng Life Sci* 20, 412–421
- 704 36. Hewitt, C.J. *et al.* (1998) The Use of Flow Cytometry to Study the Impact of Fluid  
705 Mechanical Stress on Escherichia coli W3110 During Continuous Cultivation in an Agitated  
706 Bioreactor. *Biotechnol Bioeng* 59, 612–620
- 707 37. Gomes, D.G. *et al.* (2023) Bioreactors and engineering of filamentous fungi cultivation. In  
708 *Current Developments in Biotechnology and Bioengineering*, pp. 219–250, Elsevier
- 709 38. Balakrishnan, R. *et al.* (2021) Suboptimal resource allocation in changing environments  
710 constrains response and growth in bacteria. *Mol Syst Biol* 17
- 711 39. Minden, S. *et al.* (2023) Performing in spite of starvation: How *Saccharomyces cerevisiae*  
712 maintains robust growth when facing famine zones in industrial bioreactors. *Microb*  
713 *Biotechnol* 16, 148–168

- 714 40. Magnusson, L.U. *et al.* (2005) ppGpp: A global regulator in Escherichia coli. *Trends*  
715 *Microbiol* 13, 236–242
- 716 41. Hobbs, J.K. and Boraston, A.B. (2019) (p)ppGpp and the Stringent Response: An Emerging  
717 Threat to Antibiotic Therapy. *ACS Infect Dis* 5, 1505–1517
- 718 42. Xiao, H. *et al.* (1991) Residual Guanosine 3',5'-Bispyrophosphate Synthetic Activity of  
719 relA Null Mutants Can Be Eliminated by spoT Null Mutations\*. *J Biol Chem* 266, 5980–  
720 5990
- 721 43. Zhu, M. and Dai, X. (2019) Growth suppression by altered (p)ppGpp levels results from  
722 non-optimal resource allocation in Escherichia coli. *Nucleic Acids Res* 47, 4684–4693
- 723 44. Michalowski, A. *et al.* (2017) Escherichia coli HGT: Engineered for high glucose  
724 throughput even under slowly growing or resting conditions. *Metab Eng* 40, 93–103
- 725 45. Ziegler, M. *et al.* (2021) Transcriptional profiling of the stringent response mutant strain  
726 E. coli SR reveals enhanced robustness to large-scale conditions. *Microb Biotechnol* 14,  
727 993–1010
- 728 46. Chubukov, V. *et al.* (2017) Engineering glucose metabolism of Escherichia coli under  
729 nitrogen starvation. *NPJ Syst Biol Appl* 3, 16035
- 730 47. Velazquez, D. *et al.* (2022) Glucose transport engineering allows mimicking fed-batch  
731 performance in batch mode and selection of superior producer strains. *Microb Cell Fact* 21,  
732 183
- 733 48. Heins, A.L. *et al.* (2019) Quantitative flow cytometry to understand population  
734 heterogeneity in response to changes in substrate availability in Escherichia coli and  
735 Saccharomyces cerevisiae chemostats. *Front Bioeng Biotechnol* 7, 187

- 736 49. Xiao, Y. *et al.* (2016) Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis.  
737 *Nat Chem Biol* 12, 339–344
- 738 50. Rugbjerg, P. *et al.* (2021) Short and long-read ultra-deep sequencing profiles emerging  
739 heterogeneity across five platform Escherichia coli strains. *Metab Eng* 65, 197–206
- 740 51. Woo, A.C. *et al.* (2018) Heterogeneity of spontaneous DNA replication errors in single  
741 isogenic Escherichia coli cells. *Sci. Adv* 4, eaat1608
- 742 52. Sprouffske, K. *et al.* (2018) High mutation rates limit evolutionary adaptation in Escherichia  
743 coli. *PLoS Genet* 14, e1007324
- 744 53. Fijalkowska, I.J. *et al.* (2012) DNA replication fidelity in Escherichia coli: A multi-DNA  
745 polymerase affair. *FEMS Microbiol Rev* 36, 1105–1121
- 746 54. Consuegra, J. *et al.* (2021) Insertion-sequence-mediated mutations both promote and  
747 constrain evolvability during a long-term experiment with bacteria. *Nat Commun* 12, 980
- 748 55. Csörge, B. *et al.* (2012) Low-mutation-rate, reduced-genome Escherichia coli: An improved  
749 host for faithful maintenance of engineered genetic constructs. *Microb Cell Fact* 11, 11
- 750 56. Rugbjerg, P. *et al.* (2018) Diverse genetic error modes constrain large-scale bio-based  
751 production. *Nat Commun* 9, 787
- 752 57. Meer, K.M. *et al.* (2019) High Transcriptional Error Rates Vary as a Function of Gene  
753 Expression Level. *Genome Biol Evol* 12, 3754–3761
- 754 58. Li, W. and Lynch, M. (2020) Universally high transcript error rates in bacteria. *Elife* 9,  
755 e5489

- 756 59. Mordret, E. *et al.* (2019) Systematic Detection of Amino Acid Substitutions in Proteomes  
757 Reveals Mechanistic Basis of Ribosome Errors and Selection for Translation Fidelity. *Mol*  
758 *Cell* 75, 427-441.e5
- 759 60. Evans, C.R. *et al.* (2019) Increased mistranslation protects E. coli from protein misfolding  
760 stress due to activation of a RpoS-dependent heat shock response. *FEBS Lett* 593, 3220–  
761 3227
- 762 61. Samhita, L. *et al.* (2020) Global mistranslation increases cell survival under stress in  
763 Escherichia coli. *PLoS Genet* 16, e1008654
- 764 62. Boo, A. *et al.* (2019) Host-aware synthetic biology. *Curr Opin Syst Biol* 14, 66–72
- 765 63. Ceroni, F. *et al.* (2018) Burden-driven feedback control of gene expression. *Nat Methods*  
766 15, 387–393
- 767 64. Bertaux, F. *et al.* (2021) External control of microbial populations for bioproduction: A  
768 modeling and optimization viewpoint. *Curr Opin Syst Biol* 28, 100394
- 769 65. Rugbjerg, P. *et al.* (2018) Synthetic addiction extends the productive life time of engineered  
770 Escherichia coli populations. *Proc Natl Acad Sci U S A* 115, 2347–2352
- 771 66. Mori, M. *et al.* (2017) Quantifying the benefit of a proteome reserve in fluctuating  
772 environments. *Nat Commun* 8, 1225
- 773 67. Ziegler, M. and Takors, R. (2020) Minimal Cells: Design, Construction, Biotechnological  
774 Applicationspp. 1–32
- 775 68. Ziegler, M. *et al.* (2021) Engineering of a robust Escherichia coli chassis and exploitation  
776 for large-scale production processes. *Metab Eng* 67, 75–87

- 777 69. Lastiri-Pancardo, G. *et al.* (2020) A quantitative method for proteome reallocation using  
778 minimal regulatory interventions. *Nat Chem Biol* 16, 1026–1033
- 779 70. de la Cruz, M. *et al.* (2020) Plasmid DNA production in proteome-reduced Escherichia coli.  
780 *Microorganisms* 8, 1444
- 781 71. Sharma, A.K. *et al.* (2020) A novel knock out strategy to enhance recombinant protein  
782 expression in Escherichia coli. *Microb Cell Fact* 19, 148
- 783 72. Borkowski, O. *et al.* (2018) Cell-free prediction of protein expression costs for growing  
784 cells. *Nat Commun* 9, 1457
- 785 73. Yang, L. *et al.* (2018) Modeling the multi-scale mechanisms of macromolecular resource  
786 allocation. *Curr Opin Microbiol* 45, 8–15
- 787 74. Dourado, H. *et al.* (2021) On the optimality of the enzyme–substrate relationship in bacteria.  
788 *PLoS Biol* 19, e3001416
- 789 75. Wetmore, K.M. *et al.* (2015) Rapid quantification of mutant fitness in diverse bacteria by  
790 sequencing randomly bar-coded transposons. *mBio* 6, e00306-15
- 791 76. Eng, T. *et al.* (2021) Engineering *Pseudomonas putida* for efficient aromatic conversion to  
792 bioproduct using high throughput screening in a bioreactor. *Metab Eng* 66, 229–238
- 793 77. Löffler, M. *et al.* (2016) Engineering *E. coli* for large-scale production – Strategies  
794 considering ATP expenses and transcriptional responses. *Metab Eng* 38, 73–85
- 795 78. Vernyik, V. *et al.* (2020) Exploring the fitness benefits of genome reduction in Escherichia  
796 coli by a selection-driven approach. *Sci Rep* 10, 7345

- 797 79. Sandberg, T.E. *et al.* (2019) The emergence of adaptive laboratory evolution as an efficient  
798 tool for biological discovery and industrial biotechnology. *Metab Eng* 56, 1–16
- 799 80. Lambros, M. *et al.* (2021) Emerging Adaptive Strategies Under Temperature Fluctuations  
800 in a Laboratory Evolution Experiment of Escherichia Coli. *Front Microbiol* 12, 724982
- 801 81. Choe, D. *et al.* (2019) Adaptive laboratory evolution of a genome-reduced Escherichia coli.  
802 *Nat Commun* 10, 935
- 803 82. Phaneuf, P. v. *et al.* (2021) Escherichia coli Data-Driven Strain Design Using Aggregated  
804 Adaptive Laboratory Evolution Mutational Data. *ACS Synth Biol* 10, 3379–3395
- 805 83. Koch, M. *et al.* (2019) Custom-made transcriptional biosensors for metabolic engineering.  
806 *Curr Opin Biotechnol* 59, 78–84
- 807 84. Wehrs, M. *et al.* (2020) You get what you screen for: on the value of fermentation  
808 characterization in high-throughput strain improvements in industrial settings. *J Ind*  
809 *Microbiol Biotechnol* 47, 913–927
- 810 85. Haringa, C. *et al.* (2016) Euler-Lagrange computational fluid dynamics for (bio)reactor  
811 scale down: An analysis of organism lifelines. *Eng Life Sci* 16, 652–663
- 812 86. Sievert, C. *et al.* (2017) Experimental evolution reveals an effective avenue to release  
813 catabolite repression via mutations in XylR. *Proc Natl Acad Sci U S A* 114, 7349–7354
- 814 87. Löffler, M. *et al.* (2017) Switching between nitrogen and glucose limitation: Unraveling  
815 transcriptional dynamics in Escherichia coli. *J Biotechnol* 258, 2–12
- 816 88. Neubauer, P. *et al.* (1995) Influence of substrate oscillations on acetate formation and  
817 growth yield in Escherichia coli glucose limited fed-batch cultivations. *Biotechnol Bioeng*  
818 47, 139–146

- 819 89. Schellhorn, H.E. (2020) Function, Evolution, and Composition of the RpoS Regulon in  
820 Escherichia coli. *Front Microbiol* 11, 560099
- 821 90. Sanchuki, H.B.S. *et al.* (2017) Dynamics of the Escherichia coli proteome in response to  
822 nitrogen starvation and entry into the stationary phase. *Biochim Biophys Acta Proteins*  
823 *Proteom* 1865, 344–352
- 824 91. Switzer, A. *et al.* (2018) New insights into the adaptive transcriptional response to nitrogen  
825 starvation in Escherichia coli. *Biochem Soc Trans* 46, 1721–1728
- 826 92. Gerhardt, E.C.M. *et al.* (2015) The Bacterial signal transduction protein GlnB regulates the  
827 committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of  
828 acetyl-CoA carboxylase. *Mol Microbiol* 95, 1025–1035
- 829 93. Huergo, L.F. *et al.* (2013) PII signal transduction proteins: Nitrogen regulation and beyond.  
830 *FEMS Microbiol Rev* 37, 251–283
- 831 94. Tan, J. *et al.* (2020) Independent component analysis of E. coli’s transcriptome reveals the  
832 cellular processes that respond to heterologous gene expression. *Metab Eng* 61, 360–368
- 833 95. Li, Z. *et al.* (2021) Transcriptional network analysis identifies key elements governing the  
834 recombinant protein production provoked reprogramming of carbon and energy metabolism  
835 in Escherichia coli BL21 (DE3). *Engineering Reports* 3, e12393
- 836 96. Li, Z. and Rinas, U. (2021) Recombinant protein production-associated metabolic burden  
837 reflects anabolic constraints and reveals similarities to a carbon overfeeding response.  
838 *Biotechnol Bioeng* 118, 94–105
- 839 97. Lynch, M. and Marinov, G.K. (2015) The bioenergetic costs of a gene. *Proc Natl Acad Sci*  
840 *U S A* 112, 15690–15695

- 841 98. Li, Z. and Rinas, U. (2020) Recombinant protein production associated growth inhibition  
842 results mainly from transcription and not from translation. *Microb Cell Fact* 19, 83
- 843 99. Mori, M. *et al.* (2021) From coarse to fine: the absolute Escherichia coli proteome under  
844 diverse growth conditions. *Mol Syst Biol* 17, e9536
- 845 100. Guyet, A. *et al.* (2018) Mild hydrostatic pressure triggers oxidative responses in Escherichia  
846 coli. *PLoS One* 13, e0200660
- 847 101. Maslowska, K.H. *et al.* (2019) The SOS system: A complex and tightly regulated response  
848 to DNA damage. *Environ Mol Mutagen* 60, 368–384
- 849 102. Christodoulou, D. *et al.* (2018) Reserve Flux Capacity in the Pentose Phosphate Pathway  
850 Enables Escherichia coli's Rapid Response to Oxidative Stress. *Cell Syst* 6, 569-578.e7
- 851 103. Sandoval-Basurto, E.A. *et al.* (2005) Culture of Escherichia coli under dissolved oxygen  
852 gradients simulated in a two-compartment scale-down system: Metabolic response and  
853 production of recombinant protein. *Biotechnol Bioeng* 89, 453–463
- 854 104. Liu, D. *et al.* (2017) Enhancing fatty acid production in Escherichia coli by Vitreoscilla  
855 hemoglobin overexpression. *Biotechnol. Bioeng* 114, 463–467
- 856 105. Baez, A. *et al.* (2011) Simulation of dissolved CO<sub>2</sub> gradients in a scale-down system: A  
857 metabolic and transcriptional study of recombinant Escherichia coli. *Biotechnol J* 6, 959–  
858 967
- 859 106. Baez, A. *et al.* (2009) Metabolic and transcriptional response of recombinant Escherichia  
860 coli to elevated dissolved carbon dioxide concentrations. *Biotechnol Bioeng* 104, 102–110

- 861 107. Gecse, G. *et al.* (2022) Impact of Elevated Levels of Dissolved CO<sub>2</sub> on Performance and  
862 Proteome Response of an Industrial 2'-Fucosyllactose Producing *Escherichia coli* Strain.  
863 *Microorganisms* 10, 1145
- 864 108. Morales, M. *et al.* (2018) The impact of environmental factors on carbon dioxide fixation  
865 by microalgae. *FEMS Microbiol Lett* 365, fnx262
- 866 109. Blombach, B. and Takors, R. (2015) CO<sub>2</sub> - intrinsic product, essential substrate, and  
867 regulatory trigger of microbial and mammalian production processes. *Front Bioeng*  
868 *Biotechnol* 3, 108
- 869