Milligrams to Kilograms: Making Microbes Work at Scale

William T Cordell[1], Gennaro Avolio[4], Ralf Takors[4], Brian Pfleger[1,2,3,*]

[1] Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, USA

[2] DOE Center Advanced Bioenergy and Bioproducts Innovation, University of Wisconsin-Madison, Madison, WI 53706, USA

[3] DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI 53706, USA

[4] Institute of Biochemical Engineering, University of Stuttgart, Stuttgart 70569 Germany

*Correspondence: brian.pfleger@wisc.edu (B. F. Pfleger)

Keywords: Scale Down, Genome Minimization, Cell Heterogeneity, Strain Engineering, Computational Fluid Dynamic, Bioprocess Engineering

Abstract: If biomanufacturing can become a sustainable route for producing chemicals, it will provide a critical step in reducing greenhouse gas emissions to fight climate change. However, efforts to industrialize microbial synthesis of chemicals have met with varied success – in part due to challenges in translating laboratory successes to industrial scales. With particular focus on *Escherichia coli,* this review examines the lessons learned when studying microbial physiology and metabolism under conditions that simulate large-scale bioreactors and methods to minimize cellular waste through reduction of maintenance energy, optimizing the stress response, and minimizing culture heterogeneity. With general strategies to overcome these challenges, biomanufacturing process scale-up could be de-risked and potentially reduce the time and cost of bringing promising syntheses to market.

Scaling the Bioeconomy

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

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

The Challenges of Bioprocess Engineering at Scale

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

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

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

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

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

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

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

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

Optimization of the Stress Response

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

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

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

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

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

Minimizing Heterogeneity and Cellular Adaption

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

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

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

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

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

An alternative method to increase homogeneity of cultures takes advantage of expression noise through metabolic coupling. Also known as synthetic addition, this strategy couples a product detecting biosensor or native product sensing transcriptional regulator with expression of an essential or beneficial gene for cell growth. As a result, product synthesis confers a fitness benefit. The first demonstrations of this methodology explored producing free fatty acids and tyrosine. In the free fatty acid producing strains, the authors coupled production to a tetracycline resistance marker; creating a selection pressure for fatty acids by supplementing the antibiotic to the growth media. Using this approach, a greater proportion of cells were classified as high fatty acid producers. Coupling the fatty acid production pathway to leucine biosynthesis in an auxotrophic *E. coli* (instead of an antibiotic resistance marker), led to 21.5g/L FFA in a fed batch cultivation, a greater than four times improvement compared to a strain without metabolic coupling [49]. Iterative flask cultivations suggested improvements were due to non-genetic variations. When high producing cells were isolated and returned to fresh media without selection, the enhanced product titer was not reproduced, and genome sequencing did not suggest any beneficial genetic mutations [49]. Analogously, coupling essential gene expressions to a mevalonate biosensor a synthetically addicted strain could maintain mevalonate production over 85 generations. The un-addicted strain by comparison had reduced mevalonate production due to the accumulation of insertion sequences

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

Minimize basal maintenance energy:

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

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

Concluding Remarks:

Transitioning laboratory success to the industrial scale relies on predictable and consistent performance parameters at many volumes. Scale down technology offers an excellent methodology to study cells in a controlled environment that mimics industrial scale bioreactors without significant capital and operational investment. With design guided by computational fluid dynamics and industrial data, scale down reactors can accurately represent both the timescale and severity of industrial nutrient gradients. With this information, novel strains can be designed and tested for industrial scale challenges to evaluate their performance, and de-risk scale up. These industrial strains must be metabolically efficient, expressing only genes which are required for maintaining high productivity, titer, and yield. Strategies for selecting genes to remove can be either rational or random but must be evaluated in production strains at industrial conditions to understand their effect. Further, strains must be both capable of surviving and metabolically impartial to transient nutrient limitations (i.e., carbon, oxygen, etc.) and stresses (i.e., high CO_2, O_2, etc.) at scale. To optimize the cellular response to these environments, engineers may derive information from functional genomics and transcriptional network analysis to select either wasteful genes or transcriptional regulators for removal or modification. Additionally, a growing subset of microfluidics and cell sorting strategies are expanding our understanding of culture heterogeneity with single cell resolution. Methods like synthetic addition take advantage of this heterogeneity to improve culture performance, reduce escape mutants, and reduce heterogenous populations. Future studies evaluating how heterologous burden and maintenance energy interact with transcriptional, translational, and replication error will likely offer additional opportunities for engineering. However, adjusting the cellular response to multiple different stress has not yet 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.

Glossary

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

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

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

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

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

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

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

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

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

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

FIGURE 1: Pirt visualization of cellular expenditures during cultivation.

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

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

From single cell to lab scale bioreactors, titer, rate, and yield is often quantified in an ideal or well mixed environment in which nutrients are readily available. Gradients modeled in large scale reactors however show that nutrient availability can vary from excess to starvation at a timescale from seconds to minutes [21–23]. With their increased risk and capital cost along with their low throughput, large-scale reactors are often excluded from strain evaluation. Scale down reactors offer an alternative with a relatively higher throughput and lower cost. These reactors can mimic individual and combined stresses with the added benefit of product quantification and functional genomics on a minute timescale [85]. Samples, however, will show the average cellular response and may not be able to capture individual cell variations unless cell sorting is utilized. Scale down microfluidics, however, can allow observations of single cells with fluorescent microscopy-based measurements on a timescale of seconds [85]. The frequency of the stress fluctuations is limited only by the speed of media flushing in the microfluidic design. However quantitative analysis is primarily limited by low analytic concentrations and low volumes making product quantification and functional genomics difficult, if not impossible, with current methods [22,27].

BOX 1: Carbon dependent response

Figure I: Simplification of carbon metabolism and signaling in *E. coli*. Note that larger chemical or signaling pathways (TCA, FAB, stringent response, etc.) are underlined for clarity. Enzymes and proteins are represented by ovals and individual chemicals are represented as text. Positive regulation is represented with a dotted green arrow and negative regulation is represented with a dotted red line. Columns represented by color and boxed header (High, Nominal, Low, and Competing carbon) show the signaling pathway for each condition. The levels of core signaling molecules are boxed at the bottom of each condition, in this case cAMP-Crp levels.

Transporters and catabolic enzymes responsible for carbon assimilation are regulated by both substrate-specific (e.g., XylR sensing of xylose) and global (e.g., cyclic-AMP-CRP) regulation mechanisms (**Figure I**). Glucose as a preferred carbon source is imported into the cell from the periplasm via the EllABC transporter at the cost of 1 ATP. The EllA component is not membrane bound and allows for sensing glucose availability. When cells sense a lack of preferred carbon sources, the activity of adenylate cyclase (AC) is upregulated, increasing cyclic-AMP (cAMP) synthesis. Conversely, cAMP synthesis is negatively regulated by accumulation of α-keto acids, (e.g., α-ketoglutarate, pyruvate, and oxaloacetate) hallmarks of abundant catabolic flux or reduced synthesis of amino acids and other biomass precursors. The latter mechanism connects regulation of catabolic flux in fatty acid biosynthesis (FAB) via acetyl-CoA carboxylase (ACC) and the tricarboxylic acid (TCA) cycle to the stringent response. A complex (cAMP-CRP) of cAMP with its cognate receptor protein (CRP) activates transcription of promoters associated with secondary carbon uptake pathways. These pathways are typically positively regulated by the presence of

cognate carbon sources (e.g., XylR) to ensure that pathways are expressed only when the carbon-source is present [29,86]. The absence of cAMP-CRP returns transcription of regulons to basal, often very low, levels. In addition to direct regulation, increased expression of the cAMP-CRP regulon decreases the pool of available cellular resources and indirectly downregulate genes outside the regulon [29]. Conversely, downregulating the cAMP-CRP regulon makes resources available for translating other proteins. These global effects are because a cell's translation capacity is finite.

In scale down reactors, glucose starvation quickly (over about 2 hours) induces expression of the cAMP-CRP regulon including sugar transporters and the outer membrane porin OmpF to improve carbon uptake [87]. In continuous STR-PFR experiments, when exposed to a glucose excess along the PFR, cells showed a rapid, short-term metabolic response leading to byproducts such as acetate and lactate. A function of overflow metabolism, these metabolic by-products negatively affect the substrate to biomass yield [88]. When instead cells are starved of glucose along a PFR, cells experience a decrease in the adenylate energy charge (a measure of energy availability in the cell) and an accumulation of ppGpp accumulation (signaling the stringent response) [77]. After 110s of glucose starvation, 400-600 genes were up- or downregulation by at least 1.5-fold. Carbohydrate and amino acid metabolism, glucose transport, post translational modification, protein turnover and folding were among the upregulated genes and translation, ribosomal structure, replication, recombination, and repair genes were on average downregulated. After 28 hours of intermittent glucose starvation, *E. coli* expressed genes associated with amino acids biosynthesis/metabolism, the general stress response and stationary phase growth [77,89].

BOX 2: Nitrogen dependent response

Figure II: Simplification of nitrogen signaling pathway and metabolism in *E. coli*. Note that larger chemical or signaling pathways (stringent response) are underlined for clarity. Enzymes and proteins are represented by ovals and individual chemicals are represented as text. Glutamine is simplified to Gln and glutamate to Glu. Positive regulation is represented with a dotted green arrow and negative regulation is represented with a dotted red line. Columns represented by color and boxed description show the signaling pathway for each condition at the top of the figure. Boxed description of core signaling molecules, in this case glutamine levels and the ratio of glutamine to α-ketoglutarate are located within each column.

Like carbon stress, nitrogen extremes can be caused by concentrated nitrogen feeds or poor mixing. Summarized in **Figure II**, nitrogen availability is primarily sensed by internal levels of glutamine and α-ketoglutarate in *E. coli*. This is achieved by the uridylyl-transferase/uridylyl-removing enzyme (GlnD) which, depending on the glutamine concentration, determines the uridylayation state of the PII homologs GlnK and GlnB. In nitrogen limited conditions or low glutamine levels, uridylated GlnK and GlnB accumulate. Uridylation of GlnK reduces its affinity for the AmtB transporter, allowing increased ammonia uptake. Whereas uridylated GlnB has two effects to increase nitrogen uptake. First, uridylated GlnB encourages de-adenylation of glutamine synthase (GS), increasing its activity and encouraging nitrogen assimilation. Second, uridylation of GlnB inhibits its interaction with the histidine kinase NtrB and maintaining the phosphorylate the nitrogen response transcriptional regulator NtrC. While phosphorylated, NtrC increases expression of σ^{54} and RelA leading to a coordinated expression of nitrogen assimilation genes and the stringent

response [90,91]. α-ketoglutarate additionally regulates nitrogen assimilation through direct binding to the GlnB protein. High concentrations of α-ketoglutarate lead to direct binding of α-ketoglutarate to GlnB, encouraging uridylation of GlnB, and deadenylation and thus activation of GS. High α-ketoglutarate also maintains a phosphorylated NtrC by reducing binding of GlnB to NtrB, in general acting antagonistically to glutamine concentrations. In relation to carbon flux, GlnB binds to and inhibits the initiating enzyme of the fatty acid biosynthesis pathway, acetyl-CoA carboxylase (ACC). The Uridylation of GlnB reduces this inhibitory binding, and increases flux through the fatty acid biosynthesis pathway, allowing carbon flux in nitrogen limited conditions. α-ketoglutarate binding to GlnB also encourages dissociation from ACC, increasing ACC activity. In this way GlnB and α-ketoglutarate allow full ACC activity depending on both carbon and nitrogen availability [92]. Additional modes of α-ketoglutarate, ATP, and glutamine regulation have been reviewed elsewhere [93]. With excess glutamine (high nitrogen concentrations), GlnD will deuridylayate PII homologs GlnB and GlnK, preventing the nitrogen stress response summarized above. In nitrogen excess, diffusion of ammonia is thought to be the primary nitrogen transport method rather than any transporters. Under normal growth conditions glutamate dehydrogenase (GDH) acts as the primary assimilation pathway with glutamine synthase (GS) and glutamate synthase (GOGAT) is primarily used under nitrogen limitation requiring ATP [87].

When observing a STR shifting from nitrogen limitation to carbon limitation and finally dual nitrogen and carbon limitation, α-ketoglutarate was again shown to balance carbon and nitrogen utilization. In this case, α-ketoglutarate only remains high with carbon excess, but decreases under carbon limitations and dual carbon-nitrogen limitations. Thereby maintaining nitrogen assimilation pathways only if carbon was available [87]. In nitrogen limited STR-PFR cultivations,

553 (p)ppGpp accumulated along the PFR (128 seconds), corresponding with the stringent response
554 and NtrC mediated regulation of genes. After 5 minutes and 28 hours of cultivation, genes related
555 to translation, nucleotide metabolism, and ribosomal structure and biogenesis were downregulated
556 showing the expected nitrogen starvation response occurs even with transient limitations [45].

BOX 3: Exploring Protein Expression Stress

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

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

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

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

For *E. coli*, carbon dioxide concentrations are not typically considered to be a critical variable in laboratory cultivations. However, industrial reactors can accumulate significantly more CO_2 due to the height of the reactors and high cell density cultures. Growth in consistently high CO_2 conditions has been shown to induce expression of amino acid pH resistance systems and down regulation of the tricarboxylic acid (TCA) cycle leading to a decrease in cell biomass, growth rate, and heterologous protein production [105,106]. Oscillatory dCO_2 concentrations, however, were not found to have a significant effect on cells in an STR-STR reactor mimicking industrial mixing times, suggesting high CO_2 environments on their own do not lead to significant

metabolic cost [105,106]. In a recent fed batch study, cells grown with CO_2 concentrations modeled after a 450m^3 tank only showed reduced productivity when increased CO_2 concentrations were combined with carbon and nitrogen limitations [107]. However, CO_2 enrichment was implemented by a constant feed of gas and is thus unable to completely mimic the fluctuating concentrations seen in large scale reactors [107]. Constant high CO_2 enrichment did however lead to premature termination of the culture due to acidification [107]. In total, the results suggest that carbon dioxide is unlikely to be the primary stress for industrial fermentations, but in combination with other limitations may affect core process parameters like biomass yield [107]. It is important to stress that while CO_2 may not be a significant factor for *E. coli*, the effect varies extremely by host [108,109].

Acknowledgements:

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

Bibliography:

1. MohammadiPeyhani, H. *et al.* (2022) Expanding biochemical knowledge and illuminating metabolic dark matter with ATLASx. *Nat Commun* 13, 1560

2. Voigt, C.A. (2020) Synthetic biology 2020–2030: six commercially-available products that are changing our world. *Nat Commun* 11, 6379

3. Lee, S.Y. *et al.* (2019) A comprehensive metabolic map for production of bio-based chemicals. *Nat Catal* 2, 18–33

4. Nielsen, J. and Keasling, J.D. (2016) Engineering Cellular Metabolism. *Cell* 164, 1185–1197

5. Casini, A. *et al.* (2018) A Pressure Test to Make 10 Molecules in 90 Days: External Evaluation of Methods to Engineer Biology. *J Am Chem Soc* 140, 4302–4316

6. Biggs, B.W. *et al.* (2021) Enabling commercial success of industrial biotechnology. *Science (1979)* 374, 1563–1565

7. Jullesson, D. *et al.* (2015) Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnol Adv* 33, 1395–1402

8. Kampers, L.F.C. *et al.* (2021) From Innovation to Application: Bridging the Valley of Death in Industrial Biotechnology. *Trends Biotechnol* 39, 1240–1242

9. Crater, J.S. and Lievense, J.C. (2018) Scale-up of industrial microbial processes. *FEMS Microbiol Lett* 365, fny138

10. Zieringer, J. *et al.* (2021) Data-driven in silico prediction of regulation heterogeneity and ATP demands of Escherichia coli in large-scale bioreactors. *Biotechnol Bioeng* 118, 265–278

11. Yang, L. *et al.* (2018) Modeling the multi-scale mechanisms of macromolecular resource allocation. *Curr Opin Microbiol* 45, 8–15

12. Chen, K. *et al.* (2021) Bacterial fitness landscapes stratify based on proteome allocation associated with discrete aero-types. *PLoS Comput Biol* 17, e1008596

13. Pirt, S.J. (1965) The maintenance energy of bacteria in growing cultures. *Proceedings of the Royal Society of London Biology* 163, 224–231

14. Hafner, J. *et al.* (2020) Updated ATLAS of Biochemistry with New Metabolites and Improved Enzyme Prediction Power. *ACS Synth Biol* 9, 1479–1482

15. Hicks, M.A. and Prather, K.L.J. (2014) Bioprospecting in the Genomic Age. In *Advances in Applied Microbiology* 87, pp. 111–146, Academic Press Inc.

16. Dueber, J.E. *et al.* (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat Biotechnol* 27, 753–759

17. Pfleger, B.F. *et al.* (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat Biotechnol* 24, 1027–1032

18. Montaño López, J. *et al.* (2022) Physiological limitations and opportunities in microbial metabolic engineering. *Nat Rev Microbiol* 20, 35–48

19. Rajpurohit, H. and Eiteman, M.A. (2022) Nutrient-Limited Operational Strategies for the Microbial Production of Biochemicals. *Microorganisms* 10, 2226

20. Menacho-Melgar, R. *et al.* (2021) Optimization of phosphate-limited autoinduction broth for two-stage heterologous protein expression in Escherichia coli. *Biotechniques* 71, 566–572

21. Bisgaard, J. *et al.* (2021) Automated compartment model development based on data from flow-following sensor devices. *Processes* 9, 1651

22. Haringa, C. *et al.* (2018) From industrial fermentor to CFD-guided downscaling: what have we learned? *Biochem Eng J* 140, 57–71

23. Nadal-Rey, G. *et al.* (2022) Computational fluid dynamics modelling of hydrodynamics, mixing and oxygen transfer in industrial bioreactors with Newtonian broths. *Biochem Eng J* 177, 108265

24. Olughu, W. *et al.* (2019) Insight into the large-scale upstream fermentation environment using scaled-down models. *Journal of Chemical Technology and Biotechnology* 94, 647–657

25. Ho, P. *et al.* (2019) Reproduction of large-scale bioreactor conditions on microfluidic chips. *Microorganisms* 7, 105

26. Täuber, S. *et al.* (2020) DMSCC: A microfluidic platform for microbial single-cell cultivation of: Corynebacterium glutamicum under dynamic environmental medium conditions. *Lab Chip* 20, 4442–4455

27. Ho, P. *et al.* (2022) Microfluidic Reproduction of Dynamic Bioreactor Environment Based on Computational Lifelines. *Frontiers in Chemical Engineering* 4, 826485

28. Irving, S.E. *et al.* (2021) The stringent response and physiological roles of (pp)pGpp in bacteria. *Nat Rev Microbiol* 19, 256–271

29. Scott, M. and Hwa, T. (2022) Shaping bacterial gene expression by physiological and proteome allocation constraints. *Nat Rev Microbiol* 21, 327–342

30. Doran, P. (2013) Bioprocess Engineering Principles. In (Second.), pp. 428, Elsevier

31. Wilkinson, P.M. *et al.* (1994) MASS TRANSFER AND BUBBLE SIZE IN A BUBBLE COLUMN UNDER PRESSURE. *Chem Eng Sci* 49, 1417–1427

32. Follonier, S. *et al.* (2012) Pressure to kill or pressure to boost: A review on the various effects and applications of hydrostatic pressure in bacterial biotechnology. *Appl Microbiol Biotechnol* 93, 1805–1815

33. Spann, R. *et al.* (2019) CFD predicted pH gradients in lactic acid bacteria cultivations. *Biotechnol Bioeng* 116, 769–780

34. Cort es, J.T. *et al.* (2016) Physiological effects of pH gradients on Escherichia coli during plasmid DNA production. *Biotechnol. Bioeng* 113, 598–611

35. Paul, K. *et al.* (2020) Investigation of cell line specific responses to pH inhomogeneity and consequences for process design. *Eng Life Sci* 20, 412–421

36. Hewitt, C.J. *et al.* (1998) The Use of Flow Cytometry to Study the Impact of Fluid Mechanical Stress on Escherichia coli W3110 During Continuous Cultivation in an Agitated Bioreactor. *Biotechnol Bioeng* 59, 612–620

37. Gomes, D.G. *et al.* (2023) Bioreactors and engineering of filamentous fungi cultivation. In *Current Developments in Biotechnology and Bioengineering*, pp. 219–250, Elsevier

38. Balakrishnan, R. *et al.* (2021) Suboptimal resource allocation in changing environments constrains response and growth in bacteria. *Mol Syst Biol* 17

39. Minden, S. *et al.* (2023) Performing in spite of starvation: How Saccharomyces cerevisiae maintains robust growth when facing famine zones in industrial bioreactors. *Microb Biotechnol* 16, 148–168

40. Magnusson, L.U. *et al.* (2005) ppGpp: A global regulator in Escherichia coli. *Trends Microbiol* 13, 236–242

41. Hobbs, J.K. and Boraston, A.B. (2019) (p)ppGpp and the Stringent Response: An Emerging Threat to Antibiotic Therapy. *ACS Infect Dis* 5, 1505–1517

42. Xiao, H. *et al.* (1991) Residual Guanosine 3',5'-Bispyrophosphate Synthetic Activity of relA Null Mutants Can Be Eliminated by spoT Null Mutations*. *J Biol Chem* 266, 5980–5990

43. Zhu, M. and Dai, X. (2019) Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in Escherichia coli. *Nucleic Acids Res* 47, 4684–4693

44. Michalowski, A. *et al.* (2017) Escherichia coli HGT: Engineered for high glucose throughput even under slowly growing or resting conditions. *Metab Eng* 40, 93–103

45. Ziegler, M. *et al.* (2021) Transcriptional profiling of the stringent response mutant strain E. coli SR reveals enhanced robustness to large-scale conditions. *Microb Biotechnol* 14, 993–1010

46. Chubukov, V. *et al.* (2017) Engineering glucose metabolism of Escherichia coli under nitrogen starvation. *NPJ Syst Biol Appl* 3, 16035

47. Velazquez, D. *et al.* (2022) Glucose transport engineering allows mimicking fed-batch performance in batch mode and selection of superior producer strains. *Microb Cell Fact* 21, 183

48. Heins, A.L. *et al.* (2019) Quantitative flow cytometry to understand population heterogeneity in response to changes in substrate availability in Escherichia coli and Saccharomyces cerevisiae chemostats. *Front Bioeng Biotechnol* 7, 187

49. Xiao, Y. *et al.* (2016) Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. *Nat Chem Biol* 12, 339–344

50. Rugbjerg, P. *et al.* (2021) Short and long-read ultra-deep sequencing profiles emerging heterogeneity across five platform Escherichia coli strains. *Metab Eng* 65, 197–206

51. Woo, A.C. *et al.* (2018) Heterogeneity of spontaneous DNA replication errors in single isogenic Escherichia coli cells. *Sci. Adv* 4, eaat1608

52. Sprouffske, K. *et al.* (2018) High mutation rates limit evolutionary adaptation in Escherichia coli. *PLoS Genet* 14, e1007324

53. Fijalkowska, I.J. *et al.* (2012) DNA replication fidelity in Escherichia coli: A multi-DNA polymerase affair. *FEMS Microbiol Rev* 36, 1105–1121

54. Consuegra, J. *et al.* (2021) Insertion-sequence-mediated mutations both promote and constrain evolvability during a long-term experiment with bacteria. *Nat Commun* 12, 980

55. Csörgo, B. *et al.* (2012) Low-mutation-rate, reduced-genome Escherichia coli: An improved host for faithful maintenance of engineered genetic constructs. *Microb Cell Fact* 11, 11

56. Rugbjerg, P. *et al.* (2018) Diverse genetic error modes constrain large-scale bio-based production. *Nat Commun* 9, 787

57. Meer, K.M. *et al.* (2019) High Transcriptional Error Rates Vary as a Function of Gene Expression Level. *Genome Biol Evol* 12, 3754–3761

58. Li, W. and Lynch, M. (2020) Universally high transcript error rates in bacteria. *Elife* 9, e5489

59. Mordret, E. *et al.* (2019) Systematic Detection of Amino Acid Substitutions in Proteomes Reveals Mechanistic Basis of Ribosome Errors and Selection for Translation Fidelity. *Mol Cell* 75, 427-441.e5

60. Evans, C.R. *et al.* (2019) Increased mistranslation protects E. coli from protein misfolding stress due to activation of a RpoS-dependent heat shock response. *FEBS Lett* 593, 3220–3227

61. Samhita, L. *et al.* (2020) Global mistranslation increases cell survival under stress in Escherichia coli. *PLoS Genet* 16, e1008654

62. Boo, A. *et al.* (2019) Host-aware synthetic biology. *Curr Opin Syst Biol* 14, 66–72

63. Ceroni, F. *et al.* (2018) Burden-driven feedback control of gene expression. *Nat Methods* 15, 387–393

64. Bertaux, F. *et al.* (2021) External control of microbial populations for bioproduction: A modeling and optimization viewpoint. *Curr Opin Syst Biol* 28, 100394

65. Rugbjerg, P. *et al.* (2018) Synthetic addiction extends the productive life time of engineered Escherichia coli populations. *Proc Natl Acad Sci U S A* 115, 2347–2352

66. Mori, M. *et al.* (2017) Quantifying the benefit of a proteome reserve in fluctuating environments. *Nat Commun* 8, 1225

67. Ziegler, M. and Takors, R. (2020) Minimal Cells: Design, Construction, Biotechnological Applicationspp. 1–32

68. Ziegler, M. *et al.* (2021) Engineering of a robust Escherichia coli chassis and exploitation for large-scale production processes. *Metab Eng* 67, 75–87

69. Lastiri-Pancardo, G. *et al.* (2020) A quantitative method for proteome reallocation using minimal regulatory interventions. *Nat Chem Biol* 16, 1026–1033

70. de la Cruz, M. *et al.* (2020) Plasmid DNA production in proteome-reduced Escherichia coli. *Microorganisms* 8, 1444

71. Sharma, A.K. *et al.* (2020) A novel knock out strategy to enhance recombinant protein expression in Escherichia coli. *Microb Cell Fact* 19, 148

72. Borkowski, O. *et al.* (2018) Cell-free prediction of protein expression costs for growing cells. *Nat Commun* 9, 1457

73. Yang, L. *et al.* (2018) Modeling the multi-scale mechanisms of macromolecular resource allocation. *Curr Opin Microbiol* 45, 8–15

74. Dourado, H. *et al.* (2021) On the optimality of the enzyme–substrate relationship in bacteria. *PLoS Biol* 19, e3001416

75. Wetmore, K.M. *et al.* (2015) Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *mBio* 6, e00306-15

76. Eng, T. *et al.* (2021) Engineering Pseudomonas putida for efficient aromatic conversion to bioproduct using high throughput screening in a bioreactor. *Metab Eng* 66, 229–238

77. Löffler, M. *et al.* (2016) Engineering E. coli for large-scale production – Strategies considering ATP expenses and transcriptional responses. *Metab Eng* 38, 73–85

78. Vernyik, V. *et al.* (2020) Exploring the fitness benefits of genome reduction in Escherichia coli by a selection-driven approach. *Sci Rep* 10, 7345

79. Sandberg, T.E. *et al.* (2019) The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab Eng* 56, 1–16

80. Lambros, M. *et al.* (2021) Emerging Adaptive Strategies Under Temperature Fluctuations in a Laboratory Evolution Experiment of Escherichia Coli. *Front Microbiol* 12, 724982

81. Choe, D. *et al.* (2019) Adaptive laboratory evolution of a genome-reduced Escherichia coli. *Nat Commun* 10, 935

82. Phaneuf, P. v. *et al.* (2021) Escherichia coli Data-Driven Strain Design Using Aggregated Adaptive Laboratory Evolution Mutational Data. *ACS Synth Biol* 10, 3379–3395

83. Koch, M. *et al.* (2019) Custom-made transcriptional biosensors for metabolic engineering. *Curr Opin Biotechnol* 59, 78–84

84. Wehrs, M. *et al.* (2020) You get what you screen for: on the value of fermentation characterization in high-throughput strain improvements in industrial settings. *J Ind Microbiol Biotechnol* 47, 913–927

85. Haringa, C. *et al.* (2016) Euler-Lagrange computational fluid dynamics for (bio)reactor scale down: An analysis of organism lifelines. *Eng Life Sci* 16, 652–663

86. Sievert, C. *et al.* (2017) Experimental evolution reveals an effective avenue to release catabolite repression via mutations in XylR. *Proc Natl Acad Sci U S A* 114, 7349–7354

87. Löffler, M. *et al.* (2017) Switching between nitrogen and glucose limitation: Unraveling transcriptional dynamics in Escherichia coli. *J Biotechnol* 258, 2–12

88. Neubauer, P. *et al.* (1995) Influence of substrate oscillations on acetate formation and growth yield in Escherichia coli glucose limited fed-batch cultivations. *Biotechnol Bioeng* 47, 139–146

815 89. Schellhorn, H.E. (2020) Function, Evolution, and Composition of the RpoS Regulon in
816 Escherichia coli. *Front Microbiol* 11, 560099

817 90. Sanchuki, H.B.S. *et al.* (2017) Dynamics of the Escherichia coli proteome in response to
818 nitrogen starvation and entry into the stationary phase. *Biochim Biophys Acta Proteins*
819 *Proteom* 1865, 344–352

820 91. Switzer, A. *et al.* (2018) New insights into the adaptive transcriptional response to nitrogen
821 starvation in Escherichia coli. *Biochem Soc Trans* 46, 1721–1728

822 92. Gerhardt, E.C.M. *et al.* (2015) The Bacterial signal transduction protein GlnB regulates the
823 committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of
824 acetyl-CoA carboxylase. *Mol Microbiol* 95, 1025–1035

825 93. Huergo, L.F. *et al.* (2013) PII signal transduction proteins: Nitrogen regulation and beyond.
826 *FEMS Microbiol Rev* 37, 251–283

827 94. Tan, J. *et al.* (2020) Independent component analysis of E. coli's transcriptome reveals the
828 cellular processes that respond to heterologous gene expression. *Metab Eng* 61, 360–368

829 95. Li, Z. *et al.* (2021) Transcriptional network analysis identifies key elements governing the
830 recombinant protein production provoked reprogramming of carbon and energy metabolism
831 in Escherichia coli BL21 (DE3). *Engineering Reports* 3, e12393

832 96. Li, Z. and Rinas, U. (2021) Recombinant protein production-associated metabolic burden
833 reflects anabolic constraints and reveals similarities to a carbon overfeeding response.
834 *Biotechnol Bioeng* 118, 94–105

835 97. Lynch, M. and Marinov, G.K. (2015) The bioenergetic costs of a gene. *Proc Natl Acad Sci*
836 *U S A* 112, 15690–15695

98. Li, Z. and Rinas, U. (2020) Recombinant protein production associated growth inhibition results mainly from transcription and not from translation. *Microb Cell Fact* 19, 83

99. Mori, M. *et al.* (2021) From coarse to fine: the absolute Escherichia coli proteome under diverse growth conditions. *Mol Syst Biol* 17, e9536

100. Guyet, A. *et al.* (2018) Mild hydrostatic pressure triggers oxidative responses in Escherichia coli. *PLoS One* 13, e0200660

101. Maslowska, K.H. *et al.* (2019) The SOS system: A complex and tightly regulated response to DNA damage. *Environ Mol Mutagen* 60, 368–384

102. Christodoulou, D. *et al.* (2018) Reserve Flux Capacity in the Pentose Phosphate Pathway Enables Escherichia coli's Rapid Response to Oxidative Stress. *Cell Syst* 6, 569-578.e7

103. Sandoval-Basurto, E.A. *et al.* (2005) Culture of Escherichia coli under dissolved oxygen gradients simulated in a two-compartment scale-down system: Metabolic response and production of recombinant protein. *Biotechnol Bioeng* 89, 453–463

104. Liu, D. *et al.* (2017) Enhancing fatty acid production in Escherichia coli by Vitreoscilla hemoglobin overexpression. *Biotechnol. Bioeng* 114, 463–467

105. Baez, A. *et al.* (2011) Simulation of dissolved CO2 gradients in a scale-down system: A metabolic and transcriptional study of recombinant Escherichia coli. *Biotechnol J* 6, 959–967

106. Baez, A. *et al.* (2009) Metabolic and transcriptional response of recombinant Escherichia coli to elevated dissolved carbon dioxide concentrations. *Biotechnol Bioeng* 104, 102–110

857 107. Gecse, G. *et al.* (2022) Impact of Elevated Levels of Dissolved CO2 on Performance and
858 Proteome Response of an Industrial 2′-Fucosyllactose Producing Escherichia coli Strain.
859 *Microorganisms* 10, 1145

860 108. Morales, M. *et al.* (2018) The impact of environmental factors on carbon dioxide fixation
861 by microalgae. *FEMS Microbiol Lett* 365, fnx262

862 109. Blombach, B. and Takors, R. (2015) CO2 - intrinsic product, essential substrate, and
863 regulatory trigger of microbial and mammalian production processes. *Front Bioeng*
864 *Biotechnol* 3, 108

865