Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

Modeling the multi-scale mechanisms of macromolecular resource allocation

Laurence Yang[1], James T Yurkovich[1,2], Zachary A King[1] and Bernhard O Palsson[1,2,3,4]

As microbes face changing environments, they dynamically allocate macromolecular resources to produce a particular phenotypic state. Broad 'omics' data sets have revealed several interesting phenomena regarding how the proteome is allocated under differing conditions, but the functional consequences of these states and how they are achieved remain open questions. Various types of multi-scale mathematical models have been used to elucidate the genetic basis for systems-level adaptations. In this review, we outline several different strategies by which microbes accomplish resource allocation and detail how mathematical models have aided in our understanding of these processes. Ultimately, such modeling efforts have helped elucidate the principles of proteome allocation and hold promise for further discovery.

Addresses
[1] Bioengineering Department, University of California, San Diego, La Jolla, CA, USA
[2] Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA, USA
[3] Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA
[4] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

Corresponding author: Yang, Laurence (lyang@eng.ucsd.edu)

Current Opinion in Microbiology 2018, **45**:8–15

This review comes from a themed issue on **Microbial systems biology**

Edited by **Terence Hwa** and **Uwe Sauer**

https://doi.org/10.1016/j.mib.2018.01.002

Introduction

Microbes face transiently changing environments that require the expression of new proteins and the dilution or degradation of others. To adapt to these environmental changes, cells preferentially allocate these macromolecular resources to achieve certain objectives, a process typically referred to as 'resource allocation.' The total amount and allocation of these proteins is fundamentally limited by constraints such as enzyme kinetics, cell size, and nutrient availability [1–4]. Therefore, microbes are regularly under selection pressure to optimize their resource allocation.

The macromolecular state of a cell can be measured using 'omics' technologies, allowing insights into how resource allocation changes in a given condition. Omics data have revealed a highly skewed distribution of macromolecular resource allocation. For example, the most abundant 190 proteins in *E. coli* are estimated to account for about 60% of the total protein mass [5]. The functional consequences of such skewed macromolecular compositions — and how microbes regulate their state — are open questions. Over the past few years, studies that integrate omics and mathematical modeling have increased our knowledge of how microbes allocate macromolecular resources and of the genetic basis of these allocation strategies.

In this review, we summarize the current understanding of microbial resource allocation based on recent omics measurements from the perspective of biochemical networks. We discuss how computational models have been used to elucidate the functional significance of a cellular state and how these functions are linked to a genetic basis. We close with perspectives on promising directions for future modeling studies and the potential for examining resource allocation in the context of human health.

Proteome pre-allocation provides fitness benefits at a cost

The most direct way for microbes to alter the proteome is to synthesize proteins as needed. The maximum translation rate in *E. coli* is 16-20 amino acids per second per ribosome [6,7,8], implying synthesis in ~15 seconds for a copy of protein. However, protein abundances range from ~1 to >100 000 copies per cell [9••], and ribosome abundances from ~7000 to >70 000 per cell [7]. Therefore, during a nutrient shift where hundreds of thousands of additional protein copies can be needed [9••], cells must utilize efficient strategies to dynamically allocate expression machinery resources. One strategy to minimize the delay of protein synthesis is to constitutively express proteins even when they are not immediately beneficial. This pre-allocation strategy incurs the cost of using up expression machinery that could be used to express immediately useful proteins, and a metabolic (energetic) cost of expression. Combined omics and

modeling analyses have been used to test the hypothesis of pre-allocation.

In *E. coli*, up to half of expressed protein mass potentially provides no immediate benefit for a given growth condition [10]. Even when grown on glucose minimal medium, at least 13% of the proteins expressed confer no immediate fitness benefit based on ribosomal profiling and transposon mutagenesis [11]. Genome-scale model computations suggested that pre-allocating the *E. coli* proteome toward alternative carbon sources may provide a fitness benefit when alternative carbon sources are encountered [10].

Pre-allocation also applies to expressing more expression machinery than immediately needed to ensure fast expression rates when needed in a new environment. For example, when growing *E. coli* under feast-famine cycles, growth recovery during the feast phase was maximized by strategically allocating a ribosomal protein reserve [12•].

These results suggest that omics data contain information both on the immediate response to the current environment and the regulatory program shaped by the organism's evolutionary history. Computational models help to distinguish environment-specific response from pre-programmed responses shaped by evolutionary history.

Hierarchical regulation of resource allocation robustly improves fitness

Cellular metabolism has long been recognized to be regulated through a hierarchical network of regulatory processes [13]. The slow processes of transcriptional regulation and post-translational modifications act together with fast metabolite-level allosteric regulation to control metabolic fluxes [14,15]. Metabolites also modify transcription factors (TFs), leading to a coupling between metabolism and transcriptional regulation.

Recently, there has been increasing recognition of the importance of growth rate-associated global transcriptional regulation for resource allocation. For example, a recent study [16] showed that over 90% of transcriptional changes in 100 *E. coli* genes across 26 environments could be explained by a surprisingly small number of metabolite-TF interactions along with global regulation. In response to nutrient shifts, *E. coli* was found to use a global proteome reallocation strategy [17] instead of a theoretically optimal strategy of sequentially de-bottlenecking the rate limiting enzymes [18]. This apparently conservative strategy was hypothesized to be robust by confining metabolic bottlenecks to central precursors that drive global regulatory control [17].

However, growth rate-dependent regulation is not always dominant and appears to be context specific. A recent study subjected *E. coli* to a transient nutrient stress by starvation or by switching to a lower quality carbon source [19]. The study revealed a central role for proteome allocation in triggering the 'persister' phenotype, a metabolically active but non-growing state with increased antibiotic tolerance [20]. Analysis of time-course proteomics from the nutrient-stressed cells (including persisters) and proteomics from other stress conditions (pH, temperature, and osmotic) revealed that proteome allocation was mainly driven by ppGpp-mediated regulation rather than a global growth rate effect [19]. Interestingly, a recent modeling study [21] showed that the optimal control strategy for *E. coli* to dynamically allocate resources during environmental changes involves an iterative on-off control strategy that resembles the structure of ppGpp-mediated regulation of ribosomal RNA transcription [22]. It thus appears that resource allocation under a variety of stresses may be mediated through overlapping mechanisms that are distinct from those of unstressed conditions.

Laboratory evolution aids in understanding the genetic basis of cellular resource allocation

Adaptive laboratory evolution (ALE) is an experimental method of serially passaging cells under a selection pressure. The outcome of ALE is a set of strains possessing adaptive mutations. ALE has now been automated [23], enabling large-scale production of evolved strains, followed by phenotyping and system-level characterization by DNA re-sequencing, RNA-Seq, ^{13}C-metabolic flux analysis, etc. ALE has been used to reveal the genetic basis of growth rate-selection under various conditions: different carbon sources [24], thermal stress [25], osmotic and chemical stress [26,27], oxidative stress [27], and gene knockouts [28].

Multiple studies have connected the systems-level adaptations in an evolved strain to a genetic basis. For example, strains of *E. coli* were evolved for fast aerobic growth on glucose minimal media, yielding frequent key mutations in genes including *rpoB* and *hns*. Despite the potentially broad effects of mutations in these global regulators, the strains showed little change in intracellular metabolic pathway usage. Rather, the mutations enabled higher fluxes for glucose uptake, oxygen uptake, and central carbon metabolism [29]. Interestingly, TCA cycle enzymes have been reported to be transcriptionally repressed under similar selection pressure, yet TCA cycle metabolic flux did not decrease [29]. This result suggests potential nonlinearities between transcriptome abundance, proteome allocation, and flux capacity. A genome-scale model-based analysis further suggests that this nonlinearity arises in part from the flexibility of metabolic states that support optimal growth under these conditions [29]. Specifically, simulations showed that growth at 99% of the computed maximum rate could