Characterizing the metabolic limitations of protein synthesis in model systems

Hsin-Jung (Sophia) Li

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MOLECULAR BIOLOGY

Adviser: Zemer Gitai

November 2018
Abstract

The regulation of protein synthesis is tailored to the metabolic state of the cell. This thesis explores the causes and consequences of restricted translation by ribosomes due to metabolic limitations in *E. coli* and mitochondria. First, systematic characterization of *E. coli* in balanced growth under carbon, nitrogen and phosphorus limitations at the same rate revealed three ways of tuning translation in response to nutrient limitation: carbon limitation slows translation by accumulating a large pool of inactive ribosomes; nitrogen limitation slows translation by slowing translational elongation; phosphorus limitation slows translation by reducing the total number of ribosomes. In addition to the growth-rate dependency of ribosome level as previously discovered, our study shows a surprising adaptation that phosphorous-limited *E. coli* cells can grow at the same rate as carbon- or nitrogen-limited cells with fewer ribosomes. Moreover, while manufacturing extra (unused) ribosomes can be energetically inefficient, the extra ribosomes accelerate growth when nutrients reappear. Second, examination of mitochondrial 1C metabolism unfolded a novel function of mitochondrial 5,10-methylene tetrahydrofolate for modifying mitochondrial tRNA at the wobble position. Missing the taurinomethyluridine base results in unstable codon-anticodon pairing and thus ribosome stalling on specific lysine (AAG) and leucine (UUG) codons. Together these results suggest an essential role of metabolism in regulating protein synthesis and provide new regulatory nodes in translation.
### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>0 Introduction</td>
<td>2</td>
</tr>
<tr>
<td>1 E. coli translation strategies differ across carbon, nitrogen, and phosphorus limitation conditions</td>
<td>7</td>
</tr>
<tr>
<td>1.1 Abstract</td>
<td>7</td>
</tr>
<tr>
<td>1.2 Introduction</td>
<td>8</td>
</tr>
<tr>
<td>1.3 Results</td>
<td>9</td>
</tr>
<tr>
<td>1.4 Discussion</td>
<td>23</td>
</tr>
<tr>
<td>1.5 Materials and Methods</td>
<td>27</td>
</tr>
<tr>
<td>2 Mitochondrial translation requires folate-dependent tRNA methylation</td>
<td>33</td>
</tr>
<tr>
<td>2.1 Abstract</td>
<td>33</td>
</tr>
<tr>
<td>2.2 Introduction</td>
<td>34</td>
</tr>
<tr>
<td>2.3 Results and Discussion</td>
<td>34</td>
</tr>
<tr>
<td>2.4 Materials and Methods</td>
<td>51</td>
</tr>
<tr>
<td>3 Synchronized Growth Oscillation in Glucose-Limited Bacteria</td>
<td>63</td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>63</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>64</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>65</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>71</td>
</tr>
<tr>
<td>3.5 Materials and Methods</td>
<td>74</td>
</tr>
<tr>
<td>4 Conclusions</td>
<td>77</td>
</tr>
<tr>
<td>Appendix A Supporting Information for Chapter 1</td>
<td>79</td>
</tr>
<tr>
<td>Appendix B Characterization of E. coli physiology across different nutrient conditions and growth rates</td>
<td>96</td>
</tr>
<tr>
<td>Appendix C Gene expression of E. coli across different nutrient conditions and growth rates by microarray</td>
<td>101</td>
</tr>
<tr>
<td>Appendix D Measurement of metabolite pool sizes in E. coli across different nutrient conditions and growth rates by LC/MS</td>
<td>103</td>
</tr>
<tr>
<td>Appendix E Measurement of metabolite pool sizes in synchronized Caulobacter crescentus during the cell cycle by LC/MS</td>
<td>106</td>
</tr>
<tr>
<td>Appendix F The effect of dksA on ribosome abundance and activity in E. coli</td>
<td>109</td>
</tr>
<tr>
<td>References</td>
<td>124</td>
</tr>
</tbody>
</table>
Chemostat enables growing *E. coli* at a wide range of growth rates by changing the dilution rate and different nutrient limitations. (a) A picture of a chemostat experimental setup. (b) An illustrated chemostat and the ODE model for cell concentration inside the vessel.

**1.1 RNA-to-protein ratio is both growth-rate and nutrient dependent.** (a) Schematic flow of nutrients for biomass formation. Carbon (C) and nitrogen (N) combine to make amino acids. Amino acids combine with carbon precursors and phosphorus (P) to make nucleic acids. (b) RNA-to-protein ratios for chemostat cultures under C-, N-, and P-limitations at different growth rates. Each data point shows the mean value from three technical replicates. (c) Translation elongation rates (amino acids/s) as measured by the *lacZ* induction assay after correction for translation initiation. The bar height represents mean values with error bars as SEM from three biological replicates. (d) Polysome profiles of cells grown under C-, N-, and P-limitations in chemostats at dilution rate 0.1 h\(^{-1}\). Five independent experiments were repeated, with a single representative curve shown. (e) Quantification of ribosomes in the form of subunits (30S + 50S), free 70S, mRNA-bound 70S (one ribosome on one mRNA), and polysomes (multiple ribosomes on one mRNA). The bar height represents mean values with error bars as SEM from three biological replicates. (f) Cells adapt to different nutrient limitations using different strategies of translational regulation that achieve the same protein production rate.

**1.2 Lower RNA-to-protein ratio under P-limitation results from lower RNA concentration, and phosphorus metabolism-related genes are not involved in the regulation.** (a) RNA-to-protein ratio in chemostat and batch conditions. Each data point shows the mean value from three technical replicates. (b) Comparison of RNA-to-protein ratio values between this study and previously reported values from the same strain of *E. coli* NCM3722. Same conditions are connected by the dotted line. (c-d) Total protein (µg/mL/OD600) and RNA (µg/mL/OD600) concentrations were measured. These data were used to generate Fig. 1.1b of the RNA-to-protein ratio. Each data point represents the mean value from three biological replicates. (e-f) Fraction of rRNA (e) and tRNA (f) in total RNA determined from bioanalyzer assay across different nutrient limitations and growth rates. Each data point shows the mean value from three biological replicates with error bars as SEM. (g-h) RNA-to-protein ratio for chemostat cultures of wild-type and mutant cells upon N-, and P-limitations at different growth rates. Each data point shows the mean value from three technical replicates: (g) ΔphoB, (h) Δppk.

**1.3 lacZ induction curves and fittings for translational elongation rate measurements.** (a) Raw data of *lacZ* induction time course for batch glucose minimal media (Min-0.9), C-, N-, and P-limited cells grown at 0.1 h\(^{-1}\). Three biological replicates were performed and data from one representative replicate is shown for clarity. (b) Transformed *lacZ* induction data for lag time estimation. All three biological replicates are shown with estimated lag time indicated in the bracket (See Methods for detail).
1.4 Elevated [K\(^+\)] distinguishes mRNA-free from mRNA-bound 70S ribosomes. (a) The plot shows the separation of these two distinct 70S species under different growth conditions using 170 mM K\(^+\). The A254 values are vertically shifted for each condition for easier visualization. Puromycin is used as a positive control for free ribosome accumulation and fast-growing cells in glucose minimal media (Min-0.9) are used as a negative control for cells expected to lack free ribosomes. At least two independent experiments were performed with similar results. (b) Fraction of assembled ribosomes under different growth conditions. Chemostat growth conditions include C-, N-, and P-limitations grown at 0.1 and 0.6 h\(^{-1}\). Batch conditions include glucose minimal media (Min-0.9) and defined rich media (Rich-1.7). Mean values from three biological replicates are shown with error bars indicating the standard error of the mean.  

1.5 Deletion of raiA does not alter ribosome dynamics. (a) The ratio of RNA-to-protein ratio of wild type and ΔraiA at growth rate 0.1 h\(^{-1}\) under different nutrient limitations. Each bar represents the ratio of the mean R/P ratio from three technical replicates of wild type and the mutant. (b) Polysome profiles of wild type and ΔraiA cells. Same amount of RNA from WT and ΔraiA was loaded for comparison. Deletion of raiA has no effect on the size of 70S peak. 100S peak from ribosome dimerization was not detected under any condition. The thin dashed line marks the average baseline. (n = 1 independent biological sample).  

1.6 LA macroscopic model reveals different ribosome dynamics that achieve the same growth rate. (a) Averaged A-site ribosome counts within the first and last 50 codons of transcripts from ribosome profiling analysis. Ribosomes bound to the first 10 codons are defined as "initiating ribosomes" (\(R_i\)) and those bound to the rest of transcripts except the stop codon are "working ribosomes" (\(R_w\)). (n = 1 independent biological sample) (b) A macroscopic model of ribosome dynamics. \(k_f\) forward rate (s\(^{-1}\)); \(k_p\) proceeding rate (s\(^{-1}\)); \(k_{el}\) elongation rate (aa/s); \(N_{aa}\) number of amino acids in an average protein of \(E. coli\); \(R_t\) total number of ribosomes. (c) The relative proceeding rate under C-, N-, and P-limitations at growth rate 0.1 h\(^{-1}\). The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements. (d) The saturation parameter under C-, N-, and P-limitations at growth rate 0.1 h\(^{-1}\). The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements. (e) The relationship between elongation rate (\(k_{el}\)), fraction of working ribosomes (\(\Phi_{R_w}\)) and total number of ribosomes (\(R_t\)) that leads to the same growth rate at 0.1 h\(^{-1}\). Colored dots indicate the values for C-, N-, and P-limited wild type cells.  

1.7 Modeling ribosome dynamics on an mRNA with a microscopic model. (a) Comparison of ribosome density between first and second halves of genes. Scatter plot of summed RPM (reads per million) in the first and second halves of each gene. Blue line has a slope of one. Red line is the fitting result with the value of slope marked in the graph. (n = 1 independent biological sample.) (b) Illustration of the microscopic model for bound ribosome dynamics on mRNA, and the corresponding flux balance equation. (c) Dark bars are the normalized ribosome count in the first 50 and the last 50 codons from ribosome profiling studies. Red curves are the smoothed ribosome occupancies obtained by fitting the normalized ribosome count in the first 50 and the last 49 codons to two-term exponential functions, and copying the observed count number at the stop codon. (d-e) The position-dependent normalized step rate (s\(^{-1}\)) and fraction of aborted translation that best fit the smoothed ribosome occupancies in (c). Insets are magnifications for codon position 1-10.
1.8 **Parameters of the macroscopic model relate translation activity and growth.** (a) Comparison of predicted elongation rates from the macroscopic ribosome dynamics model with experimentally measured ones. The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements. (b-c) The relative proceeding rate \((N_{ab} \cdot \frac{k_p}{k_e})\) and saturation parameter \((R_i/(K_m + R_i))\) under various conditions. The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements. (d-e) Model relationships between total number of ribosomes and growth rate, while the total number of mRNAs is held constant. Dots mark the estimated values of ribosomes under each condition, and dashed line in (e) marks the half-saturation number of ribosomes under Min condition. The inset in (e) shows the region around the open circle. (f) The relationship between elongation rate \((k_e)\), fraction of working ribosomes \((\Phi_{RW})\), and total number of ribosomes \((R_t)\) that leads to the same growth rate of 0.1 h\(^{-1}\). Filled circles indicate the values for C-, N-, and P-limited wild type cells, and open circles indicate the corresponding values for the \(\Delta relA\) mutant.

1.9 **Deletion of relA disrupts translation regulation under nitrogen limitation.** (a) A-site codon occupancy under different growth conditions from ribosome profiling. Occupancy was calculated as the ratio between measured and expected counts for each gene based on codon frequency. The average of this ratio is plotted (with number of genes specified). Codons that have higher than average ratio by 2.5 standard deviations are highlighted. (b) Fraction of assembled (70S) ribosomes in wild type and \(\Delta relA\) under C-, N-, and P-limitations at growth rate 0.1 h\(^{-1}\). The bar height represents mean values with error bars as SEM from three biological replicates. (c) \(lacZ\) induction assay for wild type and \(\Delta relA\) under nitrogen limitation at growth rate 0.1 h\(^{-1}\). The lag time measures when the first functional LacZ is produced and is inversely proportional to the elongation rate. Mean values from three biological replicates are shown with error bars as SEM. (d) Ratio of codon occupancy between \(\Delta relA\) and wild type under nitrogen limitation at growth rate 0.1 h\(^{-1}\) (n = 1 independent biological sample). (e) Ratio of the relative proceeding rate \((N_{ab} \cdot \frac{k_p}{k_e})\) between \(\Delta relA\) and wild type across different conditions at growth rate 0.1 h\(^{-1}\). The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements.

1.10 **relA-dependent ppGpp production leads to free ribosome accumulation under stringent response and nitrogen limitation.** (a) Polysome profile of wild type and \(\Delta relA\) cells with and without treatment of serine hydroxamate (SHX) for ten minutes. Two independent repeats were performed and one representative data is shown. (b) Free-ribosome profiling using 170 mM KCl to distinguish mRNA-free and mRNA-bound ribosomes. (n = 1 independent biological sample.) (c) Polysome profile of wild type and \(\Delta relA\) cells at growth rate of 0.1 h\(^{-1}\). Three independent repeats were performed and one representative data set is shown. (d) Ratios of R/P ratios of wild type and \(\Delta relA\) at growth rate of 0.1 h\(^{-1}\). The bar shows the ratio of the mean R/P ratio from three technical replicates of wild type and the mutant.

1.11 **Deletion of relA disrupts translation under nitrogen limitation.** (a) Ratio of codon occupancy between \(\Delta relA\) and wild type cells under carbon limitation at 0.1 h\(^{-1}\). (b) Ratio of codon occupancy between \(\Delta relA\) and wild type cells under phosphorus limitation at 0.1 h\(^{-1}\). (c) Cumulative fraction of ribosome counts of fabI in wild type and \(\Delta relA\) under C-, N- and P-limitations at 0.1 h\(^{-1}\). The vertical dotted lines and red triangles mark the positions of glutamine codons.
1.12 Extra ribosomes confer growth advantage upon nutrient upshift. (a) Theoretically-predicted growth curves from the macroscopic model. (b) Experimental nutrient-upshift growth curves in LB + 0.4% glucose for wild type cells from C-, N-, and P-limited chemostats at growth rate 0.1 h⁻¹. The initial OD and standard deviation of three biological replicates for each condition are 0.141±0.008 (C), 0.107±0.017 (N), and 0.154±0.008 (P) respectively. Each data point shows the mean value from three biological replicates with error bars as standard deviation. (c) Model for nutrient-dependent ribosome usage: Under C- or N-limitation, the total ribosome pool is high while under P-limitation the total pool is low. C-limited cells elongate fast but have a low fraction of working ribosomes. N-limited cells elongate slower but have a higher fraction of working ribosomes than C-limited cells. Under P-limitation the low supply of ribosomes leads to both a high fraction of working ribosomes and fast elongation to meet the protein production demand.

1.13 Different allocation strategies determine growth dynamics upon modeled nutrient upshift to rich medium. After nutrient upshift to a rich condition, different regulatory strategies for Ψ_R(t), the fraction of new protein synthesis allocated to ribosomal proteins (first three rows for C-, N-, and P-limitations, respectively), and their resulting post-upshift growth curves (last row).

2.1 SHMT2 deletion-induced respiratory chain dysfunction in different cellular backgrounds and clones. (a) Change in media colour after 48 h cell growth. (b-c) Lactate secretion (b) and normalized NAD+/NADH ratio (c) of HCT116 knockout cell lines (n = 6). (d-e) Basal respiration as measured by Seahorse XF analyser (n = 3) (d) and normalized NAD+/NADH ratio (n = 3) (e) of HEK293T folate 1C gene CRISPR–Cas9 knockout cell lines. (f) Normalized levels of TCA cycle and associated metabolites (n = 3). (g) Steady-state labelling fraction into citrate from [U-13C]substances glutamine (left) and glucose (right) (n = 3). (h) Immunoblot of extracted mitochondria for subunits of respiratory chain complexes I–V (CI–CV) and markers of mitochondrial mass. (i) Mitochondrial complex I levels (NDUFS4) in independent HCT116 folate 1C gene knockout clones. Data are mean ± s.e.m. n indicates the number of biological replicates. *P <0.01, two-tailed Student’s t-test.

2.2 Mitochondrial respiratory chain function is dependent on SHMT2 catalytic activity. (a) 1C pathway and known mitochondrial products. (b) Lactate secretion of HCT116-knockout cell lines (n = 6). ΔSHMT2-A and ΔSHMT2-B denote two separate SHMT2-knockout lines. WT, wild type. (c) Oxygen consumption rate measured by Seahorse XF analyser (n = 3). FCCP denotes a mitochondrial uncoupling agent. Oligom., oligomycin; Rot./antim., rotenone/antimycin. (d) Immunoblot for mitochondrial respiratory complex I and II (CI and CII) proteins (NDUFS4 and SDHA, respectively), 1C enzymes, and a marker of mitochondrial mass (VDAC1). (e) Basal respiration (n = 3) upon re-expression of wild-type or catalytically deficient mutant forms of SHMT2 in HEK293T knockout cell lines. Data are mean ± s.e.m. n indicates the number of biological replicates, which for the Seahorse experiments refers to independent plates on separate days. *P <0.01, two-tailed Student’s t-test (see Supplementary Table 7 for exact P values). Cat. inact., catalytically inactive SHMT2; PLP bind., PLP binding-deficient SHMT2.

2.3 Catalytically deficient SHMT2 constructs. (a) Mapping of mutated amino acid residues on human SHMT1 (PDB code 1BJ4) using iCn3D and alignment of E. coli serine hydroxymethyltransferase (GLYA), H. sapiens mitochondrial serine hydroxymethyltransferase 2 (GLYM) and cytosolic serine hydroxymethyltransferase 1 (GLYC). Positions for GLYM are given with reference to GenBank NM_005412.5. (b) Sanger sequencing traces of mutant constructs. (c) Immunoblot for mitochondrial complex I levels (NDUFS4) in cell lines re-expressing catalytically deficient forms of SHMT2.
2.4 Restoring SHMT2 catalytic activity normalizes 1C flux, respiratory chain expression, glycolytic activity, and cell growth. (a) Immunoblot of re-expression of catalytically active SHMT2 (left) and the effects of its re-expression on mitochondrial complex I and II levels (right). (b-f) Effect of re-expression of catalytically active and inactive forms of SHMT2 in two different ΔSHMT2 clones in the HEK293T background. b, Normalized NAD+/NADH ratio (n = 6). c, Lactate secretion and glucose uptake (n = 6). d, Cell proliferation (n = 6). e, Purine biosynthesis intermediate 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) levels (n = 4) as an indicator of cytosolic folate 1C status. f, [2,3,3-²H]serine tracing to differentiate cytosolic from mitochondrial folate 1C unit production for incorporation into deoxynthymidine triphosphate (n = 3). Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.

2.5 SHMT2-knockout-induced respiratory chain deficiency is caused by mitochondrial methylene-THF depletion but is unrelated to dTTP synthesis. (a) Sarcosine serves as an SHMT2-independent source of mitochondrial methylene-THF. (b) NAD+/NADH ratio (n = 6) and NDUFS4 (complex I) protein expression upon sarcosine supplementation (1 mM) in SHMT2 single-knockout (ΔSHMT2) and SHMT2/MTHFD2 double-knockout (ΔSHMT2/ΔMTHFD2) cell lines compared to wild-type cells. (c-d) Functional readouts for mitochondrial dTTP status based on mitochondrial DNA (mtDNA) levels (n = 3) determined by quantitative PCR (qPCR; c) and gene expression determined by RNA-seq (d) in SHMT2-knockout and wild-type HEK293T cells. RPKM, reads per kilobase per million mapped reads. In d, each data point represents the mean gene expression of two biological replicates of two independent knockout clones (n = 4) and two wild-type replicates (n = 2). Genes linked to OXPHOS function are highlighted in red (nuclear-encoded) or blue (mitochondrial-encoded). Data are mean ± s.e.m. n indicates the number of independent biological replicates. *P < 0.01, two-tailed Student’s t-test.

2.6 Oxidative phosphorylation defect is caused by a post-transcriptional mechanism independent of methionine formylation. (a) Fraction of initiating amino acid (formyl-methionine versus methionine) of mitochondrial-expressed COX1 peptide determined by high-resolution LC–MS (wild type n = 4, ΔSHMT2 n = 3, ΔMTHFD2 n = 2). (b) Lactate secretion (n = 3) upon sarcosine supplementation (1 mM). (c) Relative mtDNA levels in HEK293T cells (n = 3). (d) Agarose gel of mtDNA long-range PCR products of HCT116 and HEK293T knockout cell lines. (e) Relative mRNA levels of mtDNA-encoded respiratory chain subunits in the HEK293T background (n = 3). (f) Gene expression levels in SHMT2-knockout cell lines compared to SHMT2 wild-type re-expressed lines by total RNA sequencing. Each dot represents mean gene expression as derived from two biological replicates of two independent knockout clones and matched re-expressed lines (n = 4). Genes linked to human OXPHOS function are highlighted in red. (g) Position-dependent next-generation sequencing coverage of mtDNA in HEK293T wild-type, SHMT2-knockout and MTHFD2-knockout cell lines supports the absence of deletions due to SHMT2 loss. (h) Corresponding variant position and frequency. Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.

2.7 Impairment of mitochondrial translation due to loss of SHMT2. (a) SDS–PAGE of [³⁵S]methionine-labelled mitochondrially translated proteins in wild-type (lane 1) and two SHMT2-knockout (lane 2 and 3) HEK293T cell lines. Decreased synthesis of COX1 and COX2/3 are evident upon short exposure and reduced synthesis of ND5 and ND6 is more easily visualized upon longer exposure. (b) Absorbance at 254 nm upon sucrose gradient fractionation of cell lysates digested by micrococcal nuclease (Fig. 2.8a). Fractions corresponding to 4 and 5 were collected for mitochondrial ribosome enrichment as shown on the matched immunoblot for mitochondrial ribosome subunit MRPL11. (c) Read length distribution (top) and read length-dependent sub-codon read phasing (bottom) across the 13 mitochondrial protein-coding transcripts. Data in c are based on the mitochondrial ribosome profiling experiment in Fig. 2.8, and represent the mean of two technical replicates of two independent samples.
2.8 Mitochondrial ribosome profiling reveals that SHMT2-knockout cells are deficient in translating specific guanosine-ending codons. (a) Workflow of mitochondrial ribosome profiling. Translation was halted using chloramphenicol and immersion into liquid nitrogen, cells were lysed and RNA was digested using micrococcal nuclease (MNase). After sucrose-gradient enrichment for mitochondrial ribosomes (shaded in red), protected fragments were sequenced. (b) Mean cumulative ribosome density along selected mitochondrial transcripts. Additional transcripts are given in Fig. 2.9a. RPF, ribosome-protected fragment; RPM, reads per million. (c) Mean codon-specific mitochondrial ribosome occupancy in HCT116 cells (ΔSHMT2/wild type). Red data points correspond to codons that are decoded by tRNAs carrying the 5-taurinomethyluridine (τm5U) modification. Red labels correspond to the subset of codons that end in guanosine and thus require wobble-base pairing. Methionine codons are highlighted in blue and show no increased codon occupancy. The insert shows mean normalized ribosome density relative to UUG and AAG codon position. aa, amino acids. Data in b and c represent two technical replicates of two independent samples.

2.9 Impairment of mitochondrial translation due to loss of SHMT2. (a) Expanded version of Fig. 2.8b, showing the mean cumulative ribosome protected fragments of all mitochondrial protein-coding genes. (b) Mean relative density of actively translating (that is, not stalled) ribosomes for mitochondrial transcripts. Data in a and b represent two technical replicates of two independent samples. (c) Enzymatic activities of citrate synthase and individual mitochondrial respiratory chain complexes from mitochondrial extracts (n = 5). Data are mean ± s.e.m. *P < 0.01, two-tailed Student’s t-test. (d) Mitochondrial genetic code table with split codon boxes depending on taurinomethylated tRNAs for translation highlighted in red. Codons decoded by anticodon formylcytidine-containing tRNA Met are highlighted in blue. (e) Mean codon-specific mitochondrial ribosome occupancy of HCT116 SHMT2/MTHFD2 double-knockout cell lines supplemented with sarcosine (1 mM). Codons highlighted in red are decoded by tRNAs carrying a 5-taurinomethyluridine modification. The supplementation with sarcosine prevents the stalling normally observed with SHMT2 deletion (n = 2).

2.10 MTO1/GTPBP3-dependent tRNA methylation requires mitochondrial methylene-THF. (a) Interaction of tRNA position 34 anticodon loop modified base with mRNA codon 3 position A/G, forming a non-Watson – Crick base pair. (b) Total ion chromatogram of 5-taurinomethyluridine monophosphate (p-τm5U) (m/z = 460.043) from digested mitochondrial tRNAs. 5-formylcytidine monophosphate was not altered (Fig. 2.11a). (c) Mean codon-specific mitochondrial ribosome occupancy for MTO1-knockout (ΔMTO1) HCT116 cell lines and primary patient-derived fibroblasts carrying MTO1 mutations or the MT-TL1 m.3243A>G MELAS variant (n = 2). Corresponding immunoblots are shown below. Individual patient data are in Fig. 2.13a. (d) Basal respiration rates measured using the Seahorse XF analyser. Data were collected after growth in the absence (−) of folate for 5 passages or in the presence of the indicated methotrexate (MTX) concentration for 96 h (n = 3, except HCT116 WT n = 4 and MTX 50 nM n = 6). HT, 100 μM hypoxanthine and 16 μM thymidine. Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.
2.11 tRNA modification status in ΔSHMT2 and effects of 5-taurinomethyluridine modification loss caused by human disease gene MTO1. (a) Total ion chromatogram of 5-formylcytidine monophosphate in digested mitochondrial tRNAs upon loss of SHMT2. The same samples were analysed for 5-taurinomethyluridine monophosphate (p-τm5U) in Fig. 2.10b. The combined data demonstrate that SHMT2 deletion causes loss of τm5U but not 5-formylcytidine. (b) Levels of τm5U, 5-taurinomethyl-2-thiouridine monophosphate (p-τm5s²U) and 2-thiouridine monophosphate (p-s²U) in wild-type HCT116 and SHMT2 deletion lines normalized to 5-formylcytidine monophosphate (p-f5C) (n = 3). (c) Taurine levels in HCT116 wild-type and SHMT2-knockout cells (n = 3). (d) τm5U levels in digested mitochondrial tRNAs upon re-expression of SHMT2 (n = 1). (e) τm5U, τm5s²U and s²U levels normalized to f5C in HCT116 SHMT2/MTHFD2 knockout lines after sarcosine supplementation and HCT116 upon loss of MTO1 (n = 2). For all panels, data are mean ± s.e.m. or individual data points only. (f) Labelling pattern of 5-taurinomethyluridine and 5-formylcytidine monophosphate extracted from mitochondrial tRNAs after growth in media containing either [3-¹³C]serine or [U-¹³C]methionine. (g) Mean cumulative count of ribosome protected fragments (RPF) mapping to mitochondrial protein coding transcripts upon ribosome profiling in HCT116 MTO1-knockout cell lines. Data were normalized to RPM (n = 2); n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.

2.12 Investigation of mRNA and protein secondary structure effects on mitochondrial ribosome stalling sites. (a) Identification of mitochondrial RNA secondary structure based on analysis of the mitochondrial transcript data from the dimethyl sulfate sequencing dataset published previously. R values and Gini differences were calculated to detect changes in nucleotide reactivity between the in vivo and denatured condition for the complete mitochondrial transcriptome. (b) Determination of ribosome stalling sites in SHMT2-knockout HCT116 cell lines. Data points represent individual codons of all 13 mitochondrial protein-coding transcripts. For each codon, the y axis indicates the ratio of normalized counts in SHMT2-knockout to normalized counts in wild-type HCT116. Two and three s.d. above the mean of all codons in the genome are indicated by the grey and black dotted line, respectively. Highlighted in red are codons with greater than 2 s.d. (c) Mapping of AAG and UUG codons from SHMT2 knockout-specific ribosome stalling sites (>3 s.d.) on protein structures. For b and c, analysis is based on ribosome profiling data in Fig. 2.8, with two technical replicates of two independent samples.

2.13 Mitochondrial transcript codon occupancy from ribosome profiling of individual patient lines. (a) Codon-specific mitochondrial ribosome occupancy ratio (patient/control fibroblasts) in individual patient derived cell lines (n = 1 for each individual patient, normalized to mean of n = 2 control fibroblast lines). Patients either had nuclear MTO1 missense mutations (patient A c.[1261-5T>G];[1430G>A], patient B c.[1222T>A];[1222T>A]) or were diagnosed with MELAS and carry the recurrent point mutation m.3243A>G in the mitochondrial gene for tRNA Leu1 (MT-TL1). (b) Next-generation sequencing of mtDNA mutation load m.3243A>G (MT-TL1) in control fibroblasts and MELAS patient cell lines. Each bar shows one biological replicate for control and patient cell lines. Integrative genomics viewer sequencing raw data are shown on the right.
2.14 Effects of targeting 1C metabolism on mitochondrial function. (a) Mitochondrial complex I and II levels after growth in the absence of folate for five passages or in the presence of the indicated methotrexate concentration for 96 h. Ethidium bromide (250 nM) was used as a positive control. (b) Cellular mtDNA levels in HCT116 cells after folate depletion (with or without 100 μM hypoxanthine and 16 μM thymidine (HT) as rescue agents) or in the presence of methotrexate for 96 h (n = 3). (c) To determine whether the decrease in respiration due to methotrexate arises from methotrexate depleting mitochondrial DNA, impairing mitochondrial translation, or a combination, in HCT116 cells we compared the effects of methotrexate (50 nM) to ethidium bromide (250 nM = 100 ng ml−1), which is classically used to deplete mitochondrial DNA, and to chloramphenicol (310 μM = 100 μg ml−1), which blocks mitochondrial translation. After 48 h of treatment, methotrexate and ethidium bromide both decreased oxygen consumption and DNA content. Importantly, despite ethidium bromide depleting mitochondrial DNA much more strongly, methotrexate had an equivalent effect on oxygen consumption, consistent with the effect of methotrexate on oxygen consumption being in part via mitochondrial translation inhibition. Data are normalized and compared to untreated control (all n = 3; except oxygen consumption methotrexate 96 h n = 6 and control n = 4). Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . ..52

3.1 E.coli exhibits oscillatory oxygen consumption and cell density in glucose-limiting chemostats. (A) The oxygen level in chemostats oscillates and the period is correlated with the dilution rate. (B) The cell density (OD600) and oxygen level during oscillation in a chemostat with dilution rate 0.6 h⁻¹. The blue dash line in OD panel is the population density at steady state. Red points are samples used for microarray analysis. 65

3.2 Functional gene clusters change expression levels during the metabolic cycle. (A) Heat map of gene expression level from microarray data. The data is centered to the row means. 24 samples in total were collected and corresponding OD and oxygen levels were plotted at the top. (B) Principle component analysis shows two major PCs and their trends are plotted. (C) Averaged gene expression pattern as a function of time for the three clusters defined in (A). The error bar shows the standard error. (Cluster 1: n=1118, Cluster 2: n=841, Cluster3: n=267). The smoothing curves are generated using loess smoothing function with grey area representing the 95% confidence interval. 72

3.3 Metabolite pool sizes change during the period of metabolic cycle. Metabolite levels are normalized to the row means and plotted in log2 scale. The corresponding OD 600 and pO2 values for the 18 samples are shown at the top. 73

3.4 Functional gene clusters in metabolic cycle exhibit growth-rate dependency. The black line shows the distribution of growth-rate dependent coefficient of gene expression under carbon-limitation. The enrichment of genes in Cluster 1 and 2 (defined in Figure 3.2) is plotted as shaded areas. The median and mean of the coefficients are plotted as solid and dashed lines respectively. 73

B.1 Growth rate prediction using physiological measurements agree with actual ones across a wide range of growth conditions. (a) Translational elongation rates measured by lacZ induction assay. (b) The fraction of bound ribosomes (bound 70S monosome + polysomes). (c) RNA-to-protein ratio. (d) Comparison of actual and predicted growth rates. The dashed line has a slope equal to 1. (e) RNA-to-DNA ratio. (f) DNA-to-protein ratio. All data points are averages from three biologically independent samples. Error bars are s.e.m. In (d) all individual samples are plotted. C: glucose-limitation; N: ammonia-limitation; P: phosphorus-limitation; LB: ΔleuB mutant in leucine-limitation; RA: ΔrplA mutant in ammonia limitation; M: Batch glucose minimal media; R: Batch defined rich media. All samples were grown in 40 mM MOPS media. 98
B.2 Fraction of different RNA species change across growth conditions. (a) Fraction of rRNA in total RNA. (b) Fraction of tRNA in total RNA. All data points are averages from three biologically independent samples. Error bars are s.e.m. C: glucose-limitation; N: ammonia-limitation; P: phosphorus-limitation; LB: leuB mutant in leucine-limitation; RA: rplA mutant in ammonia limitation; M: Batch glucose minimal media; R: Batch defined rich media. All samples were grown in 40 mM MOPS media.

B.3 Cell size parameters are dependent on more than the growth rate. (a) The average and standard deviation of normalized values for the cell length and width across conditions. (b) Averaged 3D cell volumes inferred from cell widths and lengths. (c) Average surface area-to-volume ratio. All data points are from more than 500 cells in one biological sample. C: glucose-limitation; N: ammonia-limitation; P: phosphorus-limitation; LB: leuB mutant in leucine-limitation; RA: rplA mutant in ammonia limitation; M: Batch glucose minimal media; R: Batch defined rich media. All samples were grown in 40 mM MOPS media.

C.1 Gene expression of *E. coli* is dominated by nutrient limitation than growth rate. (a) Hierarchical clustering of *E. coli* gene expression across glucose (C-), ammonia (N-), and phosphorus (P-) limitations. Six growth rates are measured from 0.1 to 0.6 h\(^{-1}\) with 0.1 h\(^{-1}\) increment. Log2 ratio is plotted. (b) Heatmap of gene expression correlation between the samples. Samples under the same limitations are closer to each other. (c) Fraction of variance explained by principle components of Principle Component Analysis. (d) Eigen-gene expression profile of the first three principle components. Growth rate is captured mostly in PC3.

D.1 Metabolite pool sizes in *E. coli* cells across different growth rates and nutrient limitations. (a) Hierarchical clustering of *E. coli* metabolites across glycerol, glucose, ammonia (N), and phosphorus (P) limitations. Six growth rates are measured from 0.1 to 0.6 h\(^{-1}\) with 0.1 h\(^{-1}\) increment. Log2 ratios to the row means are plotted. (b-d) Critical metabolic ratios relating to energy charge (ATP, b), the reducing factor (NADPH, c) and the electron acceptor (NAD+, d). The values were calculated from the absolute level of metabolite pool sizes. Batch cultures of acetate, glycerol, and glucose minimal media are plotted for comparison.

D.2 Absolute concentrations of metabolites in *E. coli* cells under different nutrient limitations and growth rates. Cells were grown in chemostat until steady states under different nutrient limitations. Cultures were collected using filter membranes and quenched with -20°C 40:40:20 (Acetonitrile:Methanol:Water) solution. Metabolites annotated with (a) indicate the acidic quenching condition (with 0.5% Formic acid and neutralized with 15% NH\(_4\)HCO\(_3\) with the ratio of 800: 67). Samples were analyzed on an LC/MS system with an Exactive Orbitrap MS (Thermo). A U-\(^{13}\)C glucose Gutnick minimal media culture was extracted on the same day with the samples and mixed 1:1 OD ratio before LC/MS analysis. The absolute concentration was converted using the ion count ratio between the sample and the glucose minimal media culture that has previously determined absolute metabolite pool sizes. Glycerol-, glucose-, and ammonia (N-) limited cultures were grown in 1x Gutnick minimal media. Phosphorus- (P-) limited cells were grown in 8 mM MOPS minimal media (5x dilution of normal 40 mM) to lower the effect of MOPS on MS. While no other obvious phenotype differences observed, the total metabolite pool size seems to be smaller in P-limited samples than the others, potentially due to the limitation of osmolarity. The right three bars are derived from previously determined glucose minimal media data with upper bound (UB) and lower bound (LB).
E.1 **Metabolite pool sizes as a function of the cell cycle in a synchronized *Caulobacter crescentus* culture.** Cells were synchronized using centrifugation (~2 hr). At each time point, 3 mL culture was collected using a filter membrane and quenched in -20°C 40:40:20 (Acetonitrile:Methanol:Water) solution. Samples were analyzed on an LC/MS system with an Exactive Orbitrap MS (Thermo). The signals were normalized to an unsynchronized culture which did not undergo mock synchronization. The log2 ratios are plotted with the color bar indicating the scale. The top images showed the representative morphology of cells at each time point.

F.1 **Deletion of dksA affects ribosome abundance but does not abort growth-rate dependent RNA-to-protein ratio under phosphorus limitation.** (a) RNA-to-protein ratio of wild type and dksA deletion mutant for chemostat cultures upon P-limitations and batch cultures at different growth rates. All of the media is supplemented with amino acids due to the amino acid auxotrophy of dksA deletion. Each data point represents three technical replicates. (b) Quantification of assembled ribosomes in the form of free 70S, mRNA-bound 70S (one ribosome on one mRNA), and polysomes (multiple ribosomes on one mRNA) at growth rate 0.1 h⁻¹ under P-limitation.
Acknowledgments

The journey toward getting a Ph. D. does not start on day one of graduate school. The seed of the curiosity and the passion for science to seek the truth about the world by one’s own hands and mind is planted in childhood and cultivated throughout a person’s development. There are many people to acknowledge for their help along the way, and I would like to present them in reverse chronological order. It is going to be long, because this is my only precious opportunity to do their efforts justice.

First, I would like to thank the present and former lab members from the Gitai, Wingreen, Rabinowitz and Botstein labs. Princeton University has been a rare and fantastic environment in which I got to work with people from different labs and disciplines. I would like to thank Dr. Max Wilson and Dr. Rachael Barry, the other members of the metabolism team in the Gitai lab. I thank Max for his guidance at the beginning when I first joined the lab, and the encouraging words he wrote to me on the front page of Contact, which he gave me as a gift at a difficult time. I thank Drs. Rachael Barry, April Williams and Julia Wittes for their lively conversations and support. Rachael and Julia have helped me through many daunting writing tasks. I thank Dr. Albert Siryaporn, whose attitude of building one’s own tools has shaped my approach to experiments. Biological experiments can become sterile and mechanical routines when everything is commercialized and standardized. Albert’s passion for developing tools that fit one’s own needs and his audacity in opening the black box of analysis codes or any piece of equipment inspire me to bring a hacker’s mindset into my daily life as a scientist. I want to thank Drs. Nick Ouzounov and Alex Lorestani, the two founders of an amazing company, Geltor. Nick has saved me numerous times from the despair of failed attempts at cloning. His creativity and original scientific ideas are impressive. Alex is the most friendly person I know. He is always optimistic in the face of uncertainty, and his cheerful attitude is contagious. I would also like to thank present and former Gitai lab members for their input and helpful discussions: Edith Blackman, Joseph Sheehan, Dr. Ben Bratton, Dr. Matthias Koch and Dr. Alex Persat.

I would like to thank Professor Joshua D. Rabinowitz for being on my graduate thesis committee and also mentoring me in his lab. I have learned so much about metabolism from his lab meetings and his teaching throughout the years. His sharp insights and suggestions have helped me focus on the most critical data from a pile of experiments. I would like to thank Li Chen, Dr. Gregory Ducker, Jonathan Ghergurovich, Dr. Lukas Tanner, Dr. Chel Nofal, and Dr. Jun Park for thoughtful conversations and collaborations throughout the years. I thank Dr. Tony Hui for teaching me how to measure total RNA and proteins in the cell, which lead me to the main discovery of my Ph.D. research. I would like to thank Dr. Xiaoyang Su, who generously trained me in mass spectrometry and also taught me how to do metabolic flux analysis.

I would like to thank the Botstein lab, notably Sandy Silverman and Dr. Scott McIsaac, for teaching me how to use a chemostat, a robust system that is an essential foundation of my graduate research. Scott’s efficiency of getting stuff done and his strict management of lab operation instilled in me the critical attitude necessary for lab work at the very beginning. I thank Professor David Botstein for letting me work in his lab. His vast knowledge of science and deep commitment to pedagogy are inspiring. He trains newcomers in the lab by making them measure microbial growth with different methods (spectrometer, serial dilutions, etc.). This simple yet effective way of teaching has become my standard protocol when I train someone who has never before done any bench work.

I thank Dr. Adam Rosenthal for exchanging ideas with me and respecting me as a colleague even when I was a junior graduate student. He is always open to sharing his excitement about
science and also gave me a mold for the mother machine he made. I thank Professor Gene-Wei Li and his lab for teaching me how to do ribosome profiling, an indispensable technique that has indeed enabled me to link metabolism and translation.

I would like to thank Professor Martin Wühr and his lab for teaching me quantitative proteomics. We have a great collaboration on quantifying the central dogma in *E. coli* and I am excited to see the answers we will find. I would like to thank Professor Tom Silhavy for his teaching. His Proks class gave me a solid knowledge of bacteria, which kickstarted my study of microbiology. While not on my committee, he is generous about giving me feedback and has a broad and deep knowledge of *E. coli*. I would like to thank Professor Jared Toettcher for his advice on how to survive graduate school and find the next exciting thing to work on as a postdoc.

I would like to thank Dr. Sean Hackett and Dr. David Robinson, two former Ph.D. students in the QCB program. They introduced statistics and programming in R to me, and that has changed me from a qualitative biologist into a quantitative computational biologist. David’s passion for teaching people how to interact with data in R has influenced me a lot, and I too have tried to help people learn R.

I would like to thank Dr. Derek Lidow. I was lucky to work as a teaching assistant with Derek for his excellent class, “Entrepreneur Leadership”. His class gave me a framework and understanding of leadership and teamwork, through which I have learned to better understand myself and to work better with other people, both in science and in life. Derek has also become one of my most important mentors at Princeton. I always walk away from our coffee meetings with new ideas and energy.

I would like to thank my advisors Professors Zemer Gitai and Ned S. Wingreen. While my research has moved in a different direction and involves distinct approaches from the rest of the lab, Zemer always supports my experimental proposals and provides feedback on questions I would like to answer. The trust and the freedom he gave me were critical to my growth as an independent scientist. I would like to thank Ned for his support and teaching. His class, taught with David Botstein, was the best class I have ever taken. His mastery of the physical perspective on biology gives him great mastery in explaining daunting equations and numbers in the context of fascinating biological phenomena. His patience and support allay my doubts and fears in the process of biological discovery. My first significant finding wouldn’t have yielded any results, had Ned not encouraged me to look deeper at ribosomes under different nutrient limitations.

I would like to thank Dr. Zhiyuan Li (李志遠). She can magically turn my experimental observations into rigorous equations. As a friend and a colleague, she has given me great support and advice on navigating through one’s career and generously taught me how to build mathematical models. Moreover, I was fortunate to get to watch her become a loving mom to her newborn, Jiujiu. It reminds me that our lives are enriched by experiences beyond the lab.

I would like to thank Dr. Raphael J. Morscher. His deep focus and passion for science are unique and special. I am honored and pleased to have had the opportunity to work with him on deciphering mitochondrial ribosome translation. His keen sense of commitment is contagious, and by working with him side-by-side for countless hours, his attitude has transformed me. My main breakthrough in understanding bacterial translation came after we broke down the mitochondrial system at an unprecedented resolution. While they were once free-living bacteria, mitochondria gave up the freedom for an endosymbiont lifestyle, and thereby generated the diversity and complexity of the eukaryotic world. The Mitochondrion is my favorite organelle because of its fascinating biology and relevance to metabolism and physiology. It also teaches us how the synergy of a team is greater than work of individuals.

I would like to thank Christopher King and Christopher Russo, the two bright and energetic undergraduate students I was fortunate to work with. Behind the data we generated together are priceless memories: working until late at night in the chemostat room, betting on the outcome of an experiment, and discussing exciting topics from modern wars to ancient Greek poems like Odyssey. They taught me as much as I taught them. Chris King has been a critical force in carrying out the systematic characterization of *E. coli* physiology in chemostat cultures. His meticulous and detailed experimental skills have generated high-quality data that will continue to yield new insights into bacterial physiology. I am glad that our adventure in science has encouraged him to
pursue a Ph.D. at Yale so he can continue to discover cool things that bacteria do.

I would like to thank Professors Rob Phillips, Wallace Marshall, Jennifer Lippincott-Schwartz, and all the friends I met during the Physiology course at Woods Hole in the summer of 2016. The course reminded me of the fun and enjoyment of science at a time when I felt the directionless in graduate school. Rob’s passion for teaching, science, and life is inspiring. I will always remember swimming to the Dock of Truth and the late night adventure to see the glowing sea. Wallace’s originality and crazy ideas freed my imagination about cell biology and provoked Maya Anjur-Dietrich and me to find a novel way to label ph ysarum. JLS is simply my role model: a great scientist, a visionary leader, and a strong woman.

I would like to thank the Massachusetts Institute of Technology (MIT), where I completed my undergraduate education in chemical engineering and biology, for giving me an opportunity to study there as an international transfer student. The generous offer of financial aid during the poor economic times of 2008 made my dream of studying science at one of the world’s top institutions come true. My experience at MIT transformed my life and opened up doors beyond my expectations. I would like to thank Professor Robert A. Weinberg for accepting me to work in his lab as a sophomore. I would also like to thank Professor Jeff Gore for letting me work in his lab and for introducing me to quantitative approaches and thinking in biology.

I would like to thank Dr. Christina Scheel, my first mentor at MIT. She helped me build the habits and solid foundations of an experimental scientist, which I continue to benefit from today. Her patience and guidance have influenced me. I have passed down her teachings to the undergraduates and junior graduate students I am fortunate to work with, as my way to pay her efforts forward.

I would like to thank my doctor, Dr. Han-Min Tseng (曾漢民). I was forced to take a leave of absence during my second year of graduate school when I was diagnosed with a benign brain tumor. Dr. Tseng used his outstanding skills to perform a successful surgery, which allowed me to quickly recover from the disease and return to my life at graduate school. This medical episode reminded me how precious life is and how fragile it can be: I need to do my best at my work and let things go when they are outside of my control.

I would like to thank my friends, whose support and laughter carry me through the dark times in science and life. I would particularly like to thank Ai-Lei Sun (孫愛蕾), my roommate for five years, and also a high school friend. She has played several roles in my life: counselor, style advisor, and car driver. I would like to thank my oldest friend, Vicky Chu (朱珮云), who I have known since we were ten. Despite vast differences in our life trajectories, she has visited me several times since I moved to the U.S. and given me the blunt advice I needed the most. I would like to thank my high school friend, Dr. Grace Liao (廖櫻美), who shares with me a passion for biology and now uses her talents in pediatrics. Our parallel careers in science and medicine have made us each other’s support for our aspirations to become professional women with great lives, including meaningful work lives and relationships. I would like to thank Ben Winer for super enjoyable conversations about science and life. Our shared go-getter personality and passion for science and entrepreneurship never invoke trivial conversation topics. I would like to thank Dr. Yogesh Goyal for his candid suggestions. I would like to thank Dr. Xin Jin for her ten-year-long sisterhood of support and guidance. I would like to thank Dr. Mazell Winikor for her company in doctors’ offices and support when I was taking a medical leave of absence. I would like to thank Dr. Daniela Garcia for being my boxing workout partner. I would like to thank Hilary Monaco, Dr. Christine Chin and Dr. Jihye Kim for giving me a boost of energy whenever I visit them in New York. I would like to thank Bianca Dumitrascu, Dr. Patty Ho (何率菩), Dr. Jen-Tang Lu (呂仁棠), Dr. Cameron Chen (陳柏旦), Ting-An Lin (林庭安), Dr. Shu-Heng Shao (邵書禎), Dr. Amy Wu (吳映女), and friends from and outside the Princeton Association of Taiwanese Students (普林斯頓台灣同學會). I would like to thank Dr. Mircea Davidescu for his unreserved and warm support during our time together at Princeton.

I would like to thank my middle school teacher, Dr. Jaw-Jing Jong (鍾兆晉), for kindling my initial passion for science. I would like to thank my teachers, Yu-yah Wu (吳育雅), Ling-Chi Hu (胡苓芝), Yu-Chen Sun (孫譽真), Ying-Zi Lin (林英子), Professor Guo-zen Yeh (葉國禎), Professor Yi-Chun Wu (吳益群), and Professor Da-Wei Chao (趙大衛), for teaching me science and supporting
me in pursuing opportunities. Throughout the years in high school and college in Taiwan, these passionate teachers patiently helped me grow the capability to work at the boundary of human knowledge. Professor Chao was the director of the Taiwanese International Biology Olympiad program which promotes biology education in high school. This program has generated a generation of bright young people who are exploring academics and biology around the globe. The Biology Olympiad program was the major turning point of my life and influenced my decision to pursue science and study abroad as early as possible.

I would like to thank my lovely family, particularly my mother, Chiao-Mei Chang (張喬湄). She has worked very hard to provide me with the best education possible. She is strict and never let me give things up half-way. In the meantime, she also allowed me to freely explore my interests with trust and supports me in any way she can provide. Through her example, I learned how to be dedicated to my responsibilities, including for my work and for the people around me. She taught me tremendously both with her actions and with her words. I would like to thank my stepfather, Roger Ou-Yang (歐陽啟源). He is an avid listener to whatever concerns I bring to him. Because of his vast repertoire of wisdom and experiences during his career in the semiconductor industry, he is always able to put me back on track when things fall apart. I would like to thank my father, Jack Li (李良傑), who since my childhood inspired me to think not only with my head but also with my hands. He is very creative at turning scrap materials into something useful in daily life. I would like to thank my sister, Ivy Li (李柏萱). Her feedback is always insightful with the wisdom she learns from the Bible. I am so happy to see her starting a happy and fulfilling marriage with her husband and my new brother-in-law, Kurihara Yutaka (栗原悠). Their complete dedication to God stands as a reminder for me to never lose sight of the pursuit of a meaningful life.
To my mother Chiao-Mei Joy Chang (張喬湄),
who raised me to be a kind, independent and courageous person.

獻給我最親愛的媽媽，張喬湄。
感謝妳從小到大對我的用心栽培，因為妳的信任和支持才能成就今天的我。
Protein synthesis is a fundamental activity all cells perform to sustain their functions and growth. Given the process requires energy such as ATP and the building blocks such as amino acids, protein synthesis activity is coupled to the metabolic state of the cell. Cellular metabolism includes uptake of external nutrients to break it down to common precursors (catabolism), and production of metabolites used by the cell for biomass accumulation (anabolism). This is achieved by a network of enzymes catalyzing conversions of metabolites between different pools as metabolic fluxes. Despite the fluctuation of the external nutrient levels and compositions, cells are capable of producing the desired biomass (including proteins) that is the proper amount in precise stoichiometry (Table 1)[163]. Understanding what metabolic limitations are and how they affect protein synthesis under different conditions will shed light on the regulatory mechanism linking the two basic cellular functions.

Protein synthesis can be viewed as the information flows from DNA to RNA (through transcription) and then to protein (through translation) in the central dogma. The metabolic state of the cell
Table 1: Percentage of total dry weight of different macromolecules in the cell

<table>
<thead>
<tr>
<th>Species</th>
<th>Proteins</th>
<th>RNA</th>
<th>DNA</th>
<th>Carbohydrates</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em></td>
<td>55</td>
<td>20.5</td>
<td>3.1</td>
<td>8.4</td>
</tr>
<tr>
<td>Budding yeast</td>
<td>45</td>
<td>6.3</td>
<td>0.4</td>
<td>40</td>
</tr>
<tr>
<td>Mammalian cells</td>
<td>70</td>
<td>1.4</td>
<td>5.9</td>
<td>5.3</td>
</tr>
</tbody>
</table>

can thus affect this process at different layers of regulation. The most widely studied layer is transcription as historically measuring RNA is technically easier than measuring proteins and previous studies have suggested RNA level is correlated to the protein level [16, 32, 137]. RNA is the second most abundant biomass in the cell (Table 1) in active growing cells and among total RNA > 95 % is stable RNA that is dedicated to translation machineries (80 % rRNA and 15 % tRNA) while ~5 % is mRNA [8, 17, 128]. The metabolic regulation on mRNA transcription is beyond the scope of this thesis but it is an interesting topic that is covered by reviews [142, 143]. On the other hand, rRNA transcription regulation has been intensively studied in *E. coli*. Ribosomes are the key cellular machineries that carry out translation. It is composed of two thirds in rRNA and one third in ribosomal protein subunits by mass fraction [107]. Given the high mass fraction of nucleotides in ribosomes, the transcription of rRNA is the bottleneck of ribosome production and is found to be a highly regulated process [114]. Two input signals are found to regulate rRNA transcription in *E. coli*: ppGpp and initiating NTP levels [19]. Each of them senses a different part of the cellular metabolic state and collectively determines the level of transcription from rRNA promoters. ppGpp, known as the magic spot, is a cellular alarmone of which the level elevates when growth rate is slowed or amino acids are deprived [118]. High level of ppGpp will reduce the activity of rRNA promoters by interfering the binding of RNA polymerases, reducing the ribosome production when nutrients are not abundant and growth is slowed [118, 119]. In addition, the transcription initiation of rRNA is highly sensitive to the cellular ATP/GTP concentrations. Low ATP/GTP reflects energy depletion of the cell and leads to down-regulation of ribosome production as well [29, 49]. Together, these studies have shown how the metabolic state of the cell affects transcription and thus ribosome abundance. It remains to be explored how cells might regulate ribosome activities in addition to its abundance.

In addition to transcription, translation carried out by ribosomes can integrate different metabolic inputs to determine the ribosome activity. Given that ribosomes are one of the largest complex in the cell and consumes a major fraction of cellular energy for translation, both of the abundance...
and the activity need to be regulated \([107]\). Protein synthesis is one of the most energy-consuming activities in the cell, taking up \(>50\%\) of cellular energy in a fast growing bacteria \([84]\). A simple scenario can illustrate why the coupling between translation and metabolic state is critical for cells. In a fast-growing \textit{E. coli} cell, there are approximately 20,000 ribosomes with elongation rate at 15 amino acids per second. Translating one amino acid costs three ATP equivalent (1 ATP for tRNA charging, 2 GTP for decoding and translocation of the ribosome). The ATP consumption per second will be \(20000 \times 15 \times 3 = 9 \times 10^5\). The size of cellular ATP pool is about 10 mM \([12]\) and cell volume is \(= 1 \mu m^3\) \([80]\), so the total ATP numbers inside the cell is \(6 \times 10^6\). As soon as ATP production is halted, translation can drain the total cellular ATP pool in seconds \([56]\). Recently, a breakthrough in technology development has made probing the dynamic of translation possible. Ribosome profiling uses the high sensitivity of next-generation sequencing technology to measure the mRNA fragments that are protected by ribosomes \([68]\). Collectively, all the fragments present a global snapshot of the ribosome activity across the genome at single nucleotide resolution. There is growing evidence that translation dynamics reveals the metabolic bottleneck \([87]\), affects protein folding \([111, 168, 175]\), and underlies the translational efficiency of mRNA \([162]\). In summary, ribosome profiling bridges the technology gap between static measurements of mRNA abundance (RNA-Seq) and protein abundance (Proteomics) and can yield new insights into post-transcriptional gene expression regulation.

A way translation can be affected by the metabolite pool sizes is through tRNA charging. Being the second most abundant RNA and the adaptor molecule between the nucleotide codes and the amino acids, tRNA is estimated to be 80 \% charged during exponential growth in bacteria \([35]\). However, when the amino acid pool becomes limited, tRNA charge ratio can drop, resulting in non-charged tRNA in the A-site of ribosome to slow translation \([150, 151]\). tRNA also contains many unique chemical modifications that are important for structural stability or facilitating decoding \([2, 58, 89, 153]\). While many studies have focused on characterizing the landscape of tRNA modification, it is not yet clear which pathways and precursors are involved and its physiological relevance.

Given all living organisms are constantly consuming energy to maintain their ordered structures, metabolic network fundamentally support all functions of life. In addition to regulating protein synthesis at specific steps, metabolism can have indirect influence through affecting the global physiology of the cell. At the organismal level, metabolic activities exhibit a robust temporal oscillation according to the circadian rhythm \([9]\). Such oscillatory metabolic activity is accompanied by time-dependent change in gene expression. An ultradian cycle in glycolytic metabolism and
respiration in the budding yeast Saccharomyces cerevisiae was found to change gene expression over half of the yeast genes [158]. Tu et al. hypothesized that the metabolic cycle is a way to compartmentalize biosynthetic activities such as translation from oxidative processes that might be harmful to the newly synthesized products [158, 160]. While they suggested metabolic cycle a single-cell and a common phenomenon beyond yeast, this behavior has not yet been described in other systems [159]. The universality of metabolic cycle and the link between cyclic metabolic activity and gene expression remains to be elucidated.

*E. coli* is an ideal model organism to study regulation of protein synthesis by metabolic limitations. Arguably the most well-known organism on earth, *E. coli* contains a 4.6 x 10^6 bp genome that has been sequenced and largely annotated [15]. *E. coli* can adapt to a wide range of growth rate and nutrient sources. Pioneering studies have investigated how ribosome level of the cell change as a function of growth rate by varying carbon sources [17, 135, 138]. More recently, researchers used tunable gene expression system to understand the allocation of proteome sectors of different functions under different nutrient uptake limitations [64]. Distinct and physiologically-relevant growth conditions are required to probe the adaptation to metabolic limitations in *E. coli*. Continuous culture system such as chemostat is suitable for the study because it provides steady states under different nutrient limitations at a wide range of growth rates [16, 20, 60]. Compared to the batch culture where nutrient quality modulates growth rate and consumption over time slows growth, chemostat controls and maintains growth rate by fixing nutrient availability at steady level in the culture [108, 109] (Fig. 1). Therefore, chemostat system avoids unintended effect from the nature of nutrient sources and can reveal the relationship between nutrient concentration and growth phenotypes.
This thesis aims to investigate how protein synthesis activity is affected by different metabolic limitations in two distinct model systems. The first chapter focuses on the *E. coli* physiology at balanced growth across different nutrient limitations. How does the cell balance making ribosomes, the complex that makes proteins, and proteins themselves, across different growth conditions? Our work systematically describes how different aspects of ribosomes are tuned to meet the protein synthesis demand. The second chapter looks at mitochondrial translation. We have uncovered a novel link between mitochondrial one-carbon metabolism and tRNA modification which underlies the oxidative phosphorylation (OXPHOS) protein deficiency observed in mutations in the lab and in patients of mitochondrial defects in clinics. Finally, the third chapter describes a serendipitous observation of an emergent behavior resembling the metabolic cycle in *E. coli* under glucose-limiting condition. I have provided another example how protein synthesis and metabolism might be coupled in a temporal manner.
E. coli translation strategies differ across carbon, nitrogen, and phosphorus limitation conditions

1.1 Abstract

For cells to grow faster they must increase their protein production rate. Microorganisms have traditionally been thought to accomplish this increase by producing more ribosomes to enhance protein synthesis capacity, leading to the linear relationship between ribosome level and growth rate observed under most growth conditions previously examined. Past studies have suggested that this linear relationship represents an optimal resource allocation strategy for each growth rate, independent of any specific nutrient state. Here we investigate protein production strategies in continuous cultures limited for carbon, nitrogen, and phosphate, which differentially impact
substrate supply for protein versus nucleic acid metabolism. Unexpectedly, we find that at slow growth rates, *E. coli* achieves the same protein production rate using three different strategies under the three different nutrient limitations. Upon phosphate (P) limitation, translation is slow due to a particularly low abundance of ribosomes, which are RNA-rich and thus particularly costly for phosphorous-limited cells. In nitrogen (N) limitation, translation elongation is slowed by processes including ribosome stalling at glutamine codons. In carbon (C) limitation, translation is slowed by accumulation of inactive ribosomes not bound to mRNA. These extra ribosomes enable rapid growth acceleration upon nutrient upshift. Thus, bacteria tune ribosome usage across different limiting nutrients to enable balanced nutrient-limited growth while also preparing for future nutrient upshifts.

1.2 Introduction

Resource allocation during growth is a fundamental challenge faced by all cells [54, 84, 139, 174]. For example, with a fixed resource budget, cells must balance production of the machinery that makes proteins (ribosomes, tRNAs, translation factors) with the production of the proteins themselves. This balance is generally represented by the RNA:protein ratio (R/P ratio) [34, 104]. The R/P ratio captures protein production capacity, as >95% of total RNA is devoted to translation (rRNAs and tRNAs [8, 34]). In single-celled organisms like *E. coli*, previous studies demonstrated that there is a linear relationship between R/P ratio and growth rate, with faster growth rates requiring more protein production capacity and therefore higher R/P ratios [34, 104, 135, 138]. Production of ribosomes is costly as each contains 52 protein subunits and three large rRNAs [107, 170]; hence, it is advantageous for the cell to saturate ribosomes with substrates. In this efficient ribosome scenario, the ribosome level should be fixed and independent of nutrient conditions for any growth rate, with the only way to increase protein synthesis rate being to increase the number of ribosomes [34, 174]. One surprise for such a seemingly optimized system is that multiple studies have demonstrated that at slow growth rates *E. coli* accumulates inactive ribosomes [30]. There are two possible explanations for the presence of inactive ribosomes. First, it is possible that *E. coli* translation is constrained in such a way that it cannot function when ribosome levels drop too low [30]. Alternatively, *E. coli* could regulate ribosome production independently of growth rate. Here we settle this debate by showing that *E. coli* ribosome production and usage differ across nutrient conditions.
1.3 Results

1.3.1 Phosphate-limited cells achieve the same growth rate with fewer ribosomes than Carbon- or Nitrogen-limited cells

To determine the generality of the relationship between growth rate and ribosome content, we examined how the R/P ratio changes as a function of growth upon different nutrient limitations. We measured R/P ratios in \textit{E. coli} under glucose- (C, carbon), ammonia- (N, nitrogen), and phosphate- (P, phosphorus) limitations over a range of different growth rates in chemostats (Fig. 1.1a). Surprisingly, P-limited cells consistently exhibited lower R/P ratios than C-limited or N-limited cells, with a roughly 2-fold difference at the slowest growth rate tested (0.1 h\(^{-1}\), Fig. 1.1b). Whereas most bacterial physiology studies have been performed with batch cultures, chemostat cultures reach steady state due to limitation of a specific nutrient. To determine if our findings are specific to chemostat-grown cultures we also determined the R/P ratios for batch cultures with different growth rates. These results confirmed previous findings that the R/P ratios of C- and N-limited cells follow the same trend regardless of whether they are grown in batch or chemostat conditions (Fig. 1.2a,b and Table S1 in Appendix A). Measured protein levels were similar in all cells regardless of growth rate or nutrient limitation (Fig. 1.2c,d). Moreover, the rRNA fraction of total RNA remained similar across different nutrient limitations at the same growth rate and decreased as growth was slowed (Fig. 1.2e,f). Since the nutrient-specific changes in R/P ratio cannot be explained by changes in rRNA fraction, and the P-limited cells with the lowest R/P ratios also have the lowest rRNA fraction, our data collectively suggest that P-limited cells produce protein at the same rate as C/N-limited cells using fewer ribosomes.

The finding that P-limited cells make the same amount of protein with fewer ribosomes, suggests that C/N-limited \textit{E. coli} cells do not use ribosomes with optimal efficiency and their “extra” ribosomes do not reflect a biophysical limitation. Since RNA accounts for two-thirds of the mass of bacterial ribosomes \cite{170}, producing fewer ribosomes upon phosphate-limitation makes sense as a way for cells to deal with a limitation that preferentially reduces an elemental substrate needed to make RNA but not protein. This lower ribosome level may be a direct consequence of low phosphate resulting in limited nucleotide pools, as deletion of genes involved in phosphate sensing or storage, \textit{phoB} \cite{167} or \textit{ppk} \cite{3}, did not alter the R/P ratio (Fig. 1.2g,h).
Figure 1.1: RNA-to-protein ratio is both growth-rate and nutrient dependent. (a) Schematic flow of nutrients for biomass formation. Carbon (C) and nitrogen (N) combine to make amino acids. Amino acids combine with carbon precursors and phosphorus (P) to make nucleic acids. (b) RNA-to-protein ratios for chemostat cultures under C-, N-, and P-limitations at different growth rates. Each data point shows the mean value from three technical replicates. (c) Translation elongation rates (amino acids/s) as measured by the lacZ induction assay after correction for translation initiation. The bar height represents mean values with error bars as SEM from three biological replicates. (d) Polysome profiles of cells grown under C-, N-, and P-limitations in chemostats at dilution rate 0.1 h⁻¹. Five independent experiments were repeated, with a single representative curve shown. (e) Quantification of ribosomes in the form of subunits (30S + 50S), free 70S, mRNA-bound 70S (one ribosome on one mRNA), and polysomes (multiple ribosomes on one mRNA). The bar height represents mean values with error bars as SEM from three biological replicates. (f) Cells adapt to different nutrient limitations using different strategies of translational regulation that achieve the same protein production rate.
Figure 1.2: Lower RNA-to-protein ratio under P-limitation results from lower RNA concentration, and phosphorus metabolism-related genes are not involved in the regulation. (a) RNA-to-protein ratio in chemostat and batch conditions. Each data point shows the mean value from three technical replicates. (b) Comparison of RNA-to-protein ratio values between this study and previously reported values from the same strain of *E. coli* NCM3722. Same conditions are connected by the dotted line. (c-d) Total protein (µg/mL/OD600) (c) and RNA (µg/mL/OD600) (d) concentrations were measured. These data were used to generate Fig. 1.1b of the RNA-to-protein ratio. Each data point represents the mean value from three technical replicates. (e-f) Fraction of rRNA (e) and tRNA (f) in total RNA determined from bioanalyzer assay across different nutrient limitations and growth rates. Each data point shows the mean value from three biological replicates with error bars as SEM. (g-h) RNA-to-protein ratio for chemostat cultures of wild-type and mutant cells upon N- and P-limitations at different growth rates. Each data point shows the mean value from three technical replicates: (g) ΔphoB, (h) Δppk.
1.3.2 N-limited ribosomes translate slowly while C-limited cells accumulate more mRNA-free ribosomes

Why do C/N-limited cells accumulate so many ribosomes if P-limited cells can achieve the same protein synthesis rates with fewer ribosomes? One possibility is that the ribosomes in these cells translate slowly. We thus used a lacZ induction assay to compare the translation elongation rates of slow-growing C-, N-, and P-limited cells (0.1 h^{-1}) [31, 136]. We observed a reduced elongation rate in N-limited cells compared to C- and P-limited cells but no difference between C- and P-limited cells (Fig. 1.1c and 1.3). Thus, N-limited cells may need higher ribosome numbers to compensate for their slow translation elongation, but something else must explain the elevated ribosome numbers in C-limited cells.

To characterize ribosome pools we performed polysome profiling, which separates ribosome species using a sucrose gradient [121]. Regardless of the growth condition, all cells exhibited similar fractional pools of dissociated 30S and 50S subunits (Fig. 1.1d). In contrast, the fraction of 70S monosomes was significantly larger in C/N-limited cells than in P-limited cells (Fig. 1.1d). Since growth rate is proportional to protein synthesis rate, growth rate can be estimated by the product of the number of active ribosomes and the translation elongation rate. However, using the assumption that all 70S monosomes are active yielded very different growth rate estimates for C-, N-, and P-limited cells (Table 1.1 and Appendix A), which is inconsistent with the fact that these cells are growing at the same rate and have the same protein content. This inconsistency

Figure 1.3: lacZ induction curves and fittings for translational elongation rate measurements. (a) Raw data of lacZ induction time course for batch glucose minimal media (Min-0.9), C-, N-, and P-limited cells grown at 0.1 h^{-1}. Three biological replicates were performed and data from one representative replicate is shown for clarity. (b) Transformed lacZ induction data for lag time estimation. All three biological replicates are shown with estimated lag time indicated in the bracket (See Methods for detail).
suggested that a fraction of the 70S ribosomes may not be active.

The mass of a single mRNA is small relative to the mass of a ribosome, such that 70S mono-

somes could represent either mRNAs with only one ribosome per transcript or inactive “free” ribosomes

that are not associated with an mRNA. To distinguish free and mRNA-bound 70S monosomes, we

utilized their differential sensitivity to high potassium levels (170 mM) [11]. High potassium causes

free ribosomes to shift to a lower density but does not shift the density of mRNA-bound mono-

somes [11]. We thus designed a high-resolution “free-ribosome profiling” method to resolve this
density shift (Fig. 1.4). As controls, we confirmed that our assay detects the potassium-dependent

shift of 70S monosomes induced by puromycin (Fig. 1.4a), which releases elongating ribosomes

from their associated mRNAs[103]. There was no potassium-dependent shift detected in fast-
growing cells that lack free ribosomes (Fig. 1.4a). As a second validation of our free-ribosome

profiling method we confirmed that free ribosome pools decreased as growth rate increased across

all three nutrient limiting conditions (Fig. 1.4b). By combining traditional and free-ribosome polysome

profiling, we quantified the relative fractions of all ribosome species in slow-growing C-, N-, and P-

limited E. coli. The fraction of free monosomes was roughly 3-fold greater in the C-limited cells

than in the P-limited cells, while the fraction of mRNA-bound monosomes remained relatively con-

stant across nutrient limitations (Fig. 1.1e). The accumulation of free 70S monosomes in C- and

N-limited cells appears to be independent of a previously-described RaiA-dependent mechanism

for ribosome storage as deletion of raiA had no impact on R/P ratios or polysome profiles (Fig.

1.5). Importantly, revising our protein synthesis rate estimates to account for the fraction of inac-
tive 70S monosomes correctly yielded similar values for all cells, regardless of nutrient limitation

(Table 1.1). These results both validate our experimental measurements and suggest that in dif-

ferent nutrient states, E. coli differentially tune ribosome number, translation rate, and active frac-
tion to produce proteins at the same rates (Fig. 1.1f).

Table 1.1: Growth rate estimation from parameters measured

<table>
<thead>
<tr>
<th>Sample</th>
<th>RPR</th>
<th>rRNA frac</th>
<th>_Rcomplex</th>
<th>_Rbound</th>
<th>kel (aa/s)</th>
<th>Estimated growth rate from all 70S (h-1)</th>
<th>Estimated growth rate from bound 70S (h-1)</th>
<th>Actual growth rate (h-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-0.1</td>
<td>0.16</td>
<td>0.64±0.03</td>
<td>0.83±0.02</td>
<td>0.31±0.03</td>
<td>12.4±0.2</td>
<td>0.29±0.02</td>
<td>0.09±0.01</td>
<td>0.09±0.00</td>
</tr>
<tr>
<td>N-0.1</td>
<td>0.16</td>
<td>0.59±0.06</td>
<td>0.84±0.02</td>
<td>0.51±0.03</td>
<td>8.1±0.1</td>
<td>0.17±0.01</td>
<td>0.09±0.01</td>
<td>0.09±0.00</td>
</tr>
<tr>
<td>P-0.1</td>
<td>0.08</td>
<td>0.51±0.03</td>
<td>0.81±0.03</td>
<td>0.65±0.04</td>
<td>12.4±0.1</td>
<td>0.12±0.01</td>
<td>0.08±0.01</td>
<td>0.09±0.01</td>
</tr>
</tbody>
</table>

Estimation of growth rate using measurements from RNA-to-protein ratio (RPR), fraction of rRNA in total RNA (rRNA frac), elongation rates (kel), and fractions of different ribosomal species (Φ_Rcomplex and Φ_Rbound). All values are mean values. RPR values are derived from three technical replicates and the rest of the parameters are from three biological replicates with the SEM shown. The propagated error is shown for growth rate estimates.
Figure 1.4: Elevated \([K^+]\) distinguishes mRNA-free from mRNA-bound 70S ribosomes. (a) The plot shows the separation of these two distinct 70S species under different growth conditions using 170 mM \(K^+\). The A254 values are vertically shifted for each condition for easier visualization. Puromycin is used as a positive control for free ribosome accumulation and fast-growing cells in glucose minimal media (Min-0.9) are used as a negative control for cells expected to lack free ribosomes. At least two independent experiments were performed with similar results. (b) Fraction of assembled ribosomes under different growth conditions. Chemostat growth conditions include C-, N-, and P-limitations grown at 0.1 and 0.6 h\(^{-1}\). Batch conditions include glucose minimal media (Min-0.9) and defined rich media (Rich-1.7). Mean values from three biological replicates are shown with error bars indicating the standard error of the mean.

Figure 1.5: Deletion of \(raiA\) does not alter ribosome dynamics. (a) The ratio of RNA-to-protein ratio of wild type and \(\Delta raiA\) at growth rate 0.1 h\(^{-1}\) under different nutrient limitations. Each bar represents the ratio of the mean R/P ratio from three technical replicates of wild type and the mutant. (b) Polysome profiles of wild type and \(\Delta raiA\) cells. Same amount of RNA from WT and \(\Delta raiA\) was loaded for comparison. Deletion of \(raiA\) has no effect on the size of 70S peak. 100S peak from ribosome dimerization was not detected under any condition. The thin dashed line marks the average baseline. (n = 1 independent biological sample.)
1.3.3 Quantitative modeling describes three different strategies of ribosome dynamics to achieve the same protein production rate

To better understand nutrient-dependent ribosome dynamics, we probed translation in slow-growing C-, N-, and P-limited cells by ribosome profiling. Analysis of ribosome densities as a function of distance from the start and stop codons revealed higher ribosome occupancy near the start codon (Fig. 1.6a). Ribosome density thereafter was similar, with no decrease in ribosome density between the first and second halves of genes (Fig. 1.6a and 1.7a), suggesting that there is little to no aborted translation after the first few codons.

In order to gain more mechanistic insights into the differences in translation across nutrient limitations, we developed a macroscopic mathematical model of ribosome dynamics (Fig. 1.6b, Fig. 1.7b-e and details in Appendix A). This macroscopic model incorporates the ribosome profiling data and models ribosomes as three species: unbound ribosomes, initiating ribosomes, and working ribosomes.
Figure 1.7: Modeling ribosome dynamics on an mRNA with a microscopic model. (a) Comparison of ribosome density between first and second halves of genes. Scatter plot of summed RPM (reads per million) in the first and second halves of each gene. Blue line has a slope of one. Red line is the fitting result with the value of slope marked in the graph. (n = 1 independent biological sample.) (b) Illustration of the microscopic model for bound ribosome dynamics on mRNA, and the corresponding flux balance equation. (c) Dark bars are the normalized ribosome count in the first 50 and the last 50 codons from ribosome profiling studies. Red curves are the smoothed ribosome occupancies obtained by fitting the normalized ribosome count in the first 50 and the last 49 codons to two-term exponential functions, and copying the observed count number at the stop codon. (d-e) The position-dependent normalized step rate (s\(^{-1}\)) and fraction of aborted translation that best fit the smoothed ribosome occupancies in (c). Insets are magnifications for codon position 1–10.
Figure 1.8: Parameters of the macroscopic model relate translation activity and growth. (a) Comparison of predicted elongation rates from the macroscopic ribosome dynamics model with experimentally measured ones. The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements. (b-c) The relative proceeding rate \( (N_{aa}/k_p/k_r) \) and saturation parameter \( (R_t/(K_m+R_t)) \) under various conditions. The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements. (d-e) Model relationships between total number of ribosomes and growth rate, while the total number of mRNAs is held constant. Dots mark the estimated values of ribosomes under each condition, and dashed line in (e) marks the half-saturation number of ribosomes under Min condition. The inset in (e) shows the region around the open circle. (f) The relationship between elongation rate \( (k_{el}/N_{aa}) \), fraction of working ribosomes \( (\Phi_{Rw}) \), and total number of ribosomes \( (R_t) \) that leads to the same growth rate of 0.1 h\(^{-1}\). Filled circles indicate the values for C-, N-, and P-limited wild type cells, and open circles indicate the corresponding values for the \( \Delta \)relA mutant.

Analysis of our model indicated that the system can be characterized by two main dimensionless parameters (SI): the "relative proceeding rate", defined as the ratio between the rate of ribosomes proceeding from initiation to elongation \( (k_p, \text{s}^{-1}) \) and the rate of elongation \( (k_{el}/N_{aa}, \text{s}^{-1}) \); and the "saturation parameter" \( (R_t/(K_m+R_t)) \), reflecting the degree of saturation of ribosome bind-
ing sites on mRNAs, where \( R \) is the total ribosome number and \( K_m = (k_r + k_p)/k_f \) (Fig. 1.6c,d and Fig. 1.8b,c). Fitting the measured ribosome densities, translation elongation rates, and pool sizes of ribosome species to the macroscopic model revealed that P-limited cells have the highest relative proceeding rate while C-limited cells have the lowest saturation parameter, and N-limited cells have the highest saturation parameter (Fig. 1.6c,d and Fig. 1.8b-e). Thus, C-, N-, and P-limited cells produce proteins at the same rate using three different strategies: P-limited cells have few ribosomes that are mostly active and elongate rapidly, N-limited cells have more ribosomes but fewer are active and they elongate slowly, and C-limited cells have many ribosomes, which elongate rapidly, but even fewer are bound to mRNA (Fig. 1.6e and Fig. 1.8f).

1.3.4 N-limited ribosomal regulation is mediated by RelA, the ppGpp alarmone synthase

Our ribosome profiling data provides additional insight into the molecular basis of nutrient-specific ribosome regulation through analysis of codon occupancies. Codon-specific ribosome stalling leads to increased codon occupancy and is a hallmark of insufficient pools of the corresponding charged tRNAs. P-limited cells exhibited no elevated codon frequencies, consistent with these cells’ efficient ribosome usage (Fig. 1.9a). In contrast, both C- and N-limited cells exhibited significant codon-specific stalling. Under N-limitation, ribosomes stalled at both of the two glutamine-encoding codons. This result is consistent with previous studies indicating that glutamine is the most strongly-depleted amino acid pool upon N-limitation and serves as an intracellular sensor for extracellular nitrogen levels [38, 66, 176]. Since glutamine codons account for 4.4% of all predicted ORF codons in \( E. coli \), the ~2-fold increase in glutamine codon occupancy indicates that additional mechanisms also contribute to the ~30% reduction in translation elongation rate observed upon N-limitation. Upon C-limitation, we observed elevated occupancy of the Leu-CUA codon, which was surprising as there is no known intracellular carbon sensor that controls translation. \( E. coli \) has six leucine codons decoded by five leucine tRNA species [145, 150]. Leu-CUA is the rarest Leu codon, accounting for only 0.4% of all predicted ORF codons.

In addition to clarifying the link between metabolism and translation for N-limitation, the observation that ribosomes stall at specific codons upon C/N-limitation suggested a molecular mechanism for nutrient-specific translation regulation (Fig. 1.9a). In bacteria, insufficient charged tRNA pools activate the stringent response, which is mediated by the accumulation of the cellular alarmone, ppGpp [62, 118]. ppGpp is known to regulate rRNA transcription in vivo. Its role in translation is less well-understood but ppGpp has been shown in vitro to inhibit translation factors such as EF-Tu and IF2 by competing with GTP [97, 98]. To test how ppGpp accumulation might af-
Figure 1.9: Deletion of relA disrupts translation regulation under nitrogen limitation. (a) A-site codon occupancy under different growth conditions from ribosome profiling. Occupancy was calculated as the ratio between measured and expected counts for each gene based on codon frequency. The average of this ratio is plotted (with number of genes specified). Codons that have higher than average ratio by 2.5 standard deviations are highlighted. (b) Fraction of assembled (70S) ribosomes in wild type and ΔrelA under C-, N-, and P-limitations at growth rate 0.1 h⁻¹. The bar height represents mean values with error bars as SEM from three biological replicates. (c) lacZ induction assay for wild type and ΔrelA under nitrogen limitation at growth rate 0.1 h⁻¹. The lag time measures when the first functional LacZ is produced and is inversely proportional to the elongation rate. Mean values from three biological replicates are shown with error bars as SEM. (d) Ratio of codon occupancy between ΔrelA and wild type under nitrogen limitation at growth rate 0.1 h⁻¹ (n = 1 independent biological sample). (e) Ratio of the relative proceeding rate (N_{aa}k_p/k_{el}) between ΔrelA and wild type across different conditions at growth rate 0.1 h⁻¹. The bar height represents mean values with error bars as standard deviation from all the possible combination of experimental measurements.
fect translation in vivo, we induced ppGpp synthesis by treating batch-grown *E. coli* with serine hydroxamate (SHX). SHX is a serine analog that competitively inhibits serine tRNA synthetase to yield uncharged serine-tRNA and thereby activate the RelA ppGpp synthase [70]. SHX treatment increased the pool of mRNA-free 70S monosomes, and this effect was completely dependent on relA (Fig. 1.10a-b). Thus, inducing ppGpp by activating RelA alters translation by increasing the fraction of inactive free 70S ribosomes. RelA is primarily required for the accumulation of ppGpp upon N-limitation but not C-limitation [118]. Consistently, deletion of relA had little impact on the accumulation of free ribosomes upon C-limitation, but significantly reduced free ribosome pools upon N-limitation to levels similar to those observed upon P-limitation (Fig. 1.9b and Fig. 1.10c). We could not probe the role of ppGpp in C-limited ribosome accumulation because unlike N-limitation, C-limitation elevates ppGpp through spoT, which is essential [52]. N-limited ΔrelA cells also contained more polysomes than wild type (Fig. 1.9b), suggesting that these cells had a higher fraction of elongating ribosomes. This result was initially confusing because the wild type and ΔrelA N-limited cells were grown at the same growth rate and maintained the same R/P ratio (Fig. 1.10d). We thus measured the rate of translation elongation and found that while N-limited ΔrelA cells have higher fractions of translating ribosomes, their ribosomes elongate more slowly (Fig. 1.9c), resulting in the same rate of protein production as wild type.

To understand how RelA influences ribosome dynamics we performed ribosome profiling. N-limited ΔrelA cells exhibited even more pronounced ribosome stalling at both glutamine codons than wild type N-limited cells, while relA deletion had no effect on P- or C-limited cells (Fig. 1.9d and Fig. 1.11). Fitting the ΔrelA cells measurements to our ribosome dynamics model revealed that ΔrelA cells specifically increase the relative proceeding rate under N-limitation, but display no effect on relative proceeding rate under P- or C-limitation and no effect on the saturation parameter in any condition (Fig. 1.9e and Fig. 1.8b,c). These results suggest that in N-limited cells, RelA-dependent ppGpp serves to restrict ribosome function by regulating the translation initiation rate and/or the frequency at which ribosomes transition from initiation to elongation. Furthermore, we speculate that in the absence of RelA more ribosomes attempt to elongate, which exacerbates the depletion of charged tRNA pools, leading to increased stalling and a slower translational elongation rate.

### 1.3.5 Extra ribosomes may facilitate growth acceleration upon nutrient upshift

While cells can modulate different aspects of ribosome dynamics to achieve the same protein production rate, what benefits might be served by the inefficient translation system used by C-
Figure 1.10: relA-dependent ppGpp production leads to free ribosome accumulation under stringent response and nitrogen limitation. (a) Polysome profile of wild type and ΔrelA cells with and without treatment of serine hydroxamate (SHX) for ten minutes. Two independent repeats were performed and one representative data is shown. (b) Free-ribosome profiling using 170 mM KCl to distinguish mRNA-free and mRNA-bound ribosomes. (n = 1 independent biological sample.) (c) Polysome profile of wild type and ΔrelA cells at growth rate of 0.1 h⁻¹. Three independent repeats were performed and one representative data set is shown. (d) Ratios of R/P ratios of wild type and ΔrelA at growth rate of 0.1 h⁻¹. The bar shows the ratio of the mean R/P ratio from three technical replicates of wild type and the mutant.
Figure 1.11: Deletion of relA disrupts translation under nitrogen limitation. (a) Ratio of codon occupancy between ΔrelA and wild type cells under carbon limitation at 0.1 h⁻¹. (b) Ratio of codon occupancy between ΔrelA and wild type cells under phosphorus limitation at 0.1 h⁻¹. (c) Cumulative fraction of ribosome counts of fabI in wild type and ΔrelA under C-, N- and P-limitations at 0.1 h⁻¹. The vertical dotted lines and red triangles mark the positions of glutamine codons.
limited cells where many ribosomes are inactive? Since free ribosomes accumulate the most at the slowest growth rates, we hypothesized that our findings could reflect a trade-off between steady-state growth rate and the ability to respond to a fluctuating environment. Such transitions could include rapidly and safely slowing growth when nutrients become depleted and rapidly increasing growth rate when nutrients are replenished. In this scenario, cells may benefit more by optimizing their ability to rapidly utilize new nutrients, for example to outcompete their neighbors or maximally utilize a transient pulse of nutrients, than by optimizing steady-state growth rate when nutrient levels are low. We extended our mathematical model of ribosome dynamics to predict cellular growth dynamics upon nutrient upshift. This modeling supported the hypothesis that the larger free ribosome pools of C/N-limited cells should enable them to increase their growth rates more quickly than P-limited cells upon nutrient upshift (Fig. 1.12a, Fig. 1.13 and details in Appendix A)[54, 115]. We experimentally tested this prediction by measuring the growth rates of slow-growing (0.1 h⁻¹) E. coli immediately after being shifted to rich media (LB + 0.4% glucose) [75]. As predicted, C/N-limited cells increased their growth rates significantly faster than P-limited cells (Fig. 1.12b). Thus, the distinct translation strategies employed under different nutrient conditions may represent nutrient-specific adaptations, with P-limited cells optimizing for current steady-state growth under slow growth conditions, and C- and N-limited cells favoring the ability to rapidly recover growth (Fig. 1.12c).

1.4 Discussion

In previous studies, E. coli were found to vary R/P ratio with growth rate independently of the specific nutrient limitation used to produce a given growth rate [34, 138]. While pioneering physiology studies acknowledged the possibility that such growth laws could depend on the specific growth state of the cell [104, 138], the reproducibility of the correlation between R/P ratio and growth rate has often been considered to be more general [138, 174]. Meanwhile, transcriptional analysis in Saccharomyces cerevisiae suggested that the primary determinant of the response to a wide range of stresses was the cellular growth rate rather than the specific stressor [16]. Together, these studies suggested that the primary regulator of microbial physiology is growth rate. However, our findings demonstrate that at the same growth rate, E. coli exhibit significantly different translation strategies across nutrient limitations; at the lowest growth rate tested, P-limited cells produced the same amount of protein with roughly half as many ribosomes as C/N-limited cells. P-limited cells also exhibited smaller inactive 70S monosome pools and higher relative pro-
Figure 1.12: Extra ribosomes confer growth advantage upon nutrient upshift. (a) Theoretically-predicted growth curves from the macroscopic model. (b) Experimental nutrient-upshift growth curves in LB + 0.4% glucose for wild type cells from C-, N-, and P-limited chemostats at growth rate 0.1 h⁻¹. The initial OD and standard deviation of three biological replicates for each condition are 0.141 ± 0.008 (C), 0.107 ± 0.017 (N), and 0.154 ± 0.008 (P) respectively. Each data point shows the mean value from three biological replicates with error bars as standard deviation. (c) Model for nutrient-dependent ribosome usage: Under C- or N-limitation, the total ribosome pool is high while under P-limitation the total pool is low. C-limited cells elongate fast but have a low fraction of working ribosomes. N-limited cells elongate slower but have a higher fraction of working ribosomes than C-limited cells. Under P-limitation the low supply of ribosomes leads to both a high fraction of working ribosomes and fast elongation to meet the protein production demand.
Figure 1.13: Different allocation strategies determine growth dynamics upon modeled nutrient upshift to rich medium. After nutrient upshift to a rich condition, different regulatory strategies for $\Psi_R(t)$, the fraction of new protein synthesis allocated to ribosomal proteins (first three rows for C-, N-, and P-limitations, respectively), and their resulting post-upshift growth curves (last row).
ceeding rates than C/N-limited cells. Furthermore, while C/N-limited cells have similar ribosome numbers, they also display differences in free ribosome pools, translational elongation rates, and sensitivity to the loss of the ppGpp synthase RelA. Thus, our results suggest that the extra ribosomes of C/N-limited cells do not reflect essential constraints, but rather reflect a selectively beneficial adaptation. Our findings also implicate ppGpp as a key mediator of ribosome activity that specifically modulates the transition from translation initiation to elongation. ppGpp is a well-characterized cellular alarmone that senses translational activity and inhibits rRNA transcription [114, 118]. However, ppGpp can competitively inhibit GTP-dependent enzymes, including the translation initiation factor IF-2 that is required for the transition to elongation [83]. Since ppGpp can also be produced by SpoT and spoT relA mutants cannot survive in the conditions of our experiments [169], the effect of complete loss of ppGpp remains unclear. Nevertheless, the phenotypes we observe in the absence of RelA provide in vivo support that ppGpp can affect translation in addition to affecting ribosome production. Our modeling suggests ppGpp affects translation by altering the relative proceeding rate, which can be influenced by both translation initiation rate and the frequency with which ribosomes transition from translation initiation to elongation. One attractive candidate target for ppGpp is IF-2, as a previous in vitro study demonstrated that ppGpp inhibits IF-2 function [97] and IF-2 has been implicated in both initiation and the transition to elongation. Since ppGpp can also affect other GTP-binding proteins involved in translation [98, 125], future studies on bacterial translation will need to address both how RelA-dependent ppGpp interferes with ribosome function and whether other sources of ppGpp are important translational modulators in other contexts. In native environments such as the mammalian gut, bacteria such as *E. coli* are faced with feast and famine cycles induced by feeding cycles [177]. Here we show at the same growth rate that the R/P ratio in *E. coli* does not reflect the optimization of steady-state growth but, instead, may reflect the ability of cells with higher inactive ribosome pools to rapidly accelerate growth upon nutrient repletion. These findings suggest that *E. coli* may improve their fitness by sacrificing their maximal growth rate in nutrient-poor periods in return for the ability to respond to a changing environment, including rapidly accelerating growth in nutrient-rich periods. Our macroscopic ribosome dynamics model highlights how cells can tune total ribosome number, the fraction of working ribosomes, and the rate of translational elongation to achieve the same total protein production rate while balancing other constraints such as reduced amino acid availability or the need to rapidly accelerate growth. This strategy could help explain recent reports of the sub-optimality of protein allocation for *E. coli* in the presence of poor carbon sources [44, 99, 157], the sub-optimal expression levels of essential genes in *B. subtilis* [116], and excess
ribosome production in *S. cerevisiae* [72, 96]. Future studies will address the consequences of adaptation strategies in dynamic conditions, for example the generality of optimizing growth rate transitions at the expense of steady-state growth.

1.5 Materials and Methods

1.5.1 Cell strains and growth conditions

*Escherichia coli* strain NCM3722 was grown in batch or continuous cultures. To achieve different growth rates, different carbon or nitrogen sources were provided in batch culture, whereas dilution rates ranging from 0.1 h\(^{-1}\) to 0.7 h\(^{-1}\) were used in continuous (chemostat) cultures. The chemostat (Sixfors, HT) volume was 300mL with oxygen and pH probes to monitor the culture. pH was maintained at 7.2 ± 0.1 and the aeration rate was set at 4.5 l/h. 40 mM MOPS media (M2120, Teknova) was used with glucose (0.4%, Sigma G8270), ammonia (9.5 mM NH\(_4\)Cl, Sigma A9434) and phosphate (1.32 mM K\(_2\)HPO\(_4\), Sigma P3786) added separately. For carbon- and nitrogen-limiting media, glucose and ammonia concentrations were reduced by 5-fold (0.08% and 1.9mM respectively). Phosphorus-limiting medium contains 0.132 mM K\(_2\)HPO\(_4\). Mutants with gene deletions were generated by P1 transduction from the KEIO collection [6] into *Escherichia coli* strain NCM3722.

1.5.2 Nutrient upshift growth measurement

Cells from chemosats were mixed with 4X volumes of fresh pre-warmed media and grown in flasks in a 37°C water bath. Cell growth was monitored every 5 minutes by checking absorbance at 600nm in a quartz cuvette (Starna, 16.160-Q-10/Z8.5) using a spectrophotometer (GENESYS™ 20, Thermo Scientific). The LB media (244610, BD) used for upshift was supplemented with 0.4% glucose.

1.5.3 Total RNA measurement

The method for RNA measurement was adapted from You et al. [174]. 1.5 mL of cultures were pelleted by centrifugation for 1 min at 13,000 X g. The pellet was frozen on dry ice and the supernatant was taken to measure absorbance at 600 nm for cell loss. The pellet was then washed twice with 0.6M HClO\(_4\) and digested with 0.3M KOH for 1 hour at 37°C. The solution was then precipitated with 3M HClO\(_4\) and the supernatant was collected. The pellet was re-extracted again with 0.5M HClO\(_4\). The supernatants were combined and absorbance measured at 260nm using
NanoDrop (ND-1000, NanoDrop). Total RNA concentration was determined by multiplying the A260 absorbance with 31 (µg RNA/mL) as the extinction coefficient.

1.5.4 Total protein measurement

The method for protein measurement is adapted from You et al. [174]. 1.5 mL of cell cultures were pelleted by centrifugation for 1 min at 13,000 X g. Cells were washed with 1 mL MOPS buffer once, re-suspended in 200 µL water and left on dry ice. All the supernatant was collected and OD600 was measured for cell loss. To measure the protein content, samples were thawed, 100µL 3M NaOH were added, and the sample was heated for 5 min at 98°C. The samples were cooled down to RT for 5 min before 300µL 0.1% CuSO₄ were added for biuret assay. The samples were incubated at RT for 5 min and centrifuged at 13,000 X g for 1 min. The supernatant was collected and absorbance was measured at 555nm for a 200µL sample volume in a microplate reader (Synergy HT, BioTek) with software Gen 5.0. Proper dilution of albumin (23209, Thermo) with known concentrations was used to infer the total protein concentration in the cell.

1.5.5 Polysome profiling and quantification of ribosome fraction

200mL of cells were collected from cultures by filtration through 90mm cellulose acetate membranes with a 0.2 µm pore size (CA029025, Strelitech) at 37°C, scratched with a clean and pre-warmed stainless steel spatula, and snap-frozen in liquid nitrogen. Cell pellets were mixed with 650µL lysis buffer frozen nuggets (20mM Tris-HCl pH 8.0, 10mM MgCl₂, 100mM NH₄Cl, 0.4% Triton X-100, 0.1% NP-40, 1 mM Chloramphenicol, 100 U/mL RNase-free DNase I (04716728001 Roche)) in a pre-chilled 10mL jar (014620331, Retsch). Pulverization was done by cryomill (Retsch) at 15 Hz for 15 minutes. The thawed cell lysates were quantified by NanoDrop and 200µL of lysates with RNA concentration ranging from 80µg to 500µg were used. For overall polysome quantification, lysates were loaded to 10%-55% linear sucrose gradients (20mM Tris-HCl pH 8.0, 10mM MgCl₂, 100mM NH₄Cl, 300 µM Chloramphenicol) made by GradientMaster (BioComp). The gradients were placed in a SW41Ti bucket and centrifuged in Optima XE-100 Ultracentrifuge (Beckman Coulter) at 35,000 rpm for 2 hours at 4°C. Gradients were fractionated by BioComp Gradient Fractionator and the absorption curves at 254nm were recorded by a UV monitor (EM-1, BioRad). Quantification of the polysome profiles was done using customized MATLAB codes. First, baselines were estimated using the readings where no peaks existed and the background was subtracted. Each peak was picked and quantified by integrating the area underneath the curve. To quantify different species of ribosomes in the 70S peak, 100mM NH₄Cl was replaced with 170mM...
KCl. Cell lysates were loaded onto 10%-30% linear gradient and centrifuged at 35,000 rpm at 4°C for 5 h. Because the two peaks for the free and mRNA-bound 70S ribosomes are very close in mass and clean separation was not possible, the MATLAB file-exchange package, peakfit, was used to fit the two overlapped peaks as two Gaussian distributions. For the free ribosome control, cells were treated with 100µM puromycin for 5 minutes and collected. For serine hydroxamate (SHX, Sigma S4503) treatment, cells were grown in MOPS glucose minimal media until OD ~0.3 and treated with SHX with final concentration of 1 mg/mL for ten minutes before collection.

1.5.6 lacZ induction and translational elongation rate measurement

The method was adapted from Zhu et al[180]. A final concentration of 5 mM IPTG (I2481C-25, Gold Biotechnology) was added to cultures. At every 15 seconds, 1 mL of culture was collected in a tube containing 10µL 100mM chloramphenicol, snap frozen in liquid nitrogen and stored at -20°C before subsequent measurement. After cells were thawed, 400µL of the sample was added to 100 µL 5x Z-buffer (0.3M Na₂HPO₄·7H₂O, 0.2M NaH₂PO₄·H₂O, 50mM KCl, 5mM MgSO₄, 20mM β-mercaptoethanol) and incubated at 37°C for 10 minutes. 100µL of 4mg/mL MUG (337210010, ACROS Organics) in DMSO was then added to each sample every 10 seconds to accurately control the reaction time. The samples were incubated at 37°C in a thermomixer at 1,400 rpm mixing rate for 30 minutes to 2 hours, depending on the enzyme expression level. The reaction was stopped by addition of 300µL of 1M Na₂CO₃. The tubes were spun down at 16,000 X g for 3 minutes to sediment the cell debris. 200 µL of supernatant were taken and measured fluorescence by a microplate reader (365 nm excitation and 450 nm emission filter). To infer translational elongation rate, the square root of the signal in excess of the signal at time zero was plotted. A linear fit was performed on the points after signal began to increase. The lag time is the x-intercept of the line. Elongation rate was corrected with the lag time in lacZα assay reported previously[180] according to Eq. S3 in Appendix A.

1.5.7 Ribosome footprinting and total RNA extraction for RNA-Seq

The cell collection step was the same as for polysome profiling except that 1mM chloramphenicol was used in the sucrose solution. The footprinting and library preparation steps were adapted from Li et al.[84]. After quantification of RNA concentration with NanoDrop, samples with 500µg RNA were digested with 750U MNase (10107921001, Roche) for 1 hour at 25°C before being quenched with 6mM EGTA. The lysates were then layered onto a 10%-55% sucrose gradient and centrifuged. The monosome fraction was collected and snap frozen in liquid nitrogen. No
polysome peaks were observed, indicating a thorough digestion. The RNA was isolated using hot phenol and size selected on 15% TBE-Urea PAGE gels run for 1 hour at 210V. Gels were stained with SYBR Gold and visualized using Dark Reader (Clare Chemical Research). RNA fragments with size between 25-40 nt were extracted using isopropanol precipitation. Total RNA was extracted with TRIZOL from the same pulverized cells used for footprinting. After DNaseI (04716728001, Roche) treatment and cleanup using RNA clean & concentrator 5 (R1016, Zymo Research), ribosomal RNA was subtracted using MicrobeExpress (AM1905, Ambion). The recovered RNA was fragmented and size selected. RNA fragments with size between 15-40 nt were extracted using isopropanol precipitation.

1.5.8 Library preparation and sequencing

Fragments from footprints and total RNA were dephosphorylated at the 3’ end by PNK (M0201, NEB). The repaired fragments were linked to the Universal miRNA Cloning Linker (S1315S, NEB), reverse transcribed (18080044, Thermo) and circularized (CL4111K, Epicentre). The circularized samples were PCR amplified (M0531L, NEB) and size selected. High quality PCR samples checked by Bioanalyzer high sensitive DNA chip. Deep sequencing was done by Illumina HiSeq 2500 on Rapid flowcells with settings of single end and 75 nt-long read length.

1.5.9 Mapping and sequencing data analysis

Data manipulation including barcode splitting, linker trimming and mapping were done using Galaxy. The processed reads were mapped to *Escherichia coli* genome escherichia_coli_k12_nc_000913_3 from the NCBI database with the BWA short read mapping algorithm. Only the reads between 20-45 nt that aligned to the coding region were used for further analysis. To infer the ribosome A-site position, python package Plastid[42] was used to align the 3’ end of reads to the stop and start codons[168], which are known to have higher ribosome densities. We found that the offsets were 12 nt for stop codon and 15 nt for start codon. Therefore, we used 11nt for A site position and 14nt for P site. The counts were normalized to the total counts in the coding region as reads per million (RPM) and transcripts per million (TPM). Further analysis was done using customized Python and R codes with packages including Plastid, dplyr, tidyr and ggplot2.

1.5.10 Analysis of ribosome profiling data

Transcripts per million (TPM) was used to identify highly expressed genes for codon occupancy analysis. After assigning each mapped read to the A-site nucleotide, the raw counts were first
normalized by the length of their mapped coding region, to yield count density for each gene. These count densities were then globally normalized to one million counts across all genes within one sample. Transcripts having total counts over 120 TPM and containing more than 200 codons were considered. To determine codon occupancy, ribosome footprint counts for the first and last 40 codons were removed to avoid possible effects from initiation and termination, and counts per codon were recorded for the remaining counts. For each gene and each codon type, the codon occupancy ratio is defined as the ratio of measured counts to expected counts. Expected counts are simply proportional to the frequency of that particular codon in the gene. The final reported codon occupancy ratio is the average of codon occupancy ratios from all genes considered weighted equally. We calculated ribosome counts along a transcript using reads per million (RPM). After assigning each mapped read to the A-site nucleotide, only genes over 100 codons long were selected and the total counts were normalized to one million reads. The selected genes account for more than 80% of the total counts. The ribosome counts are the sum of RPM at each position. To do the first and second halves comparison of ribosome counts of genes, the same set of genes and trimming processing for codon occupancy were used. The RPM from positions were summed up based on its assigned location (first or second half) on the transcript and plot against each other.

2 Mitochondrial translation requires folate-dependent tRNA methylation

2.1 Abstract

Folates enable the activation and transfer of one-carbon units for biosynthesis of purines, thymidine and methionine[40, 47, 88, 155, 171]. Antifolates are important immunosuppressive[86, 112] and anticancer agents[23, 94]. In proliferating lymphocytes[90, 126] and across human cancers[73, 105], folate enzymes localizing to the mitochondria are particularly strongly upregulated. This in part reflects the need for mitochondria to generate one-carbon units and export them to the cytosol for anabolic metabolism[39, 155]. The full range of uses of folate-bound one-carbon units in the mitochondrial compartment itself, however, has not been thoroughly explored. Here we show that loss of catalytic activity of the mitochondrial folate enzyme serine hydroxymethyltransferase 2 (SHMT2), but not other folate enzymes, leads to defective oxidative phosphorylation due to im-
paired mitochondrial translation. We find that SHMT2, presumably by generating mitochondrial 5,10-methylenetetrahydrofolate, provides methyl donors for producing the taurinomethyluridine base at the wobble position of select mitochondrial tRNAs. Mitochondrial ribosome profiling reveals that SHMT2 knockout cells, due to lack of this modified base, suffer from defective translation with preferential mitochondrial ribosome stalling at certain lysine (AAG) and leucine (UUG) codons. This results in impaired expression of distinct subunits of respiratory chain complexes I, IV, and V. Stalling at these specific codons also occurs in certain mitochondrial inborn errors of metabolism, due either to mutations in enzymes catalyzing the modification reaction or in the associated leucine or lysine tRNAs. Disrupting whole-cell folate metabolism, by folate deficiency or antifolate therapy also impairs the respiratory chain. In summary, mammalian mitochondria use folate-bound one-carbon units to methylate tRNA, and this modification is required for respiratory chain translation and thus oxidative phosphorylation.

2.2 Introduction

The major source of folate one-carbon (1C) units in mammalian cells is the amino acid serine[40, 47, 88, 155, 171]. Transfer of serine’s 1C unit to tetrahydrofolate (THF) can occur in either the cytosol or mitochondrion, via SHMT1[51] or SHMT2[51, 149], respectively (Fig. 2.2a). Evidence from stable isotope tracing indicates that cancer cells predominantly use SHMT2 to catabolize serine, exporting the resulting 1C units to the cytosol to support nucleotide synthesis[39, 117, 155]. The extent to which 1C unit production via SHMT2 is also important to support mitochondrial health has yet to be determined.

2.3 Results and Discussion

In characterizing a set of human HCT116 colon cancer CRISPR deletion cell lines lacking folate 1C enzymes, we serendipitously discovered that loss of SHMT2 induces rapid change in media color indicating extracellular acidification (Fig. 2.1a). Quantitative analysis of media confirmed increased glucose uptake and lactate secretion. This effect was SHMT2 specific: loss of other core 1C enzymes, including other mitochondrial enzymes (MTHFD2 and MTHFD1L) and SHMT1, did not induce glycolysis (Fig. 2.2b, 2.1b).

A common cause of increased glycolytic flux is respiratory deficiency[55]. Loss of SHMT2 reduced both basal respiration and maximal respiratory capacity (Fig. 2.2c) and decreased the NAD+/NADH ratio in multiple HCT116 SHMT2 knockout clones, but not knockouts of other core
Figure 2.1: SHMT2 deletion-induced respiratory chain dysfunction in different cellular backgrounds and clones. (a) Change in media colour after 48 h cell growth. (b–c) Lactate secretion (b) and normalized NAD+/NADH ratio (c) of HCT116 knockout cell lines (n = 6). (d–e) Basal respiration as measured by Seahorse XF analyser (n = 3) (d) and normalized NAD+/NADH ratio (n = 3) (e) of HEK293T folate 1C gene CRISPR–Cas9 knockout cell lines. (f) Normalized levels of TCA cycle and associated metabolites (n = 3). (g) Steady-state labelling fraction into citrate from [U-\textsuperscript{13}C]substrates glutamine (left) and glucose (right) (n = 3). (h) Immunoblot of extracted mitochondria for subunits of respiratory chain complexes I–V (CI–CV) and markers of mitochondrial mass. (i) Mitochondrial complex I levels (NDUFS4) in independent HCT116 folate 1C gene knockout clones. Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.
Figure 2.2: Mitochondrial respiratory chain function is dependent on SHMT2 catalytic activity. (a) 1C pathway and known mitochondrial products. (b) Lactate secretion of HCT116-knockout cell lines (n = 6). ΔSHMT2-A and ΔSHMT2-B denote two separate SHMT2-knockout lines. WT, wild type. (c) Oxygen consumption rate measured by Seahorse XF analyser (n = 3). FCCP denotes a mitochondrial uncoupling agent. Oligom., oligomycin; Rot./antim., rotenone/antimycin. (d) Immunoblot for mitochondrial respiratory complex I and II (CI and CII) proteins (NDUFS4 and SDHA, respectively), 1C enzymes, and a marker of mitochondrial mass (VDAC1). (e) Basal respiration (n = 3) upon re-expression of wild-type or catalytically deficient mutant forms of SHMT2 in HEK293T knockout cell lines. Data are mean ± s.e.m. n indicates the number of biological replicates, which for the Seahorse experiments refers to independent plates on separate days. *P < 0.01, two-tailed Student’s t-test (see Supplementary Table 7 for exact P values). Cat. inact., catalytically inactive SHMT2; PLP bind., PLP binding-deficient SHMT2.
folate 1C enzymes (Fig. 2.1c). This selective requirement for SHMT2 to maintain oxidative phosphorylation was also confirmed in the HEK293T cell background (Fig. 2.1d,e). Consistent with respiratory chain deficiency, SHMT2 loss decreased glucose flux into the TCA cycle intermediate citrate, with an increased fraction of citrate instead being produced by reductive carboxylation[101] (Fig. 2.1g). As reported recently in other models of mitochondrial damage[14, 152], pool size of TCA cycle metabolites and associated amino acids was also decreased (Fig. 2.1f). To identify the cause of respiratory deficiency, we examined the abundances of several mitochondrial proteins, finding decreased abundance of complex I, IV, and V subunits with retained levels of complex II, III and markers of mitochondrial mass (Fig. 2.2d and Fig. 2.1h,i). Thus, SHMT2 is required to maintain levels of multiple mitochondrial respiratory chain proteins.

Given that loss of SHMT2, but not the immediate downstream enzymes of mitochondrial 1C metabolism, caused impaired oxidative phosphorylation, we were curious if the phenotype reflected a requirement for SHMT2’s catalytic activity, or alternatively a non-catalytic role of SHMT2, perhaps related to its reported interaction with the mitochondrial nucleoid[65, 122]. Accordingly, in the SHMT2 knockout background, we stably re-expressed catalytic inactive SHMT2 (p.Glu98Leu/p.Tyr106Phe), PLP-binding mutant SHMT2 (p.Lys280Gln) or wild-type SHMT2 protein (Fig. 2.3 and Fig. 2.4a). Re-expression of wild-type protein, but not the catalytically inactive mutants, rescued oxidative phosphorylation deficiency (Fig. 2.2e) and normalized glycolytic flux (Fig. 2.4b). In addition, SHMT2 re-expression rescued the growth defect of SHMT2 knockout cells[39] (Fig. 2.4c) and normalized 1C metabolism (Fig. 2.4e,f). Thus, mitochondrial SHMT catalytic activity is critical to sustain oxidative phosphorylation.

Two compartment-specific uses of mitochondrial folate 1C units have been reported: the local biosynthesis of deoxythymidine triphosphate (dTTP)[4, 18, 132] and of formyl-methionine (f-Met)[78, 161] (Fig. 2.5a). Production of dTTP requires 5,10-methylene-THF (methylene-THF), whereas f-Met requires 10-formyl-THF (formyl-THF). SHMT2 is upstream of both compounds, and thus SHMT2 loss would be expected to deplete both. In contrast, MTHFD2 sits between methylene-THF and formyl-THF. The lack of an oxidative phosphorylation phenotype with MTHFD2 knockout led us to hypothesize that methylene-THF is the required 1C species. Consistent with this, SHMT2 knockout cell lines showed unchanged n-terminal f-Met levels of the mitochondrially translated MTCO1 peptide, which has been previously reported be an indicator of cellular f-Met status[161] (Fig. 2.6a). To confirm that methylene-THF is the required species, we generated SHMT2/MTHFD2 double deletion cells and supplemented them with methylglycine (sarcosine), which can produce mitochondrial methylene-THF via sarcosine dehydrogenase. While sarcosine is
Figure 2.3: Catalytically deficient SHMT2 constructs. (a) Mapping of mutated amino acid residues on human SHMT1 (PDB code 1BJ4) using iCn3D and alignment of *E. coli* serine hydroxymethyltransferase (GLYA), *H. sapiens* mitochondrial serine hydroxymethyltransferase 2 (GLYM) and cytosolic serine hydroxymethyltransferase 1 (GLYC). Positions for GLYM are given with reference to GenBank NM_005412.5. (b) Sanger sequencing traces of mutant constructs. (c) Immunoblot for mitochondrial complex I levels (NDUFS4) in cell lines re-expressing catalytically deficient forms of SHMT2.
Figure 2.4: Restoring SHMT2 catalytic activity normalizes 1C flux, respiratory chain expression, glycolytic activity, and cell growth. (a) Immunoblot of re-expression of catalytically active SHMT2 (left) and the effects of its re-expression on mitochondrial complex I and II levels (right). (b-f) Effect of re-expression of catalytically active and inactive forms of SHMT2 in two different ΔSHMT2 clones in the HEK293T background. b, Normalized NAD+/NADH ratio (n = 6). c, Lactate secretion and glucose uptake (n = 6). d, Cell proliferation (n = 6). e, Purine biosynthesis intermediate 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) levels (n = 4) as an indicator of cytosolic folate 1C status. f, [2,3,3-²H]serine tracing to differentiate cytosolic from mitochondrial folate 1C unit production for incorporation into deoxythymidine triphosphate (n = 3). Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.
Figure 2.5: SHMT2-knockout-induced respiratory chain deficiency is caused by mitochondrial methylene-THF depletion but is unrelated to dTTP synthesis. (a) Sarcosine serves as an SHMT2-independent source of mitochondrial methylene-THF. (b) NAD⁺/NADH ratio (n = 6) and NDUFS4 (complex I) protein expression upon sarcosine supplementation (1 mM) in SHMT2 single-knockout (ΔSHMT2) and SHMT2/MTHFD2 double-knockout (ΔSHMT2/ΔMTHFD2) cell lines compared to wild-type cells. (c-d) Functional readouts for mitochondrial dTTP status based on mitochondrial DNA (mtDNA) levels (n = 3) determined by quantitative PCR (qPCR; c) and gene expression determined by RNA-seq (d) in SHMT2-knockout and wild-type HEK293T cells. RPKM, reads per kilobase per million mapped reads. In d, each data point represents the mean gene expression of two biological replicates of two independent knockout clones (n = 4) and two wild-type replicates (n = 2). Genes linked to OXPHOS function are highlighted in red (nuclear-encoded) or blue (mitochondrial-encoded). Data are mean ± s.e.m. n indicates the number of independent biological replicates. *P < 0.01, two-tailed Student’s t-test.

not a preferred 1C source and its feeding was insufficient to restore oxidative phosphorylation in SHMT2 single gene knockout cell lines, it fully restored oxidative phosphorylation in the SHMT2/ MTHFD2 double knockout cells, in which drainage of methylene-THF to formyl-THF is blocked (Fig. 2.5b and Fig. 2.6b). Thus, mitochondrial methylene-THF is required to maintain respiratory capacity.

Depletion of mitochondrial dTTP and/or accumulation of uridine nucleotides have been shown to induce mitochondrial respiratory chain deficiency by promoting mitochondrial DNA damage (depletion[131], deletions and point mutations[4, 106, 178]). We accordingly expected that SHMT2-induced methylene-THF depletion impaired respiration through mitochondrial DNA damage. However, SHMT2 knockout cells showed no evidence of altered mitochondrial DNA copy number (Fig. 2.5c and Fig. 2.6c), deletions (Fig. 2.6d), or mutations (Fig. 2.6g,h). Moreover, whole cell RNA-sequencing revealed normal transcript levels for all mitochondrial-encoded respiratory chain protein subunits (Fig. 2.5d). The only significantly differentially expressed gene related to mitochondrial respiratory chain function[92] was the nuclear-encoded gene FIS1 (Fig. 2.6e,f), which was moderately up-regulated as has been shown to occur in response to mitochondrial deficiency[63, 165]. Thus, the dependence of oxidative phosphorylation on SHMT2 reflects a requirement for mitochondrial methylene-THF for a purpose other than supplying local dTTP to maintain mitochondrial DNA.

Whereas the vast majority of the ~1,100 mitochondrial proteins are imported from the cytosol,
Figure 2.6: Oxidative phosphorylation defect is caused by a post-transcriptional mechanism independent of methionine formylation. (a) Fraction of initiating amino acid (formylmethionine versus methionine) of mitochondrial-expressed COX1 peptide determined by high-resolution LC–MS (wild type n = 4, ΔSHMT2 n = 3, ΔMTHFD2 n = 2). (b) Lactate secretion (n = 3) upon sarcosine supplementation (1 mM). (c) Relative mtDNA levels in HEK293T cells (n = 3). (d) Agarose gel of mtDNA long-range PCR products of HCT116 and HEK293T knockout cell lines. (e) Relative mRNA levels of mtDNA-encoded respiratory chain subunits in the HEK293T background (n = 3). (f) Gene expression levels in SHMT2-knockout cell lines compared to SHMT2 wild-type re-expressed lines by total RNA sequencing. Each dot represents mean gene expression as derived from two biological replicates of two independent knockout clones and matched re-expressed lines (n = 4). Genes linked to human OXPHOS function are highlighted in red. (g) Position-dependent next-generation sequencing coverage of mtDNA in HEK293T wild-type, SHMT2-knockout and MTHFD2-knockout cell lines supports the absence of deletions due to SHMT2 loss. (h) Corresponding variant position and frequency. Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.
13 essential respiratory chain subunits are locally transcribed and translated[22]. These include components of complexes I, III, IV, and V, but not complex II. Based on the normal mitochondrial transcript abundances and complex II protein levels, we hypothesized that mitochondrial methylene-THF is required for local translation. Indeed, screening for mitochondrial translation defects using $^{35}$S-methionine incorporation showed reduced synthesis of a subset of locally-expressed subunits of complexes I and IV (Fig. 2.7a). To more specifically probe mitochondrial translation in action, we developed a protocol for mitochondrial ribosome profiling, based on digesting unprotected mRNA with the nuclease MNase, enriching the 55S-mitochondrial ribosome, and sequencing the protected footprints (Fig. 2.8a and Fig. 2.7b). This approach achieved >90 % average sequence coverage of mitochondrial transcripts with an average depth of >80 reads per codon (for phasing and read-length distribution, see Fig. 2.7c). In SHMT2 knockout cells the distribution of ribosome-protected footprints of multiple genes showed pronounced stalling at defined codon positions (Fig. 2.8b and Fig. 2.9a). This resulted in relatively fewer actively translating ribosomes (i.e. bound and not stalled) for certain subunits of respiratory chain complexes I, IV, and V (Fig. 2.9b). Consistent with the ribosome profiling data, enzymatic assays revealed intact citrate synthase and complex II and III activity, but decreased activity of complexes I, IV, and V (Fig. 2.9c).

We next aimed to elucidate the cause of ribosomal stalling. The A-site coordinates of stalled ribosomes revealed striking ribosome accumulation in SHMT2 knockout cells at particular lysine and leucine codons: Lys$^{AAG}$ and Leu$^{UUG}$ (Fig. 2.8c). This did not appear to reflect shortage...
Figure 2.8: Mitochondrial ribosome profiling reveals that SHMT2-knockout cells are deficient in translating specific guanosine-ending codons. (a) Workflow of mitochondrial ribosome profiling. Translation was halted using chloramphenicol and immersion into liquid nitrogen, cells were lysed and RNA was digested using micrococcal nuclease (MNase). After sucrose-gradient enrichment for mitochondrial ribosomes (shaded in red), protected fragments were sequenced. (b) Mean cumulative ribosome density along selected mitochondrial transcripts. Additional transcripts are given in Fig. 2.9a. RPF, ribosome-protected fragment; RPM, reads per million. (c) Mean codon-specific mitochondrial ribosome occupancy in HCT116 cells (ΔSHMT2/wild type). Red data points correspond to codons that are decoded by tRNAs carrying the 5-taurinomethyluridine (m^5U) modification. Red labels correspond to the subset of codons that end in guanosine and thus require wobble-base pairing. Methionine codons are highlighted in blue and show no increased codon occupancy. The insert shows mean normalized ribosome density relative to UUG and AAG codon position. aa, amino acids. Data in b and c represent two technical replicates of two independent samples.
Figure 2.9: Impairment of mitochondrial translation due to loss of SHMT2. (a) Expanded version of Fig. 2.8b, showing the mean cumulative ribosome protected fragments of all mitochondrial protein-coding genes. (b) Mean relative density of actively translating (that is, not stalled) ribosomes for mitochondrial transcripts. Data in a and b represent two technical replicates of two independent samples. (c) Enzymatic activities of citrate synthase and individual mitochondrial respiratory chain complexes from mitochondrial extracts (n = 5). Data are mean ± s.e.m. *P < 0.01, two-tailed Student’s t-test. (d) Mitochondrial genetic code table with split codon boxes depending on taurinomethylated tRNAs for translation highlighted in red. Codons decoded by anticodon formylcytidine-containing tRNAMet are highlighted in blue. (e) Mean codon-specific mitochondrial ribosome occupancy of HCT116 SHMT2/MTHFD2 double-knockout cell lines supplemented with sarcosine (1 mM). Codons highlighted in red are decoded by tRNAs carrying a 5-taurinomethyluridine modification. The supplementation with sarcosine prevents the stalling normally observed with SHMT2 deletion (n = 2).
of these amino acids or tRNAs in mitochondria, as much less stalling was observed at the corresponding Lys^{AAA} and Leu^{UUA} codons. Instead, it appeared to relate to difficulty reading the 3’ codon guanosine of codons where the 3’ position identity (purine vs. pyrimidine) determines the encoded amino acid ( “split codon boxes”, Fig. 2.9d), with increased codon occupancy observed also for Trp^{UGG}, Glu^{GAG} and Gln^{CAG}, but not the corresponding codons with 3’ adenosine. The sole exception was Met^{AUG/AUA}, where no stalling was observed. Decoding of A/G ending codons in split codon boxes is facilitated by methyl-derivative base modifications of the tRNA anticodon 5’ nucleotide, allowing non-Watson-Crick base-paring with the codon 3’ base[2, 58, 71, 133] (Fig. 2.10a). The mitochondrial tRNA^\text{Met} anticodon has a 5’ cytidine which is formylated with the 1C unit derived from S-adenosyl-methionine (SAM)[59, 102, 164]. In contrast, the mitochondrial tRNAs for Lys, Leu1, Trp, Glu, and Gln have uridine at the 5’ anticodon position[120]. Mammalian cytosolic tRNAs with uridine at the anticodon 5’ position are modified to produce 5-methoxycarbonylmethyluridine (mcm^{5}U), with the methoxy carbon from SAM[48, 144]. Corresponding mitochondrial tRNAs are also modified at the 5’ anticodon uridine (position 34 in the tRNA), but carry a 5-taurinomethyl modification (τm^{5}U)[153, 172, 173]. As SHMT2/MTHFD2 double knockout cell lines supplemented with sarcosine showed rescue of codon specific stalling on mitochondrial ribosome profiling (Fig. 2.9e), we hypothesized that the 5-taurinomethyl modification is dependent on mitochondrial methylene-THF production and that the observed oxidative phosphorylation defect is a consequence of impaired translation due to defective tRNA modification.

To test this hypothesis, we established a high resolution liquid chromatography-mass spectrometry (LC-MS) method for detection of modified tRNA bases from mitochondrial tRNA extracts. In SHMT2 knockout cell lines, formylcytidine levels were unchanged (Fig. 2.11a). In contrast, τm^{5}U and its 2-thio derivative (where the uracil 2-position oxygen is replaced with sulfur) were depleted to undetectable levels (Fig. 2.10b; Fig. 2.11b). The depletion was not due to lack of taurine, whose cellular levels were mostly increased (Fig. 2.11c). The loss of τm^{5}U and its 2-thio derivative was reversed upon re-expression of wild-type SHMT2 and also upon sarcosine supplementation of SHMT2/MTHFD2 double knockout cells (Fig. 2.11d, e). Thus, SHMT2 or an alternative source of mitochondrial 5,10-methylene-THF is required to make τm^{5}U.

To verify that the methyl group of τm^{5}U is coming from 5,10-methylene-THF made by SHMT2, we conducted stable isotope tracing with [^{13}C]-labeled serine (the SHMT2 substrate) versus [^{13}C]-methionine (which feeds into SAM). As expected, methionine labeled formylcytidine, the wobble-position nucleotide in mitochondrial tRNA^\text{Met}. In contrast, serine but not methionine labeled τm^{5}U.
Figure 2.10: MTO1/GTPBP3-dependent tRNA methylation requires mitochondrial methylene-THF. (a) Interaction of tRNA position 34 anticodon loop modified base with mRNA codon 3 position A/G, forming a non-Watson–Crick base pair. (b) Total ion chromatogram of 5-taurinomethyluridine monophosphate (p-5tmU) (m/z = 460.043) from digested mitochondrial tRNAs. 5-formylcytidine monophosphate was not altered (Fig. 2.11a). (c) Mean codon-specific mitochondrial ribosome occupancy for MTO1-knockout (ΔMTO1) HCT116 cell lines and primary patient-derived fibroblasts carrying MTO1 mutations or the MT-TL1 m.3243A>G MELAS variant (n = 2). Corresponding immunoblots are shown below. Individual patient data are in Fig. 2.13a. (d) Basal respiration rates measured using the Seahorse XF analyser. Data were collected after growth in the absence (−) of folate for 5 passages or in the presence of the indicated methotrexate (MTX) concentration for 96 h (n = 3, except HCT116 WT n = 4 and MTX 50 nM n = 6). HT, 100 μM hypoxanthine and 16 μM thymidine. Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.

Thus, SHMT2 is required for oxidative phosphorylation because it produces mitochondrial methylene-THF for tRNA taurinomethylation. To our knowledge this is the first direct evidence of folate-dependent macromolecule modification in mammalian cells.

The enzyme complex catalyzing the τmU base modification in mitochondria comprises MTO1 (mitochondrial tRNA translation optimization 1) and GTPBP3 (GTP binding protein 3), both of which are human mitochondrial disease genes[53, 77, 156]. Its orthologues, forming the prokaryotic tRNA modifying MnmE/GidA complex, have been shown to use THF-bound 1C units[100, 140]. We therefore explored if the mitochondrial translation defect upon SHMT2 loss matches human diseases proposed to affect taurinomethylation of mitochondrial tRNAs[53, 77, 156, 172, 173]. Indeed, LC-MS analysis for modified mitochondrial tRNA bases showed loss of τmU in HCT116 cells with MTO1 knockout (Fig. 2.11e). Mitochondrial ribosome profiling of these engineered MTO1 deletion cells showed an increase in codon occupancy at LysAAG, LeuUUG, TrpUGG, GluGAG, and GlnCAG codons analogous to that observed with SHMT2 knockout (Fig. 2.10c). As with SHMT2 deficiency, the strongest stalling occurred at LysAAG and LeuUUG, potentially because these codons have, in addition to the wobble base pairing, only A-U base pairings versus the stronger C-G base pairings in TrpUGG, GluGAG, and GlnCAG codons[36]. Mitochondrial ribosome
Figure 2.11: tRNA modification status in ΔSHMT2 and effects of 5-taurinomethyluridine modification loss caused by human disease gene MTO1. (a) Total ion chromatogram of 5-formylcytidine monophosphate in digested mitochondrial tRNAs upon loss of SHMT2. The same samples were analysed for 5-taurinomethyluridine monophosphate (p-5m^5U) in Fig. 2.10b. The combined data demonstrate that SHMT2 deletion causes loss of 5m^5U but not 5-formylcytidine. (b) Levels of 5m^5U, 5-taurinomethyl-2-thiouridine monophosphate (p-5m^5 s^2U) and 2-thiouridine monophosphate (p-s^2U) in wild-type HCT116 and SHMT2 deletion lines normalized to 5-formylcytidine monophosphate (p-f^5C) (n = 3). (c) Taurine levels in HCT116 wild-type and SHMT2-knockout cells (n = 3). (d) 5m^5U levels in digested mitochondrial tRNAs upon re-expression of SHMT2 (n = 1). (e) 5m^5U, 5m^5 s^2U and s^2U levels normalized to f^5C in HCT116 SHMT2/MTHFD2 knockout lines after sarcosine supplementation and HCT116 upon loss of MTO1 (n = 2). For all panels, data are mean ± s.e.m. or individual data points only. (f) Labelling pattern of 5-taurinomethyluridine and 5-formylcytidine monophosphate extracted from mitochondrial tRNAs after growth in media containing either [3-13C]serine or [U-13C]methionine. (g) Mean cumulative count of ribosome protected fragments (RPF) mapping to mitochondrial protein coding transcripts upon ribosome profiling in HCT116 MTO1-knockout cell lines. Data were normalized to RPM (n = 2); n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.
profiling of primary fibroblasts of two patients carrying different MTO1 missense mutations (pat_a c.[1261-5T>G];[1430G>A], pat_b c.[1222T>A];[1222T>A]) revealed increased codon occupancy selectively at Lys$^{AAG}$, but not Leu$^{UUG}$. This perhaps reflects the hypomorphic mutations causing a selective defect in taurinomethylation of the Lys tRNA. Of all adenosine ending codons, Lys$^{AAA}$ showed the greatest increase in codon occupancy across all cell lines, being most pronounced in the MTO1 patients (Fig. 2.10c).

Intriguingly, in both SHMT2 and MTO1 deletion cells, stalling did not affect all AAG / UUG codons uniformly, but occurred most strongly at the same specific gene locations. This indicates that stalling upon tRNA modification loss depends not only the defective codon-anticodon interaction, but also the transcript position or surrounding mRNA context (Fig. 2.11g). Mapping of AAG / UUG codons at stalling sites relative to mRNA secondary structure[21, 129] did not reveal any clear pattern (Fig. 2.12a, b). Mapping onto the structure of the protein being synthesized (Fig. 2.12c), however, showed a trend towards enrichment for stalling at transitions between transmembrane helices and non-membrane domains. Thus, stalling due to the defective codon-anticodon interaction might be exacerbated by particular protein sequence features.

Defined mutations in the Lys and Leu1 mitochondrial tRNAs (which decode the most strongly affected Lys$^{AAG}$ and Leu$^{UUG}$ codons) result in a tRNA-specific tm$^5$U modification defect[172, 173], causing the human mitochondrial disorders MERRF and MELAS[5]. Mitochondrial ribosome profiling of fibroblasts from two MT-TL1 m.3243A>G MELAS patients revealed, as would be predicted from previous in vitro studies[74], increased occupancy at Leu$^{UUG}$ but not Lys$^{AAG}$ or Leu$^{UUA}$ (Fig. 2.10c). For unknown reasons, increased ribosome occupancy for either Ser$^{AGU}$ or Thr$^{ACG}$ was also observed in individual patients (Fig. 2.13a). The extent of stalling was less than with nuclear SHMT2 or MTO1 mutations, presumably due to the heteroplasmic nature of the mitochondrial tRNA mutation (Fig. 2.13b). Consistent with the weaker stalling phenotype, we also observed a milder defect in complex I protein content (Fig. 2.10c). Collectively these observations highlight a common biochemical mechanism that links mitochondrial folate metabolism with a significant fraction of inborn mitochondrial diseases[5, 53, 77, 156].

Based on these observations, we were curious whether folate deficiency, which in pregnancy causes neural tube defects[40, 47, 155], could also result in mitochondrial impairment via this mechanism. Consistent with such a possibility, growth of HCT116 wild-type cells in folate-deficient media resulted in decreased complex I enzyme levels and reduced basal respiration rate (Fig. 2.10d and Fig. 2.14a), without impacting mitochondrial DNA content (Fig. 2.14b). In individuals with adequate nutrition, functional folate deficiency can be caused by antifolate therapy for au-
Figure 2.12: Investigation of mRNA and protein secondary structure effects on mitochondrial ribosome stalling sites. (a) Identification of mitochondrial RNA secondary structure based on analysis of the mitochondrial transcript data from the dimethyl sulfate sequencing dataset published previously. R values and Gini differences were calculated to detect changes in nucleotide reactivity between the in vivo and denatured condition for the complete mitochondrial transcriptome. (b) Determination of ribosome stalling sites in SHMT2-knockout HCT116 cell lines. Data points represent individual codons of all 13 mitochondrial protein-coding transcripts. For each codon, the y axis indicates the ribosome counts normalized to the gene median in RPM. The x axis indicates the ratio of normalized counts in SHMT2-knockout to normalized counts in wild-type HCT116. Two and three s.d. above the mean of all codons in the genome are indicated by the grey and black dotted line, respectively. Highlighted in red are codons with greater than 2 s.d. (c) Mapping of AAG and UUG codons from SHMT2 knockout-specific ribosome stalling sites (>3 s.d.) on protein structures. For b and c, analysis is based on ribosome profiling data in Fig. 2.8, with two technical replicates of two independent samples.
Figure 2.13: Mitochondrial transcript codon occupancy from ribosome profiling of individual patient lines. (a) Codon-specific mitochondrial ribosome occupancy ratio (patient/control fibroblasts) in individual patient derived cell lines ($n = 1$ for each individual patient, normalized to mean of $n = 2$ control fibroblast lines). Patients either had nuclear MTO1 missense mutations (patient A c.[1261-5T>G];[1430G>A], patient B c.[1222T>A];[1222T>A]) or were diagnosed with MELAS and carry the recurrent point mutation m.3243A>G in the mitochondrial gene for tRNA Leu1 (MT-TL1). (b) Next-generation sequencing of mtDNA mutation load m.3243A>G (MT-TL1) in control fibroblasts and MELAS patient cell lines. Each bar shows one biological replicate for control and patient cell lines. Integrative genomics viewer sequencing raw data are shown on the right.
to immuno
to immunity or cancer. Low doses of the antifolate methotrexate, resulting in up to 100 nM circulating drug concentrations, are commonly used to treat autoimmune diseases such as rheumatoid arthritis [24, 57, 86, 112]. In cell culture, nanomolar concentrations of methotrexate resulted in decreased complex I enzyme levels and oxygen consumption (Fig. 2.10d and Fig. 2.14a). This decrease could arise from methotrexate depleting mitochondrial DNA, impairing mitochondrial translation, or a combination. In an effort to distinguish these possibilities, we compared the effects of methotrexate to ethidium bromide, which is classically used to deplete mitochondrial DNA, and to chloramphenicol, which blocks mitochondrial translation without altering DNA content. After 48 h of treatment, methotrexate and ethidium bromide both decreased oxygen consumption and DNA content. Importantly, despite ethidium bromide much more strongly depleting mitochondrial DNA (28 % vs. 68 %, p = 0.001), methotrexate had an equivalent effect on oxygen consumption (69 % vs. 69 %), consistent with methotrexate’s effect on oxygen consumption being in part via mitochondrial translation inhibition (Fig. 2.14c). By 96 h, both ethidium bromide and chloramphenicol had greater impact on oxygen consumption, consistent with their strongly blocking mitochondrial DNA or protein synthesis, respectively, whereas methotrexate exerted a weaker effect on both processes (Fig. 2.14c). Thus, the essential role for folate metabolism in mitochondrial translation may contribute to the clinical manifestations of folate deficiency and the clinical efficacy of antifolate therapies.

2.4 Materials and Methods

2.4.1 Cell lines and growth conditions

HCT116 (CCL-247) and HEK293T/17 (CRL-11268) were purchased from ATCC®. Generation of a subset of clonal CRISPR/Cas9 knockout cell lines and detailed characterization has been reported previously [39]. Additional clonal knockout cell lines (Supplementary Table 6a) were established following the protocol published by Ran and colleagues [124]. In brief, exon targeting guide RNAs (Supplementary Table 6b) were designed against genes of interest and cloned into an expression vector containing the double nicking Cas9 variant (Addgene). Cells were transiently transfected using Lipofectamine 2000 (Life Technologies) (HEK293T) or Fugene HD (Promega) (HCT-116) and selected for 48 h with 2 μg/mL puromycin. Single clones were isolated using serial dilution into 96-well plates. Stable SHMT2 re-expression was achieved by transfecting HEK293T knockout cell lines with NM_005412.5 cDNA (GE-Healthcare) cloned into pCMV-Tag8 vector (Agilent) and selection for three weeks with 200 μg/ml Hygromycin B (Sigma-Aldrich). Catalytic in-
Figure 2.14: Effects of targeting 1C metabolism on mitochondrial function. (a) Mitochondrial complex I and II levels after growth in the absence of folate for five passages or in the presence of the indicated methotrexate concentration for 96 h. Ethidium bromide (250 nM) was used as a positive control. (b) Cellular mtDNA levels in HCT116 cells after folate depletion (with or without 100 μM hypoxanthine and 16 μM thymidine (HT) as rescue agents) or in the presence of methotrexate for 96 h (n = 3). (c) To determine whether the decrease in respiration due to methotrexate arises from methotrexate depleting mitochondrial DNA, impairing mitochondrial translation, or a combination, in HCT116 cells we compared the effects of methotrexate (50 nM) to ethidium bromide (250 nM = 100 ng ml⁻¹), which is classically used to deplete mitochondrial DNA, and to chloramphenicol (310 μM = 100 μg ml⁻¹), which blocks mitochondrial translation. After 48 h of treatment, methotrexate and ethidium bromide both decreased oxygen consumption and DNA content. Importantly, despite ethidium bromide depleting mitochondrial DNA much more strongly, methotrexate had an equivalent effect on oxygen consumption, consistent with the effect of methotrexate on oxygen consumption being in part via mitochondrial translation inhibition. Data are normalized and compared to untreated control (all n = 3; except oxygen consumption methotrexate 96 h n = 6 and control n = 4). Data are mean ± s.e.m. n indicates the number of biological replicates. *P < 0.01, two-tailed Student’s t-test.
active (p.Glu98Leu/p.Tyr106Phe)[27, 154] and PLP-binding deficient (p.Lys280Gln)[69] mutants were obtained following the QuickChange II protocol (Agilent). Knockout and re-expression cell lines were functionally verified by immunoblotting followed by targeted genomic sequencing (Supplementary Fig. 7) and, in the case of SHMT2 cell lines, also by tracing of [2,3,3H]-serine labeling into dTTP. MTO1 and MT-TL1 patient fibroblasts and controls were provided by the Department of Pediatrics, Salzburger Landeskliniken and Paracelsus Medical University, Salzburg. For studies with primary human cell lines informed consent was obtained from all subjects. Genotype of MTO1 deficient patients (GenBank NM_012123.3) was as follows: pat_a c.[1261-5T>G]; [1430G>A], (p.[?];[Arg477His]); pat_b c.[1222T>A];[1222T>A], (p.[Ile408Phe]; [Ile408Phe]). Patient MTO1_pat_b has been reported before[156]. Both MELAS patients carried the common MT-TL1 m.3243A>G mutation with the heteroplasmy rate reported in this study. All cell lines tested negative for mycoplasma and were cultured in DMEM without sodium pyruvate (Sigma-Aldrich) supplemented with 10% dialyzed fetal bovine serum (dFBS, GE Healthcare) in a 5% CO₂ incubator at 37°C. No antibiotics were used.

2.4.2 Glucose uptake and lactate secretion

Cells were seeded in 6-well plates 24 h before the start of an experiment. After reaching 50% confluency, plates were washed with phosphate buffered saline (PBS, GE Healthcare) and 3 ml of fresh media was added. Glucose uptake and lactate secretion were determined using an YSI 2900D Biochemistry Analyzer (Xylem Analytics) and normalized to cell growth as determined by μl packed cell volume (PCV, as measured using packed cell volume microfuge tubes from TPP).

2.4.3 Proliferation

Proliferation assays were conducted in 96-well plates and relative cell number was measured using resazurin sodium salt. 5,000 cells were plated in each well with 150 μl DMEM supplemented with 10% dFBS. Cell growth at each day was read as fluorescence intensity using a Synergy HT plate reader (BioTek Instruments).

2.4.4 Oxygen consumption

Oxygen consumption rates (OCR) were measured on a XF24 extracellular flux analyzer (Agilent) following the manufacturer’s instructions. In brief, XF24 cell culture microplates were coated with fibronectin (Sigma-Aldrich) and cells were seeded at 5 x 10^4 (HEK293T) and 7 x 10^4 (HCT116) cells per well. After reaching 70-90 % confluency, cells were equilibrated for 1 h in XF assay medium
supplemented with 10 mM glucose, 1 mM sodium pyruvate, and 2 mM glutamine in a non-CO\textsubscript{2} incubator. OCR was monitored at baseline and throughout sequential injections of oligomycin (1 \textmu M), carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone (1 \textmu M) and rotenone/antimycin A (0.5 \textmu M each). Data for each well were normalized to cell number as determined by CyQUANT\textsuperscript{®} Cell Proliferation Assay Kit (Invitrogen) and to the on-plate wild-type control. For absolute OCR values protein concentration was determined by BCA protein assay (Thermo Fisher Scientific).

2.4.5 Immunoblotting

Cells were cultured to sub-confluency in 6 cm plates. After removal of media, cells were rinsed with 4\textdegree C PBS and lysed in radio-immunoprecipitation assay (RIPA) buffer with phosphatase and protease inhibitors (Roche). Lysates were cleared by 10 min centrifugation at 16,000 x g and quantified using a BCA assay (Pierce). Samples were resolved by sodium dodecyl sulfate–polyacrylamide gel electrophoresis on precast gels (Bio-Rad) and transferred to a nitrocellulose membrane using the Trans-Blot Turbo system (Bio-Rad). After overnight incubation with primary antibodies, bands were visualized with horseradish peroxidase-conjugated secondary antibodies (Cell Signaling Technologies). ChemiDoc\textsuperscript{TM} XRS+ system was used for image acquisition. The following antibodies were used according to their manufacturer’s directions: Anti-SHMT1 (12612), SHMT2 (12762), MRPL11 (2066), S6RP (2217) and β-actin (5125) were from Cell Signaling Technologies; Anti-MTHFD2 (ab151447), NDUFS4 (ab139178), SDHA (ab14715) and VDAC (ab14734) were from Abcam Inc.; Anti-MTHFD1L (HPA029041) was from Sigma-Aldrich.

2.4.6 Analysis of mitochondrial specific translation

\(^{[35}S\text{-labeling of mitochondrial proteins was performed following the method described by Sasarman and Shoubridge[134]. In brief, cells were grown on 6 cm plates for 48 h to sub-confluency. Then media was changed to DMEM with 10 % dFBS without methionine (MP Biomedicals). Following a 30 min incubation, cytosolic translation was inhibited by emetine hydrochloride (0.1 mg/ml, Sigma-Aldrich) and labeling was conducted for one hour after the addition of 500 \textmu Ci \(^{[35}S\text{-methionine (EasyTagTM L}\(^{[35}S\text{-Methionine, Perkin Elmer). Protein lysates (30 \mu g) were then separated on a 15% polyacrylamide gel (8.3 x 7.3 cm) and dried using a 443 Slab Dryer (BioRad). The dried gel was exposed to a storage phosphor screen (GE-Healthcare) and imaged on a Typhoon FLA 9500 (GE-Healthcare). Equal sample loading was confirmed by Coomassie brilliant blue staining (Biorad).
2.4.7 Metabolite concentrations and labeling patterns

Cells were grown in 6 cm dishes for at least 48 h and collected at 75% confluency. Media was replaced every 24 h and additionally 6hrs before harvesting. Metabolism was quenched and metabolites were extracted by aspirating media and immediately adding 1 ml of 80:20 methanol:water at −80°C. Plates were kept on ice, scraped and non-soluble debris was pelleted at 18,000 x g for 10 min. Samples were directly analyzed by hydrophilic interaction chromatography coupled with negative-mode electrospray-ionization high resolution mass spectrometry on a quadrupole-orbitrap scanning from m/z 73 to 1,000 at 1 Hz and 140,000 resolution (Q Exactive Plus, ThermoFisher). LC separation was achieved on a XBridge BEH Amide column (2.1 mm x 150 mm, 2.5 µm particle size, 130 Å pore size; Waters) using a gradient of solvent A (20 mM ammonium acetate + 20mM ammonium hydroxide in 95:5 water:acetonitrile, pH 9.45) and solvent B (acetonitrile). Flow rate was 150 µl/min. The gradient was: 0min, 85% B; 2min, 85% B; 3min, 60% B; 9 min, 60% B; 9.5min, 35% B; 12min, 35% B; 12.5min, 0% B; 18min, 0% B; 18.5min, 85% B; 23min, 85% B. Data were processed and analyzed using in-house MAVEN software[25]. All isotope tracer experiments were conducted at isotopic steady state: [2,3,3-²H]-serine was traced into dTTP for a minimum of 6 h and [3-¹³C]-serine and [U-¹³C]-methionine were traced into tRNA for a minimum of 4 days. Isotopic tracers were purchased from Cambridge Isotope Laboratories. Isotopically labeled media was prepared from scratch and supplemented with 10% dFBS.

2.4.8 RNA-Seq

RNA was isolated from cell lines using RNeasy Plus kit (Quiagen) according to the manufacturer’s recommendation. Following the depletion of ribosomal RNA, libraries were prepared according to the TruSeq Stranded Total RNA protocol (Illumina) and sequencing was performed on a HiSeq 2500 (Illumina). Analysis was performed using the Galaxy system[1] and the R software package. Adapter sequences were trimmed using Cutadapt (Galaxy version 1.6)[91] and the trimmed reads were then mapped with TopHat (Galaxy Version 0.9) to the GRCh38 reference using ENSEMBL version 80 genes as known splice junctions. The read counts per gene were determined using htseq-count (Galaxy Version 0.6.1galaxy1) in ‘union’ mode. Differential expression analysis was performed in R using DESeq2 1.12.3 package. SHMT2 knockout gene expression (log2 reads per kilobase of transcript per million mapped reads (RPKM)) was graphed relative to the wild-type and the re-expressed cell lines.
To analyze mitochondrial DNA content, total DNA was extracted from $7 \times 10^6$ cells using Gentra Puregene Cell Kit (Quiagen) after freezing the cell pellet at -80°C for 1 h and overnight digestion with Proteinase K (Roche Diagnostics). Quantitative real-time PCR (ViiA 7, Applied Biosystems) was performed using primers targeting the mitochondrial ND2 locus (fwd: TGTTGGTTATACCCTTCCCGTACTA; rev: CCTGCAAAGATGGTAGATGATGA) and a nuclear ALU repeat sequence (fwd: CTTGCAGTGAGCCGAGATT; rev: GAGACGGGAGTCTCGCTCTGTC) as published earlier[7]. The relative mitochondrial DNA content was determined using the $\Delta \Delta \text{Ct}$ method. Each independent sample given in the figures represents the mean of 6 technical replicates. The mitochondrial genome was screened for deletions by long range PCR with two primer pairs spanning the whole coding region (fwd_1: CCAACCAAAACCCCAAAGAC, rev_1: TACTGCGACATAGGGTGCTC; fwd_2: CACCAGGCTAACCAGATTTCArev_2: TGGTACCCAAATCTGCTTTCC) and products run on a 1% agarose gel[93]. DNA from a mitochondrial DNA deletion patient was used as positive control. For mitochondrial genomic sequence analysis, mtDNA was enriched using the multiple displacement amplification strategy (REPLI-g, Quiagen) and sequencing was performed on a HiSeq 2500 after library preparation following the Nextera library prep kit protocol (Illumina). Reads were mapped to GRCh38 using Bowtie2 (Galaxy Version 0.6)[81] with default settings. Coverage plots were generated using DeepTools bamCoverage (Galaxy Version 2.3.6.0)[123]. The data were normalized to 1x coverage using an effective genome size of 16,569. Freebayes (Galaxy Version 0.4.1)[50] with frequency-based pooled settings was used to generate the variant data. The figures were generated in R using Gviz 1.18.0[61]. MT-TL1 mutation load was determined using primers specifically spanning the m.3243 position for targeted enrichment (fwd: AATGATACGGGACCACCAGATCTACACNNNNNGCCTTCCCCGTAAATGATA, rev: CAAGCAAGACGGCAATACGAGATCGAGGGTTGTAGT) followed by sequencing on a MiSeq nano flow cell using a custom sequencing primer (seq: TATTATAAACCACACCACCAAGAAGGGTGTTTAAG). Alignment to GRCh38 was performed using Bowtie2 (Galaxy Version 0.6)[81] at default settings and position specific mutation load was derived from the Integrative Genomics Viewer.

Development of our ribosome profiling method was based on concepts reported by other groups[28, 67, 127]. For mitochondrial ribosome profiling, cell lines were grown on 15 cm plates to 70-85% confluency. Sarcosine rescue of ribosome stalling in the SHMT2/MTHFD2 double-knockout back-
ground was assessed after growth in the presence of 1 mM sarcosine for 5 days. After removal of media, plates were rapidly rinsed with ice cold phosphate buffered saline containing chloramphenicol (100 µg/ml) (Sigma-Aldrich) and cycloheximide (100 µg/ml) (Sigma-Aldrich) followed by immediate immersion into liquid nitrogen. Plates were then transferred to wet ice and 1 ml of 1.5x lysis buffer was added and the lysate was collected using a cell scraper. Lysis buffer contained the following: 1.5% triton X-100 (Sigma-Aldrich), 0.15% NP40 (Sigma-Aldrich), 1x complete phosphatase and protease inhibitors (Roche), and 30 U/ml DNase 1 (Roche) in buffer base (20 mM Tris – HCl pH 7.8 (Ambion), 100 mM KCl (Ambion), 10 mM MgCl₂ (Ambion), 100 µg/ml chloramphenicol, 100 µg/ml cycloheximide). 1.6-1.8 ml were recovered per plate and homogenized by passing three times through a 32G needle at 4°C. Non-soluble debris was pelleted at 5,000 g for 10 min and 1,520 µl supernatant was used for digestion with 2,250 U Micrococal nuclease (Roche) after adding 40 µl SUPERaseIN (Ambion) and 5 mM CaCl₂ (Ambion). Digest was stopped after 1 h gentle shaking at 25°C using a final concentration of 6 mM EGTA. Buffer base was used to make 5% - 45% sucrose gradients (Gradient Master, Biocomp). After cooling to 4°C, samples were separated in an ultracentrifuge using the SW-41Ti rotor at 35,000 rpm for 2.5 h. Live UV absorption at 254 nm was used to track the mitochondrial 55S monosome enriched fractions (Extended Data Fig. 5b). The 55S fractions were pooled and mixed with 57 µl 20% SDS per ml sample before performing acid phenol chloroform RNA extraction. RNA was precipitated using 300 mM sodium acetate pH 5.5 and equal volume isopropanol and run on a 15% TBE-urea gel (Invitrogen) at 210 V for one hour for size selection. Gels were stained with Sybr®Gold (Invitrogen) and RNA fragments corresponding to mitochondrial ribosome footprints (~28-40 nucleotides) were cut and recovered from the gel using the crush and soak method. After sodium acetate / isopropanol precipitation library preparation was conducted following the TruSeq®Ribo Profile (Illumina) protocol. Sequencing of ribosome protected footprints (RPFs) was performed on a HiSeq 2500 in rapid mode followed by adapter trimming using Cutadapt (Galaxy Version 1.6)[91]. Reads were mapped to the human genome reference GRCh38 using BWA (Galaxy Version 0.9)[85] with those mapping to the mitochondrial protein-coding genes included in the subsequent analysis. The Plastid package[42] and customized Phyton and R scripts were used for analyzing mitochondrial ribosome profiling data. Each read, corresponding to a mitochondrial ribosome protected fragment (mtRPF), was assigned to a nucleotide position representing the ribosomal A-site after offset determination by metagene analysis[42]. mtRPF counts were then normalized to reads per million (RPM) mapped reads within each sample and single nucleotide positions were grouped by codon index. This transformation allows for relative quan-
tification of bound ribosomes for each nucleotide triplet along a transcript. Stalling plots were created by plotting the mean cumulative mtRPF count along each mitochondrial open reading frame. Codons were defined as stalling sites when the normalized counts mapped to the specific codon (mtRPFcodon / mtRPF gene median) exceeded 2 standard deviations from all codons in the genome. The relative abundance of actively translating ribosomes (i.e. not stalled) was calculated by subtracting mtRPF counts in stalled regions from the total sum of ribosome footprints for each gene as \[ \sum_{\text{mtRPFactive}} = \sum_{\text{mtRPFtotal}} - \sum_{\text{mtRPFstalled}} \]. Then the gene specific ratio was plotted as \[ \frac{\sum_{\text{mtRPFactive,SHMT2}}}{\sum_{\text{mtRPFactive,wild--type}}} \]. Stalling sites specific to the SHMT2 knockout condition were identified using the ratio of occupancy at each codon position relative to wildtype. Specifically, codons were defined as SHMT2-specific stalling sites when the normalized counts in the mutant relative to WT (\( \sum_{\text{mtRPFcodon,SHMT2}} / \sum_{\text{mtRPFcodon,wild--type}} \)) exceeded 2 (or, as indicated, 3) standard deviations from this ratio as determined for all codons in the genome, and the site also met the general stalling site criterion. To determine the relative abundance of mitochondrial ribosomes bound to each nucleotide triplet, codon-specific occupancy ratios were calculated. For each codon (\( \text{codon}_{i=1-64} \)), the gene-specific ratio between experimentally measured ribosome density and expected density (which is proportional to codon frequency) was calculated. Codon occupancy (\( \text{CO}_{i=1-64} \)) for each codon is the mean of the ratios from all 13 genes. The relative codon occupancy (\( \text{CO}_{i=1-64,\text{SHMT2}} / (\text{CO}_{i=1-64,\text{wild--type}}) \)) was plotted with error bars representing standard deviation across replicates after error propagation. To investigate the ribosome distribution relative to the major stalled codons (AAG and UUG), ribosome densities flanking the codons of interest within 25 amino acids were selected. Each selected fragment was first normalized to its total count so every codon of interest from the genome is weighted equally. The mean value from each position was plotted. All sequencing data presented in this manuscript, except for ribosome profiling of primary patient cell lines, has been submitted for public availability to the sequence read archive submission code SUB2743534.

**2.4.11 Investigating protein secondary structure effects on stalling at AAG / UUG**

On average stalling was most pronounced at AAG and UUG codons, but not all codons of the same sequence were equally affected. We therefore investigated whether the positioning relative to protein secondary structures (transmembrane helices) influences the extent of stalling. As Micrococcal nuclease treatment induces imprecision in A-site mapping due to sequence-biased digestion\[41, 111\], individual positions identified as SHMT2-specific stalling sites were first grouped to the adjacent codons decoded by 5-taurinomethyluridine-modified tRNAs. Then amino acid
residues corresponding to the AAG / UUG codons were mapped to Bos taurus crystal structures of mitochondrial respiratory chain complex proteins using iCn3D[166]. Structure data was retrieved using the following PDB IDs: ATP6: 5ARA_W; MT-CO2: 2Y69_B; MT-CYB: 1QCR_C; and MT-ND6: 5LDW_J. Additionally, a hidden Markov model based algorithm for transmembrane helices (TMHMM 2.0[79]) was used to predict alpha helical transmembrane domains in the Homo sapiens sequences. This method assigns each codon a probability for transmembrane helix localization which was then used for genome wide assessment of AAG / UUG localization relative to transmembrane helices. AAG / UUG codons were defined to be at a transition between a transmembrane helix and a non-membrane region if, within the five flanking codons, probabilities > 0.5 and < 0.5 for being in a transmembrane helix are found. AAG / UUG codons were defined as stalling sites based on the 3 SD cut-off as per Extended Data Figure 8b. In total, 4/5 stalling AAG / UUG and 7/23 non-stalling AAG / UUG were at a membrane transition (p = 0.04 by Chi-square).

2.4.12 Evaluation of mRNA secondary structure effects on ribosome stalling

To study a potential effect of mRNA secondary structure (i.e. base pairing) on ribosome stalling we used the previously published dimethyl sulfate sequencing datasets on human K562 cell lines from Rouskin et al.[129] to identify structured regions in mitochondrial transcripts. Following the methods described in the manuscript for nuclear transcripts, identification of sites with secondary structure was performed on mitochondrial transcript data[129]. In brief, FASTQ files (accession no. GSM1297495, GSM1297493) were retrieved from sequence read archive and mapped to GRCh38 with BWA (Galaxy Version 0.9)[85]. Reads were assigned to the nucleotide at the five-prime end with no offset using Plastid[42]. R-value (cutoff 0.75) and Gini differences (cutoff 0.1) between the in vivo and denatured dataset were calculated for the complete mitochondrial transcriptome for a window size of 50 adenosine/cytosine nucleotides and a step size of 10. This provided a list of structured mitochondrial transcript regions, with most mRNA regions unstructured. The list of structured regions was compared to the SHMT2-specific stalling sites for potential co-localization. No stalling site mapped to a structured region.

2.4.13 Mitochondrial enzyme activities

Activities of individual OXPHOS complexes I-IV, ATP synthase and citrate synthase (CS, which is nuclear encoded and was used as a marker of mitochondrial mass) were spectrophotometrically measured as previously described (Uvicon 922, Kontron)[45, 130, 147]. Measurements were performed with 2 µl of mitochondria isolated by differential centrifugation[26] (except for
complex I, where 10 µl was used). CS (EC 2.3.3.1) activity was determined following extinction dynamics at 412 nm, indicating the cleavage of Elman’s reagent (0.2 mM) after addition of oxaloacetate (0.5 mM) to the buffered reaction solution containing acetyl-CoA (0.15 mM). Rotenone-sensitive complex I (NADH:decylubiquinone oxidoreductase, EC 1.6.5.3) activity was measured by adding NADH (0.2 mM) and monitoring at 340 nm for the reduction of decyl-ubiquinone (50 µM). Complex II (succinate:ubiquinone-oxidoreductase, EC 1.3.5.1) was measured at 600 nm by monitoring the reduction of 2,6-dichlorophenol-indophenol (80 µM) after addition of succinate (10 mM). The reaction mixture to determine complex III activity (coenzyme Q:cytochrome c—oxidoreductase, EC 1.10.2.2) contained cytochrome c (100 µM) and decyl-ubiquinol (200 µM) and was measured at 550 nm. After inhibition by addition of antimycin A (1 µM) the insensitive activity was subtracted to calculate specific complex III activity. The enzyme activity of complex IV (ferrocytochrome c:oxygen oxidoreductase, EC 1.9.3.1) was read as the oxidation rate of reduced cytochrome C (60 µM) at 550 nm. Complex V (F1F0 ATP synthase, EC 3.6.3.14) was indirectly measured as oligomycin-sensitive ATPase activity in a reaction mixture containing 0.5 mM ATP. Formed ADP was coupled to a pyruvate kinase reaction, utilizing phosphoenolpyruvate (2 mM) to generate ATP and pyruvate. The latter is then used by lactate dehydrogenase in the oxidation of NADH (0.2 mM) that served as direct readout (340 nm). Reagents were obtained from Sigma-Aldrich.

2.4.14 Mitochondrial tRNA modifications

HCT116 or HEK293T cell lines and sub clones were grown to 70-85% confluency and harvested for mitochondrial extraction (1-2 x 10^8)[26]. Mitochondrial tRNAs were extracted using the MirVana™ miRNA Isolation kit (Ambion) for isolation of small RNAs followed by 10% TBE-urea gel purification and extraction as described above (tRNA fraction 65-85 nt). Quantitative analysis of 5-taurinomethyluridinemonophosphate (τm5U), 5-taurinomethyl-2-thiouridine monophosphate (τm5s2U), 2-thiouridine monophosphate (s2U) and 5-formylcytidine monophosphate (f5C) was performed using high resolution mass spectrometry following a protocol adapted from[48, 179]. In brief, 100 ng tRNA were digested at 37°C for 2 h by nuclease P1 (2 U) in 30 µl of 100 mM ammonium acetate and quenched by adding 60 µl 50/50 methanol/acetonitrile followed by centrifugation at 10,000 g for 10 min. 5 µl of sample was injected for liquid-chromatography mass spectrometry analysis. Nucleoside monophosphates were analyzed on a quadrupole-orbitrap mass spectrometer (Q Exactive plus, Thermo Fisher Scientific) operating in negative ion mode coupled to hydrophilic interaction chromatography via electrospray ionization and used to scan in SIM mode.
from m/z 459 to 482 (\(rm5U\) and \(rm5s2U\)) or m/z 338 to 355 (\(s2U\) and \(f5C\)) at 1 Hz and 140,000 resolution. Consistent loading was ensured by measuring \(f5C\) levels (which are not altered by SHMT2 knockout) in the same sample as the \(rm5U\) and \(rm5s2U\). The mass spectrometry standard for 5-taurinomethyluridine was synthesized as described by Ogata et al.[110] and was used to confirm peak identity in digested tRNA samples after an additional treatment with 1 U of alkaline phosphatase (Roche) in 100 mM ammonium carbonate at 37°C for 2 h[179].

2.4.15 Folate depletion and methotrexate treatment

Cells were grown for five passages in folic acid-deficient DMEM with 10% dFBS (US Bio), either with or without HT supplement (Gibco), containing sodium hypoxanthine and thymidine at a final concentration of 100 µM and 16 µM respectively (1x HT). To evaluate the effect of targeting 1C-metabolism on mitochondrial function, methotrexate (Sigma-Aldrich) was used at 25, 50 and 75 nM concentration. Chloramphenicol (100 µg/ml = 310 µM) and ethidium bromide (100 ng/ml = 250 nM) served as positive controls for the inhibition of mitochondrial translation and mtDNA depletion respectively. At each time point protein and DNA samples were collected for immunoblot and mtDNA content analysis. Basal respiration was assessed using the Seahorse XF analyzer with measurements conducted as described above.

2.4.16 N-terminal protein formylation

N-formyl methionine modification on MTCO1 (COXI) was assayed by mass spectrometry following the protocol of Tucker et al.[161]. In brief, mitochondria were isolated from cells by differential centrifugation[26] and complex IV was immunoprecipitated (ab109801, Abcam) before separation on a 4-20% polyacrylamide gel. A band running at the same molecular weight as a band reactive with an anti-MTCO1 antibody (ab14705, Abcam) was excised and lysed using 1.5 µg LysC (Wako) as described[141]. Samples were dried in a speedvac and re-suspended with 15 µl of 0.1% formic acid pH 3. Per run, 5 µl were injected using an Easy-nLC 1000 UPLC system. Samples were loaded directly onto a 45 cm x 75 μm nano-capillary column packed with 1.9 µm C18-AQ (Dr. Maisch) mated to a metal emitter (Thermo Scientific) in-line with a Thermo Orbitrap Elite or Thermo Orbitrap Lumos. The mass spectrometer was operated in data-dependent mode with the 120,000 resolution MS1 scan (400-1800 m/z) in the Orbitrap followed by up to 20 MS/MS scans in the ion trap. Raw files were searched using MS Amanada[37], Sequest HT[43] and Byonic[13] algorithms and validated using the Percolator algorithm[146] within the Proteome Discoverer 2.1 suite (Thermo Scientific). 10 ppm MS1 and 0.6 Da MS2 mass tolerances were specified.
Carbamidomethylation of cysteine was used as fixed modification and oxidation of methionine as dynamic modification. Additionally, acetylation, formylation and loss of methionine were specified as potential modifications at the n-terminus of proteins. The resulting msf file was used to construct a spectral library (percolator peptide q-value >0.95) and extract MS1 ion chromatographs in Skyline100,101. The fraction of modified n-terminal peptides of MT-CO1 was calculated as the area of the formylated peptide divided by the sum of the areas of all the n-terminal peptides in that sample.

2.4.17 Statistical methods

Significance was determined by two-tailed Student’s t-test comparing the indicated condition to the corresponding wild-type or control. P values <0.01 are indicated as *. Exact p-values for individual comparisons are given in Supplementary Table 7. Small filled circles are individual data points.

3 Synchronized Growth Oscillation in Glucose-Limited Bacteria

3.1 Abstract

Cells organize their metabolic activity spatially and temporally. Understanding how this organization is achieved is fundamental to understanding metabolic regulation. Eukaryotic cells are known to use organelles as compartments that organize metabolic activity spatially, however less is known about how metabolic activity is organized in time. Recent work reveals that mitochondria exhibit rhythmic respiration, indicating that this organelle also exercises temporal control of its activity. However, whether such temporal organization is a primitive trait that exists in prototrophic bacteria is not known. Here, we report an unexpected oscillation in oxygen levels in E. coli glucose-limiting chemostats, suggesting that bacteria also temporally regulate their metabolic activity.
In our experiments we discovered that population density exhibits a two-fold change during the oxygen oscillation. Interestingly, transcriptional profiling showed that two major clusters of cellular metabolic genes are expressed during different phases of the oxygen oscillation. Using GO term enrichment analysis, we discovered Cluster 1 includes genes involved in ribosomal and tRNA synthesis as well as other macromolecule biosynthetic processes. Cluster 2 includes genes involved in respiration, redox processes and catabolic activities. Moreover, these gene clusters also exhibit growth-rate dependency of expression. Future studies will focus on identifying the mechanism and function of these surprising population-level dynamics.

3.2 Introduction

Metabolism drives all physiological activities including producing energy and building blocks of biomass. Its activity must be tightly regulated both spatially and temporally at different scales. At the cellular level, spatial coordination of metabolism includes membrane-bound organelles such as peroxisomes, mitochondria and chloroplasts. Recent studies have also revealed localization of enzyme complexes drives differential metabolite concentrations in the cytosol. At the organism level, different organs express different sets of metabolic enzymes. However, less is known about the temporal regulation of metabolism.

Oscillatory behaviors are prevalent in Nature and they are important to the physiology and fitness of organisms. Circadian rhythm is the most well-known example of temporal regulation of metabolism, which drives the day-night behavioral cycles in organisms ranging from human to cyanobacteria.

In 2005, Tu et al. [158] published a study where they characterized the yeast metabolic cycle (YMC) using microarray to probe gene expression. This phenomenon has been studied for over 45 years [95] and the authors divided the oxygen consumption curve into three phases and found different functional gene clusters are enriched differently. During the oxidative phase, genes involved in amino acid synthesis and translation are up-regulated. During reductive phase, DNA replication, peroxisomal and mitochondrial genes are expressed. The authors reason such oscillatory behavior as a way to compartmentalize metabolic processes. For example, translation requires energy and thus exposes cells to high oxidative stress generated from OXPHOS. Such oxidative environment would be harmful for the fidelity of DNA replication. Therefore, temporal partitioning in metabolism and gene expression could allow cells to separate potentially incompatible reactions.
Despite the accumulating evidence of temporal regulation of metabolism, little is known if metabolic cycle exists beyond yeast. Here, we reported a serendipitous discovery that under certain condition, E.coli cells exhibit oscillatory growth and oxygen consumption patterns similar to YMC. We characterize the system further using microarray and show that the transcriptional features are similar between the two systems such that biosynthesis and respiration genes are expressed differentially in time.

3.3 Results

During the investigation of glucose-limitation growth of E.coli, we serendipitously noticed that the oxygen level oscillated over time (Figure 3.1A). The oscillation happened after cultures were switched from batch to chemostat mode for a few hours. We called this oscillatory oxygen consumption behavior “the bacterial metabolic cycle” (BMC) to refer to the resemblance of the yeast metabolic cycle (YMC)\[158\]. The bacteria metabolic cycle started about five hours after the media pump was turned on and tended to damp down within 24 hours. The period of the oscillation is correlated with the dilution rate of the culture. Despite of the lack of knowledge about what triggers the synchronization of the culture, this phenomena was reproducible using the same growth media, initial size of the inoculation population and timing of the pump engagement (See Method).

Due to the long period of slow-growing cultures and similarity between the shapes of oscillatory oxygen curves, we focused on characterizing the metabolic cycle at a faster growth rate (0.6 h\(^{-1}\)). The oscillation of oxygen consumption reflected changes in metabolic activities, such that cell growth might also be time-dependent. Indeed, cell density of the culture also oscillated about two-
fold relative to the final OD when it reached steady state (Figure 3.1B). To understand how the cellular metabolic activities change, we extracted RNA at multiple time points over one period of the bacterial metabolic cycle and analyzed the dynamics in gene expression using microarray (See method).

By characterizing the gene expression dynamics over the entire period, we discovered that of the 3242 genes detected, 70% of them showed oscillatory expression. These genes can be further grouped into three major gene clusters that exhibit different time-dependent profiles (Figure 3.2A). Interestingly, they also reveal functional enrichments by GO term analysis (Table 3.1). During low oxygen consumption phase (high oxygen level, Cluster 1), genes enriched in ribosome biogenesis and nucleotide metabolism are highly expressed. On the other hand, when oxygen consumption rate is high (low oxygen level, Cluster 2), genes involved in respiration and catabolic processes are enriched. These two functional clusters are mutually excluded from each other, as functional group in Cluster 2 is significantly under-represented in Cluster 1. These is also a third and smaller gene cluster, whose peak expression is shifted from the other two major groups, containing chemotaxis and motility-related genes (Table 3.1). The averaged expression dynamics of the three groups shows biosynthesis genes (Cluster 1) and respiration genes (Cluster 2) are anti-correlated with each other (Figure 3.2B). Principle component analysis showed two major principle components in the data, with both tracking differently to the oxygen level (Figure 3.2C).
Figure 3.2: Functional gene clusters change expression levels during the metabolic cycle. (A) Heat map of gene expression level from microarray data. The data is centered to the row means. 24 samples in total were collected and corresponding OD and oxygen levels were plotted at the top. (B) Principle component analysis shows two major PCs and their trends are plotted. (C) Averaged gene expression pattern as a function of time for the three clusters defined in (A). The error bar shows the standard error. (Cluster 1: n=1118, Cluster 2: n=841, Cluster 3: n=267). The smoothing curves are generated using loess smoothing function with grey area representing the 95% confidence interval.
<table>
<thead>
<tr>
<th>GO biological process</th>
<th>Reference</th>
<th>Sample</th>
<th>Expected</th>
<th>Fold enrichment</th>
<th>p-value</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>monocarboxylic acid catabolic process (GO:00072329)</td>
<td>79</td>
<td>1</td>
<td>16.81</td>
<td>&lt;0.2</td>
<td>6.84E-04</td>
<td>1</td>
</tr>
<tr>
<td>energy derivation by oxidation of organic compounds (GO:0015980)</td>
<td>94</td>
<td>4</td>
<td>20</td>
<td>0.2</td>
<td>1.29E-02</td>
<td>1</td>
</tr>
<tr>
<td>53 rRNA modification (GO:0015980)</td>
<td>32</td>
<td>23</td>
<td>6.81</td>
<td>3.38</td>
<td>6.52E-04</td>
<td>1</td>
</tr>
<tr>
<td>RNA methylation (GO:0001510)</td>
<td>38</td>
<td>27</td>
<td>8.08</td>
<td>3.34</td>
<td>9.18E-05</td>
<td>1</td>
</tr>
<tr>
<td>rRNA processing (GO:0006364)</td>
<td>43</td>
<td>26</td>
<td>9.15</td>
<td>2.84</td>
<td>2.94E-03</td>
<td>1</td>
</tr>
<tr>
<td>macromolecule methylation (GO:0043414)</td>
<td>44</td>
<td>31</td>
<td>9.36</td>
<td>3.31</td>
<td>1.29E-05</td>
<td>1</td>
</tr>
<tr>
<td>rRNA metabolic process (GO:0016072)</td>
<td>44</td>
<td>27</td>
<td>9.36</td>
<td>2.88</td>
<td>1.47E-03</td>
<td>1</td>
</tr>
<tr>
<td>tRNA modification (GO:0006400)</td>
<td>47</td>
<td>32</td>
<td>10</td>
<td>3.2</td>
<td>1.66E-05</td>
<td>1</td>
</tr>
<tr>
<td>methylation (GO:0032259)</td>
<td>48</td>
<td>32</td>
<td>10.21</td>
<td>3.13</td>
<td>2.66E-05</td>
<td>1</td>
</tr>
<tr>
<td>tRNA processing (GO:0008033)</td>
<td>54</td>
<td>35</td>
<td>11.49</td>
<td>3.05</td>
<td>1.21E-05</td>
<td>1</td>
</tr>
<tr>
<td>ribosome biogenesis (GO:0042254)</td>
<td>64</td>
<td>42</td>
<td>13.61</td>
<td>3.08</td>
<td>3.03E-07</td>
<td>1</td>
</tr>
<tr>
<td>ribonucleoprotein complex biogenesis (GO:0022613)</td>
<td>65</td>
<td>42</td>
<td>13.83</td>
<td>3.04</td>
<td>4.77E-07</td>
<td>1</td>
</tr>
<tr>
<td>translation (GO:0006412)</td>
<td>65</td>
<td>32</td>
<td>13.83</td>
<td>2.31</td>
<td>1.49E-02</td>
<td>1</td>
</tr>
<tr>
<td>peptide biosynthetic process (GO:0043043)</td>
<td>75</td>
<td>39</td>
<td>15.95</td>
<td>2.44</td>
<td>5.09E-04</td>
<td>1</td>
</tr>
<tr>
<td>amide transport (GO:0042886)</td>
<td>77</td>
<td>35</td>
<td>16.38</td>
<td>2.14</td>
<td>3.09E-02</td>
<td>1</td>
</tr>
<tr>
<td>RNA modification (GO:0009451)</td>
<td>80</td>
<td>57</td>
<td>17.02</td>
<td>3.35</td>
<td>7.41E-12</td>
<td>1</td>
</tr>
<tr>
<td>GO biological process</td>
<td>Reference Sample</td>
<td>Fold enrichment</td>
<td>p-value</td>
<td>Cluster</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tRNA metabolic process (GO:0006399)</td>
<td>80</td>
<td>2.59</td>
<td>17.02</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>regulation of biological quality (GO:00065008)</td>
<td>88</td>
<td>18.72</td>
<td>2.3</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>peptide metabolic process (GO:0006518)</td>
<td>42</td>
<td>18.72</td>
<td>2.24</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m-RNA processing (GO:00034470)</td>
<td>92</td>
<td>19.57</td>
<td>3.01</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>response to radiation (GO:0009314)</td>
<td>92</td>
<td>19.57</td>
<td>2.15</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA processing (GO:0006396)</td>
<td>100</td>
<td>21.27</td>
<td>3.01</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monosaccharide transport (GO:0005960)</td>
<td>35</td>
<td>5.67</td>
<td>6.34</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aerobic respiration (GO:0006081)</td>
<td>26</td>
<td>7.46</td>
<td>8.43</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monosaccharide catabolic process (GO:0006365)</td>
<td>26</td>
<td>7.46</td>
<td>8.43</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cellular aldehyde metabolic process (GO:0006081)</td>
<td>22</td>
<td>7.46</td>
<td>8.43</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monosaccharide catabolic process (GO:0006365)</td>
<td>22</td>
<td>7.46</td>
<td>8.43</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>energy derivation by oxidation of organic compounds (GO:0015980)</td>
<td>46</td>
<td>7.46</td>
<td>8.43</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>carbohydrate transport (GO:0008643)</td>
<td>46</td>
<td>7.46</td>
<td>8.43</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>archaeal or bacterial-type flagellum-dependent cell motility (GO:0097588)</td>
<td>92</td>
<td>15.24</td>
<td>2.69</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.1 Continued:

<table>
<thead>
<tr>
<th>GO biological process</th>
<th>Reference</th>
<th>Sample</th>
<th>Expected</th>
<th>Fold enrichment</th>
<th>p-value</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>bacterial-type flagellum-dependent cell motility (GO:0071973)</td>
<td>31</td>
<td>9</td>
<td>1.55</td>
<td>5.81</td>
<td>2.75E-02</td>
<td>3</td>
</tr>
<tr>
<td>cilia or flagellum-dependent cell motility (GO:0001539)</td>
<td>31</td>
<td>9</td>
<td>1.55</td>
<td>5.81</td>
<td>2.75E-02</td>
<td>3</td>
</tr>
<tr>
<td>inorganic anion transport (GO:0015698)</td>
<td>33</td>
<td>9</td>
<td>1.65</td>
<td>5.46</td>
<td>4.44E-02</td>
<td>3</td>
</tr>
<tr>
<td>localization of cell (GO:0051674)</td>
<td>36</td>
<td>10</td>
<td>1.8</td>
<td>5.56</td>
<td>1.46E-02</td>
<td>3</td>
</tr>
<tr>
<td>cell motility (GO:0048870)</td>
<td>36</td>
<td>10</td>
<td>1.8</td>
<td>5.56</td>
<td>1.46E-02</td>
<td>3</td>
</tr>
<tr>
<td>locomotion (GO:0040011)</td>
<td>53</td>
<td>16</td>
<td>2.65</td>
<td>6.05</td>
<td>1.36E-05</td>
<td>3</td>
</tr>
</tbody>
</table>
One hallmark of the yeast metabolic cycle is that the genes oscillating during the cycle also exhibit strong correlations with growth-rate [16]. If the underlying mechanism of the bacterial metabolic cycle is similar to the one in yeast, we should also observe growth-rate dependency of the gene clusters. We thus compared the microarray data to the *E. coli* microarray data from five steady-state glucose-limiting chemostats at different dilution rates (Data not published). Consistent with the yeast metabolic cycle, the expression of genes involved in biosynthesis including ribosome biogenesis and translation are more positively correlated with growth rate (Cluster 1, Figure 3.4). Genes whose expression peak at high oxygen consumption (Cluster 2) are negatively correlated with growth rate (Figure 3.4).

3.4 Discussion

The discovery of yeast metabolic cycle highlights the temporal regulation of metabolic activities. However, it has remained unknown how prevalent the metabolic cycle is in other organisms. Moreover, less is known about the relationship between metabolic cycle and circadian rhythm. Here, we reported a serendipitous finding, showing that *E. coli* can be induced to undergo synchronized metabolic cycling in glucose-limiting chemostats. Characterization of the emergent behavior draws similarities between the two metabolic cycles. For example, both cycles are induced under carbon-limiting condition and the period of the cycle is correlated, but not directly overlapped, with the dilution rate (Growth rate). Analysis of the transcriptome also shows different functional clusters are enriched at different phases during the cycle. Consistent with findings in yeast, genes involved in translation and ribosomes are positively correlated with growth rate, whereas genes involved in oxidative processes are negatively correlated with growth rate. Due to the fact that the gene functions that show oscillation are similar between *E. coli* and yeast, we can interpret the logic of the BMC may also be compartmentalizing reactions that might be antagonist to each other. Moreover, the existence of the metabolic cycle in *E. coli* suggests that it might be a primitive trait in evolution.

Given the growth-rate dependency of the metabolic cycle genes, it raises the question whether metabolic cycle underlies cell growth and also cell cycle. Comparing the oxygen traces between different dilution rates, we find the time cells spent on the OX phase is invariant. However, cells spend more time in the reductive phase when the dilution rate is lower. This shows the BMC is flexible to adapt to the environment. It is also consistent with our observation that OX phase genes are positively correlated with growth rate, given its relative fraction during BMC increases with growth.
Figure 3.3: Metabolite pool sizes change during the period of metabolic cycle. Metabolite levels are normalized to the row means and plotted in log2 scale. The corresponding OD 600 and pO2 values for the 18 samples are shown at the top.
Figure 3.4: Functional gene clusters in metabolic cycle exhibit growth-rate dependency. The black line shows the distribution of growth-rate dependent coefficient of gene expression under carbon-limitation. The enrichment of genes in Cluster 1 and 2 (Defined in Figure 3.2) is plotted as shaded areas. The median and mean of the coefficients are plotted as solid and dashed lines respectively.
rate. To understand the relationship between BMC and bacterial cell cycle, it would require synchronization of bacterial cultures in terms of cell division, which is not as clearly defined and understood as eukaryotic systems. Furthermore, yeast cells can still undergo YMC without cell division cycle.

It still remains unknown the origin of the bacterial metabolic cycle. In our experimental systems we never observed such oscillation under nitrogen- or phosphorus-limiting conditions. We did see oscillation using other buffer-based media (i.e. MOPS, data not shown). However, it is exciting to see the synchronized metabolic cycling can arise in a different organism other than yeast. This synchronous metabolic activity could potentially offer new insights into the coordination of metabolism, gene expression and physiology of the cell. Future studies on the regulatory mechanism of such oscillatory behavior is needed to elucidate the trigger of the emergent dynamics.

3.5 Materials and Methods

3.5.1 Chemostat growth conditions

*E. coli* strain NCM3722 was used in this study. Bacteria were grown Gutnick minimal media (1L of 1x 13.5g K$_2$HPO$_4$, 4.7g KH$_2$PO$_4$, 1g K$_2$SO$_4$, 0.1g MgSO$_4$-7H$_2$O, 9.5 mM NH$_4$Cl). For glucose limitation, the glucose concentration is reduced from 0.4 % to 0.08 %. The following inoculation method has successfully reproduced the oscillatory behavior. 3 mL culture were inoculated into the 300 mL chemostat from an overnight culture grown in the same media. After 8 hours, the pump was turned on to the final dilution rate. The oscillation can be observed after around 8 hours.

3.5.2 RNA extraction

RNA was collected by mixing 0.2x culture volume of 95 % EtOH + 5 % Phenol), flash freeze in liquid nitrogen. The cells were thawed and spun down at 4˚C. The supernatant was discarded and hot phenol method was used to extract RNA. Briefly, 400µL lysis buffer (10mM EDTA, 0.5% SDS, 10mM Tris pH7.5) was added to the cells. 500µL citrate saturated phenol (P4682, Sigma) was added and vortexed. The mixture was then incubated at 65˚C for 10 minute on a thermomixer with constant rigorous shaking. The samples were then chilled on ice for 10 minutes. The mixture was then loaded onto a 2 mL heavy phase lock gel then mixed with 500 µL chloroform and inverted to mix. The tubes were centrifuged for 5 minutes and the aqueous layer was recovered to a new 1.5 mL tube containing 40 µL 3M sodium acetate (pH5.2). 900µL ice-cold 100 % ethanol was added to the samples which were incubated at -20˚C for overnight. RNA was then pelleted
by centrifugation for 30 minutes at 4˚C and washed with ice-cold 70 % EtOH. RNA was resus-
pended with 100µL RNase-free water.

3.5.3 Microarray

Agilent 8x 15k microarray platform with *Escherichia coli* genome array was used in this study. 50 µL of extracted RNA was first digested with DNase (Ambion with inactivation reagent) at 37˚C for 30 minutes. The RNA was cleaned up using Qiagen RNA prep following the vendor’s instruction and eluted with 30 µL RNase-free water. 10 - 40 µg of total RNA was mixed with 1µL of random hexamers (3 mg/mL, Invitrogen) and incubated at 70˚C for 10 minutes and immediately moved onto ice. 10.6 µL of RT mix (6µL 5x FS buffer, 3µL 0.1M DTT, 0.6µL 50x dNTP mix, 1µL RNase-OUT) was added from a master mix while 3µL CyDye and 1 µL Superscript III were added individually. The mixture was incubated at 46˚C for 3 hours in the dark without heated lid. RNA was then degraded by adding 15 µL 0.1M NaOH and incubated at 70˚C for 30 minutes and the solution was neutralized by adding 15 µL of 0.1M HCl. cDNA cleanup was done using Qiagen Qiaquick PCR purification kit and eluted with 30µL water. The yield and labeling efficiency were checked using NanoDrop.

For hybridization, each channel has at least 15 pmol dye and the same amount of cDNA is loaded (no more than 1000ng per channel). 22µL of the mixture was then mixed with 5.5 µL 10x blocking agent and incubated at 95˚C for 5 minutes and RT for 5 minutes. 27.5µL 2x Hi-RPM hybridication buffer was carefully added to avoid any bubbles. 50µL was loaded onto the gasket and hybridization was done at 65˚C for 17 hours at 10 RPM. The array was then washed 3 times with wash buffer 1 for 1 minute and 1 time in wash buffer 2 for 1 minute and acetonitrile for 30 seconds. After the slides were dried, they were loaded into the scanning holder and scanned with the following parameters: 61 x 21.6 mm, resolution 5µm and two scans with 100 % and 5 % laser power. Data extraction was performed using Princeton PUMA database.

3.5.4 GO term analysis

Gene expression data was first organized using hierarchical clustering. The list of genes from the three clusters were picked and submitted to Gene Ontology Consortium for GO term enrichment analysis. Bonferroni correction for multiple testing was used. Due to the hierarchical nature and large variation in size of GO terms, only processes containing 20 to 100 genes were presented in Table 3.1.
3.5.5 Quantification of cellular metabolites

LC-MS was used to measure metabolites. 5 mL of cells were filtered through a 45 mm 0.45 µm nylon membrane and transferred to a petri dish with the cell side facing down into 1 ml cold extraction solvent (-20°C 40:40:20 methanol/acetonitrile/water) to quench metabolic activities immediately. The dishes were incubated at -20°C for at least 15 minutes. Membrane was washed, the cell extract solution was transferred to a microcentrifuge tube, and centrifuged at 13000 rcf for 10 min. The supernatant was transferred to a new microcentrifuge tube. The metabolite extracts were dried under nitrogen flow and reconstituted in HPLC-grade water for LC-MS analysis. Metabolites were measured using stand-alone orbitrap mass spectrometers (ThermoFisher Exactive) operating in negative ion mode with reverse-phase liquid chromatography. Exactive chromatographic separation was achieved on a Synergy Hydro-RP column (100 mm×2 mm, 2.5 µm particle size, Phenomenex) with a flow rate of 200 µL/min. Solvent A was 97:3 H2O/MeOH with 10 mM tributylamine and 15 mM acetic acid; solvent B was methanol. The gradient was 0 min, 5% B; 5 min, 5% B; 7 min, 20% B; 17 min, 95% B; 20 min, 100% B; 24 min, 5% B; 30 min, 5% B.
The crosstalk between protein synthesis and metabolism ensures the homeostasis of cellular energy and responses to the changing environment. This thesis has shown how metabolic limitations can influence several aspects of protein synthesis including ribosome abundance, translational elongation rate, the fraction of active ribosomes, and tRNA modification. Future studies will investigate further on the nutrient signaling between extracellular nutrient abundance and intracellular metabolites such as in the condition of carbon limitation. While dropped glutamine pool size in the cell is known to represent nitrogen limitation, it is unclear how cells perceive carbon or phosphorus limitation. The author observed an elevated NADH/NAD\(^+\) ratio and dropped nucleotide pools under phosphorus limitation. How this metabolic phenotype influences the cellular physiology remains to be characterized. The observation of ribosomal stalling at leucine CTA codon under carbon limitation is robust, and the level of stalling reduces when growth rate increases. This response implies that leucine CTA could potentially be part of the carbon sensor in *E. coli*. It also remains to be answered how the dynamics of tRNA such as its abundance and
the charged ratio change in response to the environment. Defective tRNA modification leads to the undesired reduction of protein production; however, cells might exploit this same mechanism of ribosome stalling to modulate protein level in a physiological context. Finally, ribosome profiling has been a powerful tool in probing ribosome positioning on mRNAs. New tools and measurements related to tRNA and ribosome will provide new mechanistic insights into the regulation of translation activity. These would include but not limit to high-throughput tRNA profiling of the charged ratio and abundance of all tRNA species, measurement of the amino acid and anti-codon pairs, and differentiating ribosomes composed of different subunits or stoichiometry. Together, they will open a new avenue uncovering the dynamic role of tRNA in connecting metabolism and translation, and the diverse ways cells employ to regulate protein synthesis, including the possibility of specialized ribosomes.
Supporting Information for Chapter 1
1. Comparison of RNA-to-protein ratio data with published results

Multiple groups have published the results of RNA-to-protein ratios in a variety of growth conditions, particularly in batch culture using various carbon and nitrogen sources [1-5]. We compared cells grown in our chemostat and batch culture conditions and found that the R/P ratios from batch cultures with different carbon and nitrogen sources overlapped with the R/P ratios from chemostat cultures under carbon and nitrogen limitation (Supplementary Fig. 1a). This result suggests that our findings are not chemostat-specific. We also confirmed that R/P ratio represents a good proxy for ribosome content using bioanalyzer analysis of total RNA. At all growth rates, rRNA fractions remained similar across C/N/P-limitation. We also found that rRNA fractions increased with increasing growth rate (Supplementary Fig. 1e), while tRNA fractions decreased with increasing growth rate (Supplementary Fig. 1f). These measurements showed a similar trend to previously published data regardless of the nutrient limitation [6-8]. Since P-limited cells have both the lowest R/P ratio and the lowest rRNA fraction, we concluded that under phosphorus limitation cells have lower ribosome content than under carbon or nitrogen limitation.

We also determined the R/P ratio of cells grown faster (1.2 h⁻¹) in P-limiting chemostats using defined rich media with reduced phosphate concentration. We found that even at fast growth, P-limited cells still exhibit lower R/P ratio compared to batch culture. This result implies that cells do not only maintain spare protein production capacity at very slow growth rates.

Comparing our results to published data from the same strain of E.coli, we noticed that there is a shift in the R/P ratio absolute values (Supplementary Fig. 1b). Because the change was consistent in all samples and our batch cultures of the same conditions also exhibited higher R/P ratio values, we reasoned this systematic shift might be due to differences in experimental setups, such as differences in the protein standard used.

2. Microscopic model of ribosome dynamics along an mRNA

Ribosome profiling provides positional information of ribosomes occupancy on mRNA transcripts at single nucleotide resolution. In order to investigate ribosomal dynamics on transcripts, we built a microscopic model to estimate the position-dependent behavior of mRNA-bound ribosomes. As shown in the following part of this section, the modeling results lead to a classification of mRNA-bound ribosomes into two groups: “initiating” and “working”, which helps us to build a concise macroscopic model in Section 3.

A. Construction of a representative ribosome-occupancy profile.

We first combine the experimentally obtained ribosome profiling data to obtain an average picture for ribosome occupancy along a “representative” gene with length $N_{aa} = 300$ amino acids [9]. The average ribosome occupancy $O(x)$ for codon position $x = 1:301$, including the stop codon, was constructed via the following steps:
1. All genes that code for proteins longer than 100 aa and have counts larger than 10 transcripts per million (TPM) were selected. This set of genes accounts for more than 75% of total reads.

2. Total counts from the selected genes were normalized to one million (yielding reads per million, RPM) and the ribosome counts at the first and last 50 codons were obtained by summing the RPM at the corresponding position from all the selected genes (result shown in Fig. 2a and Supplementary Fig. 5c) and converted to fraction of total counts. This procedure effectively weights genes in the representative ribosome-count profile according to their level of expression.

3. We used the data from the first 50 codons and the last 49 codons before the stop codon to produce a smooth fit to the experimentally obtained ribosome-count profile. A gap with length 201 was inserted between the first 50 and last 49 codons before the stop codon to produce an mRNA producing 300 amino acids. In order to capture the sharp drop of ribosome counts for the first several codons, together with the slower decrease over the rest of the profile, we used a sum of two decaying exponential functions to fit the data. The experimentally measured ribosome count from the stop codon \( x = 301 \) was directly taken from the data without fitting. The final results are shown in Supplementary Fig. 5c.

B. The microscopic model for bound ribosomes.
In order to decipher the information about ribosome dynamics contained in ribosome occupancy \( O(x) \), we built a microscopic model shown in Supplementary Fig. 5b. In this simplified model, bound ribosomes on codon \( x \) move with a step rate \( r_s(x) \), and a certain fraction \( f_{at}(x) \) abort translation during this transition. At steady state, the in- and out-flux of ribosomes at every codon \( x \) must be balanced:

\[
O(x - 1) \cdot r_s(x - 1) \cdot (1 - f_{at}(x - 1)) = O(x) \cdot r_s(x) .
\]

According to Eq.S 1, under steady state, the relative occupancy of neighboring codons is determined by the step rate and fraction of aborted translation. The profile of \( O(x) \) constrains the values of \( r_s(x) \) and \( f_{at}(x) \). We would like to infer the value of \( r_s(x) \) and \( f_{at}(x) \) given \( O(x) \) within some assumptions. Two possibilities concerning the profile of \( O(x) \) and the corresponding assumptions are listed below:

1. \( O(x) \) decreases with \( x \). This implies an increased step rate \( r_s \) and/or a non-zero fraction of aborted translation \( f_{at} \). For simplicity, we assume a constant fraction \( f_{dec} \) of the decrease of occupancy \( \frac{O(x-1)}{O(x)} - 1 \) can be attributed to an increased step rate \( \frac{r_s(x)}{r_s(x-1)} - 1 \):
2. $O(x)$ does not decrease with $x$. This implies that the step rate $r_s$ does not increase with $x$. Under this condition, the higher $f_{at}$, the more $r_s$ decreases with $x$. According to previous research, the translation elongation rate of ribosomes generally increases from 5' to 3'[10, 11]. Therefore, we assume $f_{at} = 0$ under this condition.

With the former assumptions, we can derive $r_s(x - 1)$ and $f_{at}(x - 1)$ given $r_s(x)$:

$$f_{dec} = \frac{r_s(x)}{r_s(x - 1)} - 1 \quad \text{Eq. S 2}$$

$$r_s(x - 1) = \begin{cases} 
\frac{r_s(x)}{f_{dec} \cdot \left(\frac{O(x - 1)}{O(x)} - 1\right) + 1}, & \text{if } \frac{O(x - 1)}{O(x)} > 1 \\
\frac{r_s(x)}{O(x)} \cdot \frac{O(x)}{O(x - 1)}, & \text{if } \frac{O(x - 1)}{O(x)} \leq 1
\end{cases} \quad \text{Eq. S 3}$$

$$f_{at}(x - 1) = \begin{cases} 
(1 - f_{dec}) \cdot \left(1 - \frac{O(x)}{O(x - 1)}\right), & \text{if } \frac{O(x - 1)}{O(x)} > 1 \\
0, & \text{if } \frac{O(x - 1)}{O(x)} \leq 1
\end{cases} \quad \text{Eq. S 4}$$

C. Calculation of step rate $r_s(x)$ and fraction of aborted translation $f_{at}(x)$.

By applying Eq. S 3 - Eq. S 4 to the ribosome occupancy $O(x)$ obtained experimentally in Section 2A, we can calculate $r_s$ and $f_{at}$ from $x = 300$ to $x = 1$, for assumed values of $f_{dec}$ and the step rate at the last coding codon $r_s(300)$. The existence of aborted translation has been suggested by previous studies [10, 12], however the quantitative fraction of aborted translation under different nutrient-limitations is difficult to quantify. Given the uncertainty of the aborted translation rate, we evenly sampled values of $f_{dec}$ from 0 to 1 and values of $r_s(300)$ from 5 to 20 codon/sec. Each combination of $f_{dec}$ and $r_s(300)$ gives a solution for $r_s(x)$ and $f_{at}(x)$ from $x = 300$ to $x = 1$. The results are shown in Supplementary Fig. 5d-e: regardless of $f_{dec}$ and $r_s(300)$, aborted translation predominantly happens within the first 10 codons, and the step rate stays relatively flat after the first 10 codons.

D. Classification of ribosomes and mRNAs.

Guided by the above results, we classify mRNA-bound ribosomes into two groups:

1. Initiating ribosomes (symbolized by $R_i$): ribosomes located within codon 1 to 10 are classified as “initiating”. Given the results in Section 2C, we assume that aborted translation, if it exists at all, only happens to this group of ribosomes. As one ribosome roughly occupies 10 codons on an mRNA [13, 14], the binding
of one ribosome at the initiating region excludes the association of another ribosome. Thereby mRNAs are divided into two groups: the freely initiable mRNAs (symbolized by \( M_f \)), and the un-initiable mRNAs with one ribosome occupying the 1-10 codon region (symbolized by \( M_b \)). By definition, the number of un-initiable mRNAs is equal to the number of initiating ribosomes.

2. Working ribosomes (symbolized by \( R_w \)): ribosomes located after the 10-th codon are classified as “working”. They contribute to the increase of biomass. We assume that these ribosomes elongate at a constant rate \( k_{el} \) without terminating translation prematurely. Under this assumption, there is no aborted translation in transition from the last coding codon \((x = 300)\) to the stop codon \((x = 301)\), and the number of ribosomes transiting from the last coding codon to the stop codon is equal to the number of ribosomes leaving stop codon in a steady state. Therefore, the ribosome occupancy at stop codon does not influence the overall protein production rate. In addition, ribosome occupancies at the stop codon are similar across different nutrient-limitations (Fig. 2a), and only take about 2% of total ribosomes. Therefore, we count ribosomes at stop codon also as “working ribosomes” instead of classifying them into another group.

3. Macroscopic model of ribosome dynamics among different states

A. Dynamics of ribosomes and mRNAs between different states.

In Section 2, we classified the mRNA-bound ribosomes into “initiating” and “working” states. This result allowed us to construct a macroscopic model that provides a concise picture of ribosomal dynamics. The transition of ribosomes between three states is shown in Fig. 2b. As ribosomal subunits and free 70S ribosomes may convert between each other, they are treated as one group, named unbound ribosomes (symbolized by \( R_u \)). A sub-population of the unbound ribosomes can associate with freely initiable mRNAs (\( M_f \)) with an effective rate constant \( k_f \). This reaction forms a complex composed of an initiating ribosome \( (R_i) \) and an un-initiable mRNA \( (M_b) \). The initiating ribosomes either become unbound ribosomes through aborted translation (with rate constant \( k_r \)), or proceed to the 11-th codon to become working ribosomes \( (R_w) \) (with rate constant \( k_p \)). Both processes can release the mRNA back to the freely initiable state. The working ribosomes elongate with constant rate \( k_{el} \). At the last coding codon, a fraction of bound ribosomes \((\sim 1/N_{aa})\) enters stop codon with rate \( k_{el} \), and the same number of ribosomes completes translation and releases mRNA and become unbound ribosomes.

The non-redundant kinetic ordinary differential equations for the number of ribosomes and mRNAs in each state are:

\[
\frac{dR_i}{dt} = k_f \cdot M_f \cdot R_u - k_r \cdot R_i - k_p \cdot R_i, \quad \text{Eq.S 5}
\]

\[
\frac{dR_w}{dt} = k_p \cdot R_i - \frac{k_{el}}{N_{aa}} \cdot R_w, \quad \text{Eq.S 6}
\]

\[
\frac{dM_f}{dt} = -k_f \cdot M_f \cdot R_u + k_p \cdot R_i + k_r \cdot R_i. \quad \text{Eq.S 7}
\]
Under steady-state growth conditions, the fluxes between ribosomal states should be balanced, and the right-hand sides of Eq.S 5 - Eq.S 7 should all be equal to zero (the production and degradation/dilution by growth of new ribosomes and mRNAs occurs at a negligible rate compared to their recycling). The total number of ribosomes ($R_t$) and the total number of mRNAs ($M_t$) are constant parameters for this model:

$$R_u + R_i + R_w = R_t,$$

**Eq.S 8**

$$M_b + M_f = M_t,$$

**Eq.S 9**

and by the definition in Section 2D,

$$M_b = R_i.$$

**Eq.S 10**

B. Estimation of $R_t$, $M_t$, $k_{el}$, and $k_p$ from experimental measurements.

We are interested in the biological mechanisms that lead to distinct ribosomal dynamics under different nutrient conditions. Possible regulatory processes can be represented by parameters in the macroscopic model, including the total number of ribosomes and mRNAs and the rates for ribosomes to initiate, dissociate, and elongate. Some of these parameters, specifically $R_t$, $M_t$, $k_{el}$, and $k_p$, can be directly estimated from experimental measurements as follows:

1. The total number of ribosomes ($R_t$): The total mass of rRNA can be obtained by multiplying the RNA-to-protein ratio ($RPR$), protein mass in a cell ($P_m$), and the fraction of RNA as rRNA ($f_r$). Given the mass of the rRNA in a ribosome as $m_r$, we estimated the total number of ribosomes in a cell:

$$R_t = P_m \cdot RPR \cdot \frac{f_r}{m_r}.$$

**Eq.S 11**

2. The total number of mRNAs in a cell ($M_t$): The method is similar to the calculation of $R_t$ in Eq.S 11. Given the fraction of RNA as mRNA as $f_m$, and the average weight of a nucleotide as $m_{nuc}$, we have:

$$M_t = P_m \cdot RPR \cdot \frac{f_m}{N_{aa} \cdot m_{nuc} \cdot 3}.$$

**Eq.S 12**

3. The protein synthesis rate ($j_p$): The amount of newly synthesized proteins in a cell per second is defined as $j_p$ (aa/sec). $j_p$ can be calculated by two ways: One, the growth rate $\mu$ (h$^{-1}$) can be calculated as the relative rate of protein mass accumulation: $\mu = \frac{j_p}{P_m/m_{aa}}$, where $m_{aa}$ is the average mass of amino acid (g). Therefore, $j_p$ is linearly proportional to the growth rate:
\[ J_P = \mu \cdot \left( \frac{P_m}{3600 \cdot m_{aa}} \right) \]  

Eq.S 13

\[ J_P \] can also be calculated from the contribution of working ribosomes to the growth of the total protein pool:

\[ J_P = R_w \cdot k_{el}. \]  

Eq.S 14

Combining Eq.S 13 and Eq.S 14, the growth rate \( \mu \) is determined by three factors: the total number of ribosomes \( (R_t) \), the fraction of working ribosomes \( (\phi_{RW}) \), and the average elongation rate of the working ribosomes \( (k_{el}) \):

\[ \mu = R_t \cdot \phi_{RW} \cdot k_{el} \cdot \left( \frac{3600 \cdot m_{aa}}{P_m} \right). \]  

Eq.S 15

The contribution of these three factors to the growth rate is illustrated by Table 1. The surface in the space of these three factors that corresponds to growth rate \( \mu = 0.1 \text{ h}^{-1} \) is shown in Fig. 2e and Supplementary Fig. 6f, with the estimated values of \( R_t, \phi_{RW}, \) and \( k_{el} \) under various conditions indicated on this surface.

Substituting \( R_t \) in Eq.S 15 by Eq.S 11, the relationship between the four experimentally measured values – growth rate \( (\mu) \), RNA-to-protein ratio \( (RPR) \), average elongation rate \( (k_{el}) \), and fraction of working ribosomes \( (\phi_{RW}) \) – is given by:

\[ \mu = RPR \cdot \phi_{RW} \cdot k_{el} \cdot \left( \frac{3600 \cdot m_{aa} \cdot f_r}{m_r} \right). \]  

Eq.S 16

Eq.S 16 can be used to estimate growth rate as shown in Table 1. Using different definitions of \( \phi_{RW} \), we show that it can lead to different estimations of growth rate. The \( \phi_{RW} \) we used for modeling is from the following Section 3B4. While growth rate, RNA-to-protein ratio and fraction of working ribosomes are global measurements, the elongation rate was measured using the \textit{lacZ} induction assay. For the rRNA fraction \( (f_r) \) we used the experimentally measured values from Supplementary Fig. 1e.

In order to obtain a global estimate of the elongation rate \( k_{el} \), we derived \( k_{el} \) from the measurements of growth rate, RNA-to-protein ratio, and the fraction of working ribosomes by the following steps 4-5.

4. Quantifying the fraction of different ribosomal species: The fraction of unbound ribosomes is the sum of the free 70S fraction and the subunit fraction \( (\phi_{RU} = \phi_{R70Sf} + \phi_{Rs}) \). The fraction of initiating ribosome is
calculated by multiplying the fraction of bound ribosomes by the sum of ribosome occupancies at the first 10 codon ($\phi_{Ri} = \sum_{x=1}^{10} O(x) \cdot (\phi_{R70Sb} + \phi_{Rpoly})$). The fraction of working ribosomes is calculated by multiplying the fraction of bound ribosomes with the summation of ribosome occupancy at the 11-th to 301-th codon ($\phi_i = \sum_{x=11}^{301} O(x) \cdot (\phi_{R70Sb} + \phi_{Rpoly})$).

5. Estimating the average working ribosome elongation rate $k_{el}$: From Eq.S 16 and the definition of $\phi_{RW}$, $k_{el}$ can be calculated by measurements of RPR and $\phi_{RW}$:

$$k_{el} = \frac{\mu}{RPR \cdot \phi_{RW} \cdot 3600 \cdot m_{aa} \cdot \frac{f_r}{m_r}}.$$  

**Eq.S 17**

The resulting $k_{el}$ is within the error range of the elongation rates measured from the $lacZ$ induction assay (Supplementary Fig. 6a).

6. The rate constant $k_p$ for ribosomes to proceed from initiating to working:

According to Eq.S 6, at steady state,

$$k_p = \frac{k_{el}}{N_{aa}} \cdot \frac{\phi_{RW}}{\phi_{Ri}}.$$  

**Eq.S 18**

C. Fraction of working ribosomes ($\phi_{RW}$) as a function of kinetic parameters.

Among the three factors in Eq.S 15 that determine growth rate, $R_t$ and $k_{el}$ are parameters of the macroscopic model. In contrast, the fraction of working ribosomes, $\phi_{RW}$, is the outcome of dynamic processes described by the parameters in the macroscopic model. We would like to infer the relationship between $\phi_{RW}$ and these parameters, in order to identify the mechanisms that lead to the different observed $\phi_{RW}$ under different conditions.

The parameters $k_r$ and $k_f$ cannot be obtained experimentally. Nevertheless, we will show in the following part of this section that these two parameters work in combination to influence the value of $\phi_{RW}$, and the combined parameter can be obtained from experimental measurements.

According to Eq.S 5-Eq.S 10, in steady state, we obtained an expression for $\phi_{RW}$:

$$\phi_{RW} = \frac{F}{2} \cdot \left( \frac{1}{S \cdot (F + 1) + \frac{M_t}{R_t}} - \left( \frac{1}{S \cdot (F + 1) + \frac{M_t}{R_t}} \right)^2 - 4 \cdot \frac{M_t}{R_t} \cdot \frac{1}{F + 1} \right),$$  

with

$$F = N_{aa} \cdot \frac{k_p}{k_{el}}.$$  

**Eq.S 19**
\[ K_M = \frac{k_r + k_p}{k_f R_t}, \]
\[ S = \frac{K_M + R_t}{M_t}. \]

Eq.S 19 depends on three combined parameters: \( \frac{M_t}{R_t} \) and \( F, S \). According to Eq.S 11 and Eq.S 12, \( \frac{M_t}{R_t} = \frac{f_m m_r}{N_{aa} m_{nuc}^3 f_r} \) is a constant across different conditions under our assumptions. Physically, the combined parameter \( F \) can be interpreted as the “relative proceeding rate”, reflecting the rate for ribosomes to proceed from an initiating to a working state relative to the average elongation rate. If there were no regulation at this step, \( F \) would be constant across different conditions regardless of elongation rates. However, if there is nutrient-specific regulation on the step rate or aborted translation rate within the first 10 codons, \( F \) will have different values for different conditions. The value of \( F \) can be calculated from the ribosomal profiling data according to Eq.S 18: 
\[ F = \frac{\sum_{x=11}^{30} O(x)}{\sum_{x=1}^{10} O(x)}. \]
Supplementary Fig. 6b shows the values of \( F \) under C-, N-, P-limitations (\( \mu = 0.1 \text{ h}^{-1} \)) and Minimal condition (\( \mu = 0.9 \text{ h}^{-1} \)) for WT and the relA mutant.

The lumped parameter \( S \) can be interpreted as a “saturation parameter”: According to Eq.S 5-Eq.S 10 and Eq.S 13-Eq.S 14, the growth rate can be expressed as a function of \( R_u \), mimicking the form of Michaelis-Menten:

\[ \mu(R_u) = \frac{3600 \cdot N_{aa} \cdot m_{aa}}{P_m} \cdot M_t \cdot k_p \cdot \frac{R_u}{K_m + R_u}. \]

Eq.S 20

In analogy to enzymatic reactions, in our model, ribosome subunits can be viewed as the substrates, mRNA as the enzyme, and working ribosomes as the product, and therefore \( K_m \) is interpreted as the half-saturation concentration of unbound ribosomes. Therefore, \( S = \frac{R_t}{K_m + R_t} \) positively correlates with the degree of saturation for ribosomes in translation: A value of \( S \) near zero implies a large fraction of mRNAs are initiable, waiting for ribosomes, while \( S \) approaching one implies most mRNAs are occupied in the initiation region and increasing in ribosome number cannot substantially boost growth rate. The value of \( S \) can be calculated by quantifying the ribosomal species according to Eq.S 5, Eq.S 11, and Eq.S 12: 
\[ S = 1/(1 + \left( \frac{f_m m_r}{N_{aa} m_{nuc}^3 f_r} - \phi_{RI} \right) \cdot \frac{\phi_{RU}}{\phi_{RI}}). \]
Supplementary Fig. 6c shows the values of \( S \) under C-, N-, P-limitations (\( \mu = 0.1 \text{ h}^{-1} \)) and Minimal condition (\( \mu = 0.9 \text{ h}^{-1} \)) for wild type and the relA mutant.

According to Eq.S 19, \( \varphi_{RW} \) is an increasing function of both \( F \) and \( S \). In wild type, the relative proceeding rate \( F \) and the saturation parameter \( S \) both increase from C-, N- P-limitation to Minimal condition, leading to an increasing \( \varphi_{RW} \). The relA mutant under P-limitation has a very high relative proceeding rate.
consistent with the known biological role of \( \text{relA} \). By contrast, the loss of \( \text{relA} \) does not affect the saturation parameter.

**D. Growth rate \( \mu \) as a function of total ribosome number \( R_t \) and total mRNA number \( M_t \).**

Given our detailed estimation of the parameters associated with translation under different nutrient conditions, we are interested in how the total number of ribosomes may influence the protein production rate, and whether ribosome number is the limiting factor for cell growth.

Combining Eq.S 15 and Eq.S 19, yields the growth rate \( \mu \) as a function of \( R_t \), and other kinetic parameters of the macroscopic model:

\[
\mu(R_t) = \left( \frac{3600 \cdot m_{aa}}{P_{m}} \right) \cdot \frac{k_p}{2} \cdot N_{aa} \times \left( A \cdot K_M + A \cdot R_t + M_t \right) - \left( (A \cdot K_M + A \cdot R_t + M_t)^2 - 4 \cdot M_t \cdot R_t \cdot A \right)\frac{1}{2},
\]

with

\[
A = \frac{1}{F + 1} = \frac{1}{\frac{N_{aa} \cdot k_p}{K_{el}} + 1}.
\]

This function resembles the shape of a Michaelis-Menten function of both \( R_t \) and \( M_t \). Intuitively, \( \mu = 0 \) when \( R_t \) or \( M_t \) equal to zero. If \( R_t \) increases while \( M_t \) is held constant, \( \mu \) monotonically increases and saturates at

\[
\mu(\infty) = \left( \frac{3600 \cdot m_{aa}}{P_{m}} \right) \cdot \frac{k_p}{2} \cdot N_{aa} \cdot M_t.
\]

The half-saturation value of \( R_t \) is also linearly increasing with \( M_t \):

\[
R_{t,1/2} = \frac{1}{2} \cdot A \cdot M_t + K_M.
\]

Supplementary Fig. 6d-e show the functions \( \mu(R_t) \) and the estimated values of \( R_t \) under different nutrient conditions while \( M_t \) is held constant.

As shown in Supplementary Fig. 6d-e, in all conditions, the estimated numbers of ribosomes are much smaller than the half-saturation value, falling into the highly linear region, it supports that ribosomes are the limiting factors for cell growth. In future, it will be interesting to directly probe the degree of saturation for ribosomes experimentally.

**E. Predicting the growth rate after nutrient upshift to rich medium.**

As mentioned in the main text, there appear to be diverse strategies that achieve the same growth rate under different nutrient limitations, but these strategies may lead to different outcomes when the environment changes. One possibility is that having extra ribosomes at slow growth rates can enable a faster recovery when nutrients become abundant again. Therefore, we expanded our...
macroscopic model beyond steady state to predict growth dynamics upon nutrient upshift.

The rate of cell growth positively correlates with the rate of protein synthesis. Among newly synthesized proteins, a certain fraction of mass is allocated to ribosomal proteins ($\Psi_t^\ast(t)$). The steady-state value of this mass fraction ($\Psi_t^\ast$) under a given condition ($i$) can be calculated from the estimated number of total ribosomes:

$$
\Psi_{R,i}^\ast = R_t \cdot N_{Ra,a} \cdot \frac{m_{aa}}{P_m}
$$

Eq.S 24

$N_{Ra,a}$ is the number of amino acids in a ribosome, and $m_{aa}$ is the average mass of an amino acid in E. coli protein. The time-dependent growth rate $g(t)$ of a bacterial population can be defined in terms of the increase in total cell volume:

$$
g(t) = \frac{V'(t)}{V(t)}.
$$

Eq.S 25

Assuming the concentration of protein ($C_p$) remains constant, the rate of increase of total volume will be linearly related to the rate of increase of protein mass, which is proportional to the concentration of working ribosomes $n_b(t)$ and the translation completion rate $k_{el}/N_{aa}$:

$$
V'(t) = V(t) \cdot k_{el} \cdot n_b(t) \cdot \frac{m_{aa}}{C_p}.
$$

Eq.S 26

Based on Eq.S 25 and Eq.S 26, the growth rate can be expressed as:

$$
g(t) = k_{el} \cdot n_b(t) \cdot \frac{m_{aa}}{C_p}.
$$

Eq.S 27

However, during the process of reaching steady state, $\Phi(t)$ might be regulated, e.g. to accelerate cell growth, and therefore could be a time-dependent function.

Applying the chain rule of differentiation to intracellular concentrations,

$$
\frac{d(x)}{dt} = \frac{1}{V} \cdot \frac{dx}{dt} - \frac{x}{V} \cdot \left(\frac{1}{V} \cdot \frac{dv}{dt}\right),
$$

the rates of change in the concentration of different ribosome species can be expressed as:

$$
\frac{dr_i}{dt} = k_f \cdot M_f \cdot (r_i(t) - r_i(t) - r_w(t)) - k_r \cdot r_i(t) - k_p \cdot r_i(t) - g(t)
$$

Eq.S 28

$$
\cdot n_b(t),
$$

Eq.S 29

$$
\frac{dr_w}{dt} = k_p \cdot r_i(t) - k_{el} \cdot N_{aa} \cdot r_w(t) - g(t) \cdot r_w(t).
$$

Eq.S 30
We are interested in the growth dynamics after switching cells from C-, N-, or P-limitation to rich medium. We therefore solved Eq. S28-Eq. S30 for parameter values obtained under rich condition. The initial states of \( r_t, n, \) and \( r_w \) were taken as their steady-state values (\( r_t^*, n^*, \) and \( r_w^* \)) under C-, N-, and P-limitations, respectively.

The fraction of newly synthesized ribosomal proteins \( \psi_R(t) \) reflects the regulation of protein allocation by the cell. Eventually this value needs to reach the steady-state value \( \psi_R^* \) in rich media. Nevertheless, during the transition period, different strategies of regulating \( \psi_R(t) \) yield different growth dynamics. In our model, we tested three different control strategies for \( \psi_R(t) \): Bang-Bang control

\[
\psi_R^*(t) = \begin{cases} 
1 & \text{if } r_t < r_{\text{rich}}^* \\
\psi_{R,\text{rich}}^* & \text{if } r_t = r_{\text{rich}}^* \\
0 & \text{if } r_t > r_{\text{rich}}^* 
\end{cases}
\]

smooth control (\( \psi_R^*(t) = 2 \cdot \psi_{R,\text{rich}}^*/(r_t/r_{\text{rich}}^* + 1) \)), and steady-state control (\( \psi_R^*(t) = \psi_{R,\text{rich}}^* \)). “Bang-bang control” has been discovered in multiple domains of biology, and it has been shown that Bang-Bang control of regulation can maximize the accumulated increase of cell volume [7-9]. In our model, the control mechanism in Eq. S33 allocates all resources into synthesis of ribosomes if the ribosome concentration is less than \( r_{\text{rich}}^* \), and shuts off the ribosomal protein production if the cellular concentration of ribosomes exceeds \( \psi_{R,\text{rich}}^* \). Interestingly, we found that a Bang-Bang control strategy gave the prediction closest to our experimentally observed dynamics for nutrient-upshift (Supplementary Fig. 9).

The other two possible mechanisms give rise to qualitatively similar dynamics for nutrient-upshift, where the C/N-limited cells recovers faster than P-limited cells in Rich condition; however, the predicted cell growth curves are slower than the experimental observation.
Table S1. Definitions and values of parameters used in ribosome dynamics models.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$RPR$</td>
<td>Mass ratio of total RNA to total Protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT C-limit 0.16</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT N-limit 0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT P-limit 0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA C-limit 0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA N-limit 0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA P-limit 0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT Min 0.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT Rich 0.56</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>Steady-state growth rate (h$^{-1}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT C-limit 0.09</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT N-limit 0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT P-limit 0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA C-limit 0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA N-limit 0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA P-limit 0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Min 0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rich 1.6</td>
<td></td>
</tr>
<tr>
<td>$\phi_{Ru}$</td>
<td>Fraction of ribosomes that are not bound to mRNA, may abort translation prematurely</td>
<td>WT C-limit 0.69</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT N-limit 0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT P-limit 0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA C-limit 0.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA N-limit 0.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA P-limit 0.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT Min 0.34</td>
<td></td>
</tr>
<tr>
<td>$\phi_{Rw}$</td>
<td>Fraction of ribosomes locating after the first 10 codon of mRNA, contribute to the protein production in constant elongation rate</td>
<td>WT C-limit 0.29</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT N-limit 0.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT P-limit 0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA C-limit 0.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA N-limit 0.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relA P-limit 0.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT Min 0.63</td>
<td></td>
</tr>
<tr>
<td>$f_r$</td>
<td>Fractional mass of rRNA among total RNA</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT C-limit 0.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT N-limit 0.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT P-limit 0.51</td>
<td></td>
</tr>
<tr>
<td>$f_m$</td>
<td>Fractional mass of mRNA</td>
<td>0.02</td>
<td>[15]</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Value</td>
<td>Reference</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>$V_c$</td>
<td>Cell volume (m$^3$)</td>
<td>$10^{-18}$</td>
<td>[1]</td>
</tr>
<tr>
<td>$C_p$</td>
<td>Concentration of proteins (g/m$^3$)</td>
<td>$2.4 \times 10^5$</td>
<td>[16]</td>
</tr>
<tr>
<td>$m_r$</td>
<td>Mass of the rRNA component in a ribosome (g)</td>
<td>$2.8 \times 10^{-18}$</td>
<td>[17]</td>
</tr>
<tr>
<td>$m_{nuc}$</td>
<td>Average mass of a nucleotide in RNA in <em>E. coli</em> (g)</td>
<td>$5.4 \times 10^{-22}$</td>
<td>[18]</td>
</tr>
<tr>
<td>$m_{aa}$</td>
<td>Average mass of an amino acid in <em>E. coli</em> protein (g)</td>
<td>$1.8 \times 10^{-22}$</td>
<td>[19]</td>
</tr>
<tr>
<td>$N_{Raa}$</td>
<td>Number of amino acids in the ribosome</td>
<td>7459</td>
<td>[20]</td>
</tr>
<tr>
<td>$N_{aa}$</td>
<td>Average mRNA length (aa)</td>
<td>300</td>
<td>[9]</td>
</tr>
<tr>
<td>$P_m$</td>
<td>Total protein mass in a cell (g)</td>
<td>$C_p \cdot V_c$</td>
<td></td>
</tr>
<tr>
<td>$O(x)$</td>
<td>Average occupancy of ribosomes at the <em>x</em>-th codon</td>
<td>Obtained by fitting the data from ribosome profiling</td>
<td></td>
</tr>
<tr>
<td>$r_s(x)$</td>
<td>Step rate of ribosomes from codon <em>x</em> to the next codon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_{at}(x)$</td>
<td>Fraction of ribosomes that abort translation during transition from codon <em>x</em> to the next codon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_t$</td>
<td>Total number of ribosomes per cell</td>
<td>$P_m \cdot RPR \cdot \frac{f_r}{m_r}$</td>
<td></td>
</tr>
<tr>
<td>$R_i$</td>
<td>Initiating ribosomes: ribosomes located within the first 10 codons</td>
<td>$R_t \cdot \varnothing_{Ri}$</td>
<td></td>
</tr>
<tr>
<td>$R_w$</td>
<td>Working ribosomes: ribosomes located within codon 11 to 301.</td>
<td>$R_t \cdot \varnothing_{Rw}$</td>
<td></td>
</tr>
<tr>
<td>$R_u$</td>
<td>Unbound ribosomes, including 70S free ribosomes and subunits</td>
<td>$R_t \cdot \varnothing_{Ru}$</td>
<td></td>
</tr>
<tr>
<td>$M_t$</td>
<td>Total number of mRNAs per cell</td>
<td>$P_m \cdot RPR \cdot \frac{f_m}{N_{aa} \cdot m_{nuc} \cdot 3}$</td>
<td></td>
</tr>
<tr>
<td>$M_b$</td>
<td>Un-initiable mRNAs, with one ribosome bound within the first 10 codons</td>
<td>$R_i$</td>
<td></td>
</tr>
<tr>
<td>$M_f$</td>
<td>Freely initiable mRNAs, with no ribosome bound within the first 10 codons</td>
<td>$M_t - M_b$</td>
<td></td>
</tr>
<tr>
<td>$k_f$</td>
<td>Effective rate constant for unbound ribosomes and free mRNAs to initiate translation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_r$</td>
<td>Rate constant for initiating ribosomes to abort translation (1/s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{el}$</td>
<td>Elongation rate for working ribosomes (aa/s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_p$</td>
<td>Rate constant for initiating ribosomes to transition into working ribosomes (1/s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$J_p$</td>
<td>Total rate of protein synthesis (aa/s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F$</td>
<td>Relative proceeding rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S$</td>
<td>Saturation parameter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V(t)$</td>
<td>Volume of the population at time $t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$g(t)$</td>
<td>Growth rate of the population at time $t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_R(t)$</td>
<td>Fraction of newly-synthesized proteins allocated to ribosomal proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_t$</td>
<td>Concentration inside cells of total number of ribosomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_i$</td>
<td>Concentration inside cells of initiating ribosomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_w$</td>
<td>Concentration inside cells of working ribosomes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Equations

- Rate constant for initiating ribosomes to abort translation:
  
  \[
  k_r = \frac{\mu}{RPR \cdot \phi_{Rw} \cdot 3600 \cdot m_{aa} \cdot \frac{f_r}{m_r}}
  \]

- Elongation rate for working ribosomes:
  
  \[
  k_{el} = \frac{\phi_{Rw}}{N_{aa} \cdot \phi_{Ri}}
  \]

- Total rate of protein synthesis:
  
  \[
  J_p = \mu \cdot \left( \frac{P_m}{3600 \cdot m_{aa}} \right)
  \]

- Relative proceeding rate:
  
  \[
  F = \frac{N_{aa} \cdot k_p}{k_{el}}
  \]

- Saturation parameter:
  
  \[
  S = \frac{k_r + k_p}{k_f} + \frac{R_t}{R_t}
  \]
Supplementary References


Characterization of *E. coli* physiology across different nutrient conditions and growth rates

*E. coli* exhibit amazing adaptation to grow in a range of rates and nutrient limiting conditions. Several aspects of its physiology are tuned for its survival. Among them the translational activity is one of the most important aspect as growth requires new protein synthesis. To understand how cells tune its translational activity, 12 conditions were assayed for different genetic backgrounds (wild-type, ΔleuB, ΔrplA), nutrient limitations (Carbon, Nitrogen, Phosphorus, Leucine) and growth rates (0.1 and 0.6 h⁻¹ in chemostat, 0.9 and 1.8 h⁻¹ in batch). ΔleuB mutant is auxotrophic for leucine.ΔrplA mutant lacks a ribosomal subunit which results in slow translocation. These diverse conditions can probe the phenotype space more widely to distinguish effects from different sources.

Characterization of translational elongation rate showed that at slow growth (0.1 h⁻¹), leucine limitation exerts strong inhibition on elongation rate. ΔrplA at slow growth has no obvious defects
compared to wild type under the same nitrogen limitation. Carbon- and phosphorus-limited cells have equally faster elongation rates, most likely due to the unlimited supply of amino acids. At faster growth (0.6 h⁻¹), translational elongation rates increase in all conditions. The phenotype of ΔrplA is now dominated by its slow translation, resulting in higher expression of ribosomes shown in Fig. B.1(c). In general, the fraction of assembled ribosomes that are bound to mRNA increases with growth rate. Carbon limitation however has the smallest fraction of all. Combining the measurements done in Fig. B.1(a-c) and fraction of RNA that is rRNA (Fig. B.2a) gives the growth rate prediction in Fig. B.1(d). While at faster growth (0.6 h⁻¹) and defined rich minimal media the predicted growth rates are slower than actual ones, most conditions are consistent between the predicted and the actual. The discrepancy is most likely from rRNA fraction in total RNA given the noisy nature of bioanalyzer. However, other factors that are not considered in our model can lead to the underestimate as well.

In the cell, RNA-to-protein ratio can be an approximate for the fraction of proteome dedicated to ribosome production, thus ribosome efficiency. Other ratios might offer a different insight for the physiology. Fig. B.1e and f show the RNA-to-DNA and Protein-to-DNA ratios. Given most RNA is rRNA, R/D ratio can be considered the transcriptional efficiency from DNA to rRNA on the rRNA loci. There are seven loci of rRNA in E. coli genome. R/D ratio is highly growth rate dependent but not nutrient dependent, which is very different from R/P ratio. This tight growth rate dependence might imply that RNA polymerase activity is balanced to growth rate despite gene dosage difference or metabolite concentrations. This different behavior from R/P ratio is worth taking a further investigation. The regulation might imply the critical control point cells use to regulate the ribosome abundance.

Fig. B.2 shows similar trend of RNA composition from previous results. While tRNA and rRNA are transcribed together, they do exhibit a growth rate dependence. The effect of different genetic background and nutrients are small. The composition also shows convergence when growth rate increases.

Fig. B.3 shows parameter of cell size. Carbon limiting conditions are closest to batch systems studied before and indeed our results show increase cell length and width when growth rate increases. However, nutrient limitations have a strong effect on cell size through a combination of length and width. Interestingly, cell length has a tighter distribution of the variation in length, but not in width (Fig. B.3a, lower panel). This might suggest that regulation of cell length might be more important than of cell width. The surface area-to-volume(SAV) ratio shows carbon limited cells at slow growth has the largest SAV ratio, while it is not clear what physical parameter cells
Figure B.1: Growth rate prediction using physiological measurements agree with actual ones across a wide range of growth conditions. (a) Translational elongation rates measured by lacZ induction assay. (b) The fraction of bound ribosomes (bound 70S monosome + polysomes). (c) RNA-to-protein ratio. (d) Comparison of actual and predicted growth rates. The dashed line has a slope equal to 1. (e) RNA-to-DNA ratio. (f) DNA-to-protein ratio. All data points are averages from three biologically independent samples. Error bars are s.e.m. In (d) all individual samples are plotted. C: glucose-limitation; N: ammonia-limitation; P: phosphorus-limitation; LB: ΔleuB mutant in leucine-limitation; RA: ΔrplA mutant in ammonia limitation; M: Batch glucose minimal media; R: Batch defined rich media. All samples were grown in 40 mM MOPS media.
Figure B.2: Fraction of different RNA species change across growth conditions. (a) Fraction of rRNA in total RNA. (b) Fraction of tRNA in total RNA. All data points are averages from three biologically independent samples. Error bars are s.e.m. C: glucose-limitation; N: ammonia-limitation; P: phosphorus-limitation; LB: leuB mutant in leucine-limitation; RA: rplA mutant in ammonia limitation; M: Batch glucose minimal media; R: Batch defined rich media. All samples were grown in 40 mM MOPS media.

are sensing to regulate their sizes.
Figure B.3: Cell size parameters are dependent on more than the growth rate. (a) The average and standard deviation of normalized values for the cell length and width across conditions. (b) Averaged 3D cell volumes inferred from cell widths and lengths. (c) Average surface area-to-volume ratio. All data points are from more than 500 cells in one biological sample. C: glucose-limitation; N: ammonia-limitation; P: phosphorus-limitation; LB: leuB mutant in leucine-limitation; RA: rplA mutant in ammonia limitation; M: Batch glucose minimal media; R: Batch defined rich media. All samples were grown in 40 mM MOPS media.
The growth rate is suggested to be the major dominant effect on cellular physiology in microorganisms. In yeast cells growing under different nutrient limitations and at different rates, gene expression shows mostly a function of growth rate[16]. However, it is not clear whether a prokaryotic microorganism would behave the same. Here, I measured *E. coli* gene expression at six different growth rates under three different nutrient limitations (carbon, nitrogen, phosphorus) using microarray and showed that unlike yeast, gene expression in *E. coli* is dominated mostly by nutrient limitation (Fig. C.1). Samples under the same nutrient limitation resemble each other (Fig. C.1b). Principle component analysis shows that growth rate only explains less than 10% of the total variance (PC3, Fig. C.1c and d). This result suggests that growth rate dependent gene expres-
Figure C.1: Gene expression of *E. coli* is dominated by nutrient limitation than growth rate. (a) Hierarchical clustering of *E. coli* gene expression across glucose (C), ammonia (N), and phosphorus (P) limitations. Six growth rates are measured from 0.1 to 0.6 h⁻¹ with 0.1 h⁻¹ increment. Log2 ratio is plotted. (b) Heatmap of gene expression correlation between the samples. Samples under the same limitations are closer to each other. (c) Fraction of variance explained by principle components of Principal Component Analysis. (d) Eigen-gene expression profile of the first three principle components. Growth rate is captured mostly in PC3.

Expression is not a universal feature of microorganism. Possibly, yeast cells have TOR pathway that can integrate all nutrient inputs to regulate global gene expression; whereas in *E. coli* each nutrient input has different control node. Future studies on the translation and protein expression level can reveal how much responses at transcription level reflects overall responses in gene expression.
Measurement of metabolite pool sizes in *E. coli* across different nutrient conditions and growth rates by LC/MS

The cellular metabolic network consists of enzymes that convert small molecules to provide building blocks and energy. In *E. coli* there are about 4000 genes and more than 1000 of them are annotated as metabolic enzymes [46]. The metabolic network exhibit incredible flexibility to support cell growth under different metabolic constraints. In order to characterize different metabolic constraints across conditions, I measured metabolite pool sizes of *E. coli* across 24 conditions (six growth rates and four limitation: glycerol, glucose, ammonia, phosphate). Interestingly, while glucose and glycerol feed into glycolysis via different nodes, they show very similar metabolite pool sizes. This similarity between glycerol and glucose limited cells indicate that there is a certain pattern of metabolite pool size distribution that reflect carbon limitation, despite we do not know the
carbon-abundance sensing molecule. N-limited cells have high level fructose-1,6-bisphosphate, indicating high flux of glycolysis[76]. They also show high level of TCA intermediates. Importantly, N-limited E. coli cells do not lower all amino acid pool sizes, but only a few including glutamine, serine, asparagine and aspartate. Glutamine pool showed the strongest decrease and it is known intercellular glutamine pool is an indicator of external nitrogen abundance[66]. P-limited cells have decreased pools of nucleotides. However, the energy charge is not drastically decreased (Fig. D.1b). There is no clear trend in the NADPH/NADP⁺ ratios across conditions. On the other hand, NADH/NAD⁺ does have differences in different conditions. N-limited cells have overall very low values, while P-limited cells show a decrease in the NADH/NAD⁺ ratio as phosphate concentration increases. This is the first time to observe that P-limitation induces high NADH/NAD⁺ ratio.

All the samples were also converted to absolute concentrations. The majority is glutamate, which is served as an osmolite and a main anion ion in the cell. The concentration of glutamate
Figure D.2: Absolute concentrations of metabolites in *E. coli* cells under different nutrient limitations and growth rates. Cells were grown in chemostat until steady states under different nutrient limitations. Cultures were collected using filter membranes and quenched with -20°C 40:40:20 (Acetonitrile:Methanol:Water) solution. Metabolites annotated with (a) indicate the acidic quenching condition (with 0.5% Formic acid and neutralized with 15% NH₄HCO₃ with the ratio of 800:67). Samples were analyzed on an LC/MS system with an Exactive Orbitrap MS (Thermo). A U-¹³C glucose Gutnick minimal media culture was extracted on the same day with the samples and mixed 1:1 OD ratio before LC/MS analysis. The absolute concentration was converted using the ion count ratio between the sample and the glucose minimal media culture that has previously determined absolute metabolite pool sizes. Glycerol-, glucose-, and ammonia- (N-) limited cultures were grown in 1x Gutnick minimal media. Phosphorus- (P-) limited cells were grown in 8 mM MOPS minimal media (5x dilution of normal 40 mM) to lower the effect of MOPS on MS. While no other obvious phenotype differences observed, the total metabolite pool size seems to be smaller in P-limited samples than the others, potentially due to the limitation of osmolarity. The right three bars are derived from previously determined glucose minimal media data with upper bound (UB) and lower bound (LB).

is consistent across samples with the same osmolality except P-limitation, which has lower osmolality due to decreased MOPS concentration to avoid damage on the MS. N-limitation showed an increase in FBP level, close to the concentration in glucose batch culture. P-limited cells have an overall high concentration of glutathione and high fraction of glutathione disulfide, which indicates that P-limitation might induce higher oxidative stress due to the substrate for ATP synthase, phosphate, is being limited. In general, the total concentration of metabolites are consistent across different conditions.
Measurement of metabolite pool sizes in synchronized *Caulobacter crescentus* during the cell cycle by LC/MS

The bacterium *Caulobacter crescentus* has clearly defined cell cycle with asymmetric cell division. It has serve as a model system to understand bacterial cell cycle regulation. A previous study has shown that the gene expression of *Caulobacter crescentus* is temporally regulated according to the cell cycle[82]. During the cell cycle, different types of biomass need to be synthesized to replicate the cell fully. Thus, the metabolic activity can be modulated to support different anabolic demands. To understand how metabolism changes during a cell cycle, *Caulobacter crescentus* was synchronized and metabolites were measured. The result is presented in Fig. E.1. One thing to note that the culture was undergone synchronization process for two hours without carbon sources. Thus, when comparing to an unsynchronized culture, carbon starvation signal
such as low in carbon-rich molecules are apparent (acetyl-CoA, pyruvate, α-ketoglutarate). Pre-
cursors in amino acids and nucleotides, on the other hand, seems to have oscillatory concentra-
tions. Purine related metabolites including dGMP, AMP and IMP increased early soon after the
division. Pool sizes of amino acids peaked after 75 minutes.

The notion of temporal compartmentalization of metabolism is an interesting one, and the ob-
servation of yeast metabolic cycle might offer a glimpse to this less well-testified idea. The cycle
in metabolic activity might influence the gene expression and vice versa, creating an entrenched
cycle of cell behavior. Future studies can refine the current approach to characterize the metabolism
of *Caulobacter crescentus* through a cell cycle. It will also be interesting to look into bacteria that
do not have defined cell cycle phases such as *E. coli*. The current method for synchronizing *Caulobac-
ter crescentus* can be disruptive to its metabolism and thus not ideal for the downstream study.
The use of baby machines[10] might be able to circumvent this issue and provide new insights
into the interaction between metabolism and cell cycle.
Figure E.1: Metabolite pool sizes as a function of the cell cycle in a synchronized *Caulobacter crescentus* culture. Cells were synchronized using centrifugation (~2 hr). At each time point, 3 mL culture was collected using a filter membrane and quenched in -20°C 40:40:20 (Acetonitrile:Methanol:Water) solution. Samples were analyzed on an LC/MS system with an Exactive Orbitrap MS (Thermo). The signals were normalized to an unsynchronized culture which did not undergo mock synchronization. The log2 ratios are plotted with the color bar indicating the scale. The top images showed the representative morphology of cells at each time point.
The effect of *dksA* on ribosome abundance and activity in *E. coli*

*dksA* is found to affect RNA polymerase with ppGpp to affect rRNA transcription and thus ribosome abundance[113]. It is found to independently and collectively influence gene expression with ppGpp. Deletion of *dksA* is found to mimic the deletion of ppGpp including multiple amino acid auxotroph and aborption of growth-rate-dependent rRNA transcription[113, 114] (reflected by RNA-to-protein ratio (R/P ratio) as most RNA is rRNA). In Fig. F.1, we found that in batch we observed the same over expression of rRNA transcription. However, the growth rate dependence of R/P ratio did not change under phosphorus limitation using chemostat with a slight overall higher level of R/P ratio. Thus, this data suggest that ppGpp and *dksA* are not the only pathway affecting the growth-rate-dependent rRNA regulation, at least under phosphorus limitation. Other mechanism such as nucleotide pool might be important, as well.
Figure F.1: Deletion of dksA affects ribosome abundance but does not abort growth-rate dependent RNA-to-protein ratio under phosphorus limitation. (a) RNA-to-protein ratio of wild type and dksA deletion mutant for chemostat cultures upon P-limitations and batch cultures at different growth rates. All of the media is supplemented with amino acids due to the amino acid auxotrophy of dksA deletion. Each data point represents three technical replicates. (b) Quantification of assembled ribosomes in the form of free 70S, mRNA-bound 70S (one ribosome on one mRNA), and polysomes (multiple ribosomes on one mRNA) at growth rate 0.1 h⁻¹ under P-limitation.
References


[27] R Contestabile, S Angelaccio, F Bossa, H T Wright, N Scarsdale, G Kazanina, and
V Schirch. Role of tyrosine 65 in the mechanism of serine hydroxymethyltransferase. *Bio-

[28] Mary T Couvillion, Illiana C Soto, Gergana Shipkovenska, and L Stirling Churchman. Syn-

[29] Schneider DA, Gaal T, and Gourse RL. NTP-sensing by rRNA promoters in Escherichia coli
is direct. *Proceedings of the National Academy of Sciences of the United States of America*,

Okano, James R Williamson, Kurt Fredrick, Yi-Ping Wang, and Terence Hwa. Reduction
of translating ribosomes enables Escherichia coli to maintain elongation rates during slow

[31] D G Dalbow and R Young. Synthesis time of β-galactosidase in Escherichia coli B/r as a

1526, 2009.

[33] Cristian Del Campo, Alexander Bartholomäus, Ivan Fedyunin, and Zoya Ignatova. Sec-
ondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA

[34] Patrick P Dennis and Hans Bremer. Modulation of Chemical Composition and Other Para-
eters of the Cell at Different Exponential Growth Rates. *EcoSal Plus*, 3(1):1–56,
September 2008.

[35] Kimberly A Dittmar, Michael A Sørensen, Johan Elf, Mans Ehrenberg, and Tao Pan. Select-
tive charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO reports*, 6(2):

[36] E A Doherty, R T Batey, B Masquida, and J A Doudna. A universal mode of helix packing in

[37] Viktoria Dorfer, Peter Pichler, Thomas Stranzl, Johannes Stadlmann, Thomas Taus,
Stephan Winkler, and Karl Mechtler. MS Amanda, a Universal Identification Algorithm Optim-

[38] Christopher D Doucette, David J Schwab, Ned S Wingreen, and Joshua D Rabinowitz. α-
Ketoglutarate coordinates carbon and nitrogen utilization via enzyme I inhibition. *Nature

[39] Gregory S Ducker, Li Chen, Raphael J Morscher, Jonathan M Ghergurovich, Mark Esposito,
Xin Teng, Yibin Kang, and Joshua D Rabinowitz. Reversal of Cytosolic One-Carbon Flux
1153, June 2016.

[40] Gregory S Ducker and Joshua D Rabinowitz. One-Carbon Metabolism in Health and Dis-

[41] J G Dunn, C K Foo, N G Belletier, E R Gavis, and J S Weissman. Ribosome profiling re-
veals pervasive and regulated stop codon readthrough in Drosophila melanogaster. *eLife*,
2, 2013.

[42] Joshua G Dunn and Jonathan S Weissman. Plastid: nucleotide-resolution analysis of next-


[48] Ye Fu, Qing Dai, Wen Zhang, Jin Ren, Tao Pan, and Chuan He. The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. *Angewandte Chemie (International ed. in English)*, 49(47):8885–8888, November 2010.


A Novick and L Szilard. Experiments with the chemostat on spontaneous mutations of bacteria. … of the National Academy of Sciences, 1950.


Kun Wang, Bo Long, Jian-Qin Jiao, Jian-Xun Wang, Jin-Ping Liu, Qian Li, and Pei-Feng Li. miR-484 regulates mitochondrial network through targeting Fis1. *Nature communications*, 3(1):781, April 2012.


