Elucidating the regulatory role of CRP in coordinating protein biosynthesis machinery with metabolism that defines growth optimality in *Escherichia coli*

Ankita Pal[1, #], Mahesh S. Iyer[1, #], Sumana Srinivasan[1], Aswin S Seshasayee[2, *], K.V. Venkatesh[1, *]
[1]Department of Chemical Engineering, Indian Institute of Technology-Bombay, Powai, Mumbai-400076, India,
[2]National Centre for Biological Sciences, GKVK, Bellary Road, Bangalore-560065, India.
[#]Co-first author
[*]Corresponding author

Abstract

Bacterial gene expression is governed by two synergistic mechanisms — growth-rate dependent global machinery (e.g. ribosomes, RNA polymerase, metabolites and cofactors) and transcription factors that work together for optimal growth. However, in presence of glucose, coordination of global transcriptional regulator cAMP-CRP with such physiological resources, remain elusive. Here, we reveal that the deletion of CRP results in metabolic dysregulation with significant perturbation in protein biosynthesis machinery coupled to impaired glucose import. Additionally, we quantitatively demonstrate that the loss of CRP unbalances proteome allocation, favouring stress or hedging related functions over growth-enhancing functions. To address how *Escherichia coli* can cope up with such a system-wide upset, we adaptively evolved *Δcrp* mutant in the presence of glucose. We explicitly show that adaptation to loss of CRP frequently occurs via mutations in the *ptsG* promoter. This causes higher glucose uptake, resulting in the restoration of the levels of the protein biosynthesis machinery and metabolic rewiring of ATP towards synthesis of costly amino acids along with reduction in unnecessary cAMP synthesis. Thus we elucidate the molecular events underlying growth optimality of the organism, driven by CRP.

Introduction

In bacteria like *E. coli*, efficient allocation of cellular resources for optimal growth under changing environmental conditions is well appreciated (1–3). Transcriptional regulators along with growth-rate dependent global machinery (e.g. RNA polymerase and associated sigma factors, ribosomes, metabolites) ensure proteomic resource allocation in accordance with the metabolic demands. CRP, a global transcriptional regulator in *E. coli*, along with its cognate signalling molecule cAMP regulates a large repertoire of genes specifically, a) metabolic genes (4–7) , b) genes involved in the transport of carbon sources (8, 9), c) genes pertaining to the stress response (8–10), and d) several other genes coding for transcriptional regulators under diverse nutritional conditions (11). However, it is relatively unknown how deletion of a global transcriptional regulator like CRP perturbs such global physiological factors in glucose minimal media conditions. Further deciphering the adaptive rewiring of systemic response to balance these global factors, will provide crucial insights into mechanisms that govern the growth optimality of the organism.

Adaptive laboratory evolution (ALE) entails orchestration of genetic as well as phenotypic behaviour in response to mutations that provide growth fitness benefits to organisms under strict selection pressures (12). While a majority of the ALEs have focussed on understanding the adaptive rewiring in response to the loss of metabolic genes (13–16), fewer studies have focussed on ALEs on the loss of global transcriptional regulators (17–20). Such evolutionary studies on global regulators can presumably elaborate their importance in complementing the pleiotropic effects of growth machinery.

Here, using a multi-omics (genomics, transcriptomics, metabolomics and phenomics) approach, we illustrate the evolutionary importance of CRP in glucose minimal media conditions. First, we elucidate that evolution redress the loss of CRP by repeatedly accumulating mutations in the intergenic region of the *ptsG* gene, enabling increased glucose uptake rates and fitness benefit for the organism. Next, we unravel how such adaptation bring about system-wide changes by modulating the global gene expression states and the levels of several key intracellular metabolites that coordinate protein biosynthesis machinery and metabolism. Finally, restoration mechanism as a result of evolution involved fine tuning of proteome allocation in favour of

growth and away from stress or hedging functions. Further, using a genome-scale model we were able to quantitatively detail the energetic inefficiencies of the evolved populations which explained their inability to grow optimally as wild-type. Overall, the evolution of *Δcrp* strain suggests an underlying paradigm which delineates the inherent constraints of genetic and metabolic networks in *E. coli*.

Materials and Methods

Strains

An *E. coli* K-12 MG1655 (CGSC#6300), was used as the parent strain in this study. We constructed *Δcrp* knockout in this genetic background by **λ**-Red mediated recombination (21), using plasmids pKD46, pKD13 and pCP20. The generation of the knockout strain was confirmed by PCR with custom made oligo-nucleotides against the genomic DNA followed by Sanger sequencing. Generation of *Δfis, Δmlc, ΔfisΔcrp* and *ΔmlcΔcrp* strains for growth studies were also generated using the same procedure.

Physiological Characterization

For transcriptome, metabolome and phenotype characterizations, experiments were performed in 500 mL bioreactor (Applikon) containing 200 mL M9 media (6 g/L anhydrous Na_2HPO_4, 3 g/L KH_2PO_4, 1g/L NH_4Cl, 0.5 g/L NaCl + 2 mM $MgSO_4$ + 0.1 mM $CaCl_2$), with 2g/L glucose and 40mM MOPS. The temperature of the bioreactor was maintained at 37°C and the pH of the media was maintained at pH 7.2 using 40 mM MOPS buffer to prevent effects of pH change on growth. Aeration was done by sparging air in the bioreactor at 700 mL/min and at all times the dissolved oxygen (DO) levels were maintained above 40% saturation using mass flow controllers (MFC). Growth was monitored by collecting samples at regular intervals and measuring the optical density (OD) at 600 nm in a spectrophotometer (Thermo Multiskan GO) until the organism reached its stationary phase. The growth rate was calculated from the slope of the linear regression line fit to ln(OD600) vs. time (in hrs) plot. The dry cell weight (DCW) was determined from the O.D.600 values by using the experimentally derived relationship that 1.0 O.D.600 corresponds to 0.44 gDCW*h^{-1}. The phenotypic characterizations were done for three biological replicates (n = 3). Samples were also collected during regular intervals to determine the rate of

glucose uptake as well the rate of secretion of extracellular metabolites. The samples were centrifuged and the supernatant was collected which was then used to determine the concentrations and rates using HPLC (Agilent 1200 Series) equipped with Bio-Rad Aminex HPX- 87H ion exclusion column with 5mM H_2SO_4 as the mobile phase. The column was maintained at temperature 50°C and the flow rate at 0.6 mL/min. The biomass yield for each of the strains was determined by normalizing with their specific glucose uptake rate (gDCW/gGlucose). Oxygen uptake rate was measured using dissolved oxygen probe (Applikon) while growing the cultures in the bioreactor during its exponential phase. All physiological measurements were checked for statistical significance using unpaired two-tailed Student's t-test.

Adaptive Laboratory Evolution Protocol

Adaptive evolution of replicate populations of *Δcrp* was carried out in shake flask with M9 minimal media with 2g/L glucose and 40 mM MOPS. MOPS was added to maintain the populations at a constant pH during the course of evolution. All the replicate flasks were passaged serially into fresh media strictly in the exponential phase to ensure that fitness gain occurs via increased growth rates. This also prevents the cultures from entering the stationary phase where glucose limited and thereby avoids complexities associated with the onset of stationary phase. As the growth rate of the organism changed during evolution, the volume of culture passaged was adjusted to prevent entry into the stationary phase. Glycerol stocks were made during each passage. This procedure was followed for 10 days (approximately 100 generations) and after it had reached a stable growth and no further increase in growth rate was observed, ALE was terminated. Genotype and phenotype characterizations of the evolved replicates were done for the population and not for individual clonal samples to ensure that all the traits are taken into consideration. Characterization at the population level is a more efficient and feasible approach to understand the underlying mechanisms of evolution in an unbiased manner as it better reflects the properties of the population as a whole (16).

Whole Genome Resequencing (WGS)

Genomic DNA from all the endpoint populations were extracted using GenElute Bacterial Genomic DNA Kit (NA2120; Sigma-Aldrich, St. Louis,MO) using the manufacturer's protocol.

The integrity of the extracted genomic DNA was analysed by running it on an agarose gel and the quality was assessed and quantified using Multiskan GO (Thermo Scientific). The genomic DNA library was prepared using Illumina TruSeq DNA Nano Kit. The quality of the libraries was checked using Agilent Bioanalyzer. The libraries were then sequenced from both ends (paired-end) on Illumina HiSeq250 platform with 2 X 100 cycles. All the samples had an average of 275X mapped coverage.

The raw reads obtained from the sequencer were trimmed using CUTADAPT to remove TruSeq adapter sequences. The breseq pipeline was used to identify the point mutations (SNPs), insertions and deletions (22) mapped to *E. coli* genome K-12 MG1655 (GenBank accession no. NC_000913.3). Breseq was run using the -p option (for population samples) with default parameters to identify mutations present in only a subset of the population at a frequency less than 100%. Mutations that had 100% frequency were assumed to be parental mutations and were not considered. Only the mutations predicted with high confidence under the category "predicted mutations" were further analyzed in this study.

Mutation Validation

To determine the causality of the mutations detected by WGS, the mutations were introduced into the *Δcrp* background using *in vivo* site-directed mutagenesis. However, to limit the scale of this study and given the consistent occurrence of mutations in the Fis binding region, only the IG116 mutation (SNPs) detected in the promoter of *ptsG* gene in the EvoCrp1 strain was selected for validation. Additionally, IG116 mutation was also investigated for Mlc binding. To construct the strains, scarless editing of the *E. coli* K-12 MG1655 *Δcrp* genome was done using a two-step recombination method using pSLTS plasmid (Addgene plasmid #59386) (23). The primers used in this method are given in Supplementary Table S2. After construction of the strains, they were sequenced to verify the correct introduction of the SNPs in the genome. This strain was further used to add 3X FLAG-tag to Fis protein or Mlc protein as described in the section below.

ChIP qPCR

A 3X FLAG-tag was added to the C-terminus of the Fis or Mlc protein using pSUB11 plasmid (24) as the amplification template. The primers used for amplification for tagging are given in Supplementary Table S2. The amplified construct was then introduced into *E. coli* K-12 MG1655

Δcrp strain. The constructed strains were verified by PCR and Sanger sequencing. ChIP experiment was carried out using the protocol as described previously (25) . The DNA samples immune-precipitated by this method was recovered by PCR purification and was quantified by qPCR using primers for the specific as well as non-specific region (*frr*). Fold change in IP-ed DNA compared to mock DNA was calculated as $2^{-\Delta\Delta Ct}$ as described previously (26, 27).

RNA extraction and enrichment of mRNA

For each sample/strain, RNA extraction of two biological replicates were performed. The cells were grown till mid- exponential phase and then 50 mL cells were harvested by centrifugation. RNA extraction was done using the TRIzol-chloroform method as previously described (19, 28). DNase treatment was done to remove any DNA contamination after which enrichment for mRNA was done using the MicrobExpress Kit following the manufacturer's protocol. Single-end, strand specific libraries for RNA sequencing were prepared using NEBNext Ultra II Directional RNA library kit. The quantity of the mRNA was measured using qubit and the quality as well as integrity was checked using BioAnalyzer. The sequencing was carried out on HiSeq 2500 Rapid Run Mode using 1x50 bp format.

Transcriptome data analysis

The raw files were first trimmed using CUTADAPT to remove adapter sequence and low-quality reads. The reads were then mapped to the *E. coli* genome K-12 MG1655 (GenBank accession no. NC_000913.3) using BWA (29). The sam file generated was then converted to compressed bam format using Samtools (30). FeatureCounts (31) was used to assign counts at the gene level using reference genome provided in the GTF format. Ecocyc database (version 21.5) (32) was used to retrieve the annotations for 4662 genes. Differential gene expression (DGE) was analyzed from the raw counts by EdgeR (33) after removing genes having less than 10 reads. The genes that showed ≥ 2-fold change in expression and had adjusted $P < 0.05$ (Benjamini-Hochberg) were considered for further analysis. The DE genes were enriched for metabolic pathways using KEGG pathway classification (34) as defined in Proteomaps (www.proteomaps.net). The mapped genes were represented as Voronoi treemaps (version 2.0) (35). The significance of the upregulated and downregulated genes within each category were validated using a hypergeometric test with P-value correction using Benjamini Hochberg in R (R Core Team

2019). For each up/down regulated category, we arbitrarily choose atleast five DE genes enriched to be considered for significance analysis. Those DE genes which do not have an assigned "Accession ID" (Ecocyc Version 21.5) such as phantom genes were excluded from the above analysis.

For sigma factor enrichment analysis, the DE genes enriched by KEGG pathway classification were further enriched based on regulators using data available in EcoCyc and Regulon DB (36). The upregulated and downregulated genes belonging to each regulator were then validated using hypergeometric test in R (P-value < 0.05). Only those sigma factors which regulated atleast 10% of the KEGG enriched upregulated and downregulated DE genes, were retained for this over-representation analysis. Pearson correlation of log2 fold changes of EvoCrp vs WT and Δ*crp* vs WT and statistical significance were performed in R.

Metabolomics

1. Chemicals for metabolomics

LC- MS grade methanol, acetonitrile and ammonium hydroxide (≥ 25% in water) were purchased from Honeywell. Analytical grade chloroform, HPLC-grade water and LC-MS grade ammonium acetate was purchased from Sigma. All metabolites used as external standards were purchased from Sigma. Uniformly labelled U-^{13}C (> 99% purity) glucose was purchased from Cambridge Isotope Laboratories.

2. Extraction of metabolites

Metabolite samples were harvested in the mid-exponential phase in biological triplicates and technical duplicates. A Fast-Cooling method (37, 38) was used to quench the harvested cells as reported previously. Briefly, ~10 mL culture (~ 6-7 OD cells) was directly poured into 2 mL chilled (4°C) M9 (without glucose) in a 50 mL falcon tube. The tube was then dipped in liquid nitrogen for 10 secs to bring down the sample temperature to 0°C. To prevent the formation of ice crystals, the sample was vigorously agitated with the help of a digital thermometer. Samples were then immediately centrifuged at 0°C, 7800 rpm for 5 min. The supernatant was discarded and the pellet was snap frozen in liquid nitrogen and stored at -80°C until metabolite extraction was done.

For extraction of metabolites, the sample pellet was dissolved in 700 µL chilled (-80°C) methanol: 300 µL chilled (-20°C) chloroform (38), followed by snap freezing in liquid nitrogen and homogenization using a hand-held pestle (Sigma #Z359971) all within 2 min. ^{13}C-labelled extracts as internal standards were generated for quantification of key metabolite pool sizes using isotope-based dilution method (39). The samples were spiked at the earliest stage of extraction with a fixed volume of internal standard taken from the same batch. The sample tubes were then placed overnight on a Thermo-shaker maintained at 0°C and 400 rpm. To the sample tubes, 500 µL of 2% ammonium hydroxide (4°C) prepared in HPLC-grade water was added and incubated on ice for 10 min to enable phase separation. The tubes were then centrifuged at 4°C, 13000 rpm for 15 min and the aqueous layer was collected in a chilled 1.5 mL tube. The samples were then completely dried in a vacuum concentrator and stored at -80°C. Prior to analysis in LC-MS/MS, samples were reconstituted in 100 µL chilled (-20°C) acetonitrile: buffered water (60:40, v/v), centrifuged at 4°C for 10 min and the supernatant was transferred to pre-chilled glass vials. Buffered water consisted of 10 mM ammonium acetate, pH 9.23 adjusted with ammonium hydroxide prepared in HPLC- grade water (40). The volume of the pooled internal standard was standardized such that the external standard and internal standard peak height differed less than 5-fold (41). Metabolites were extracted from three biological and two technical replicates (n = 6).

3. LC-MS/MS settings

The samples were analyzed using high resolution mass spectrometer in Orbitrap Q Exactive Plus (Thermo) equipped with a SeQuant ZIC-pHILIC column (150 mm x 2.1 mm x 5 micron packing, Merck) and a ZIC-pHILIC guard column (20 mm x 2.1 mm x 5 micron packing, Merck) under alkaline mobile phase conditions with ESI ion source. The ESI was operated in positive (M+H) and negative (M-H) polarity switching mode. The spray voltage was set at 4.2 kV and 3.5 kV for the positive and negative mode respectively. The temperature was maintained at 300°C and 320°C for the ion transfer capillary and probe heater respectively. A heated electrospray ionization probe II (H-ESI-II) probe was used with the following tune parameters: sheath gas, 29; auxiliary gas, 7; sweep gas, 0; S-lens at 45 arbitrary units. A full scan range of 66.7 to 1000 m/z was applied for positive as well as negative modes and the spectrum data type was set to profile mode. The automatic gain control (AGC) target was set at 1e6 with resolution of 70,000 at 200 m/z. Prior to analysis, cleaning of LC-MS system and ion transfer capillary (ITC) along with mass calibration was done for both positive and negative ESI polarities by using Thermo Calmix solution along with MS contaminants to take into account lower mass ranges. The signals

of compounds 83.06037 m/z (2 x ACN+H) and 119.03498 m/z (2 x Acetate-H) were selected as lock masses for positive and negative mode respectively with lock mass tolerance of 5 ppm (42).

The mobile phase for chromatographic separation comprised of non-polar phase A (acetonitrile: water mixed in the ratio 9:1, 10 mM ammonium acetate, pH 9.23 using ammonium hydroxide) and polar phase B (acetonitrile: water mixed in the ratio 1:9, 10 mM ammonium acetate, pH 9.23 ammonium hydroxide). A linear gradient with flow rate of 200 µL/min was set as follows: 0-1 min: 0% B, 1- 32 min: 77.5% B, 32- 36 min: 77.5% B to 100% B, 36-40 min: hold at 100% B, 40-50 min: 100% B to 0% B, 50-65 min: re-equilibration with 0% B (40). Injection volume of 5 µL was used for all the samples and standards.

4. Metabolomics data analysis

The data from the machine was processed using the software package Xcalibur 4.3 (Thermo Fisher Scientific) Quan Browser. A semi-quantitative analysis was performed using peak heights of precursor ions with signal/noise (S/N) ratio more than 3, a retention time window of less than 60 secs and less than 5 ppm mass error. The peak heights of the samples were normalized to the peak height of the internal standards to obtain a height ratio. Only those metabolites were retained for analysis, which had naturally occurring ^{12}C peak height less than ~10% of the ^{13}C- labelled peak height in the internal standard. MetaboAnalyst (43) was used for identifying statistically significant metabolites and for Partial Least Square Discriminant Analysis (PLS-DA) on biomass normalized and g-log transformed metabolite concentrations. Missing value imputation were performed using SVD impute function. The relative concentration of metabolites is expressed as height ratio normalized to biomass (as height ratio/gDCW). Metabolites levels with false discovery rate (FDR) < 0.05 were considered for further analysis. For PLS-DA, out of the 40 metabolites, only those that were statistically significant in all the three conditions, i.e., *Δcrp* vs WT, EvoCrp vs *Δcrp* and EvoCrp vs WT, were considered for analysis.

ME- model simulations

The ME-model (44, 45) was simulated by constraining the glucose uptake rates in a range as reported previously (46). Any gene predicted to be expressed in any of the 20 simulations was classified as “utilized ME”; genes within the scope of the ME-Model but not expressed or having very low expression are classified as “non-utilized ME”; The transcriptome fraction was determined from the product of TPM (transcript per million) (47) and gene length divided by the

sum-product of these calculated over all the genes. TPM for all the genes were calculated on raw counts (genes < 10 reads were not considered). Next, a unique list of genes and corresponding TPMs were calculated from all the genes differentially expressed (adj-P value < 0.05, aFC >=2) in atleast one condition (*Δcrp* vs WT, EvoCrp vs *Δcrp* and EvoCrp vs WT). These genes were then enriched for utilized ME or non-utilized ME followed by summing the transcriptome fractions of all utilized ME and non-utilized ME genes respectively (48).

Calculation of non- growth ATP maintenance (ATPM)

The experimentally measured growth rates, glucose uptake and acetate secretion rates were used as flux constraints in the iJO1366 metabolic model (49). Constrained-based flux balance analysis was performed using COBRApy (50) by maximization of ATP maintenance (ATPM) reaction flux. This predicted ATPM flux was normalized to glucose uptake flux to calculate the yield (gATP/gGlucose) for each of the strains.

Results

Loss of CRP caused large shifts in the transcriptome of key metabolic pathways

CRP occupies an apex node in the hierarchical transcriptional regulatory network (TRN) of *E. coli* regulating multitude of genes under diverse nutritional conditions. However, very few studies address the role of CRP in regulating key metabolic pathways with glucose as a carbon source. To assess the systemic effect caused by the loss of CRP, we performed high-coverage RNA sequencing of Δ*crp* mutant in glucose minimal media during the mid-exponential phase. The transcriptome of this strain when compared to its parent wild-type (WT) strain, showed ∼722 differentially expressed (DE) genes (Absolute Fold change (aFC) ≥ 2, Adjusted P-value (adj-P) < 0.05) of which ∼544 genes (75%, adj-P < 0.05) were downregulated and ∼178 genes (25%, adj-P < 0.05) were upregulated in the mutant (Figure 1A), indicating a large upset of the global transcriptome.

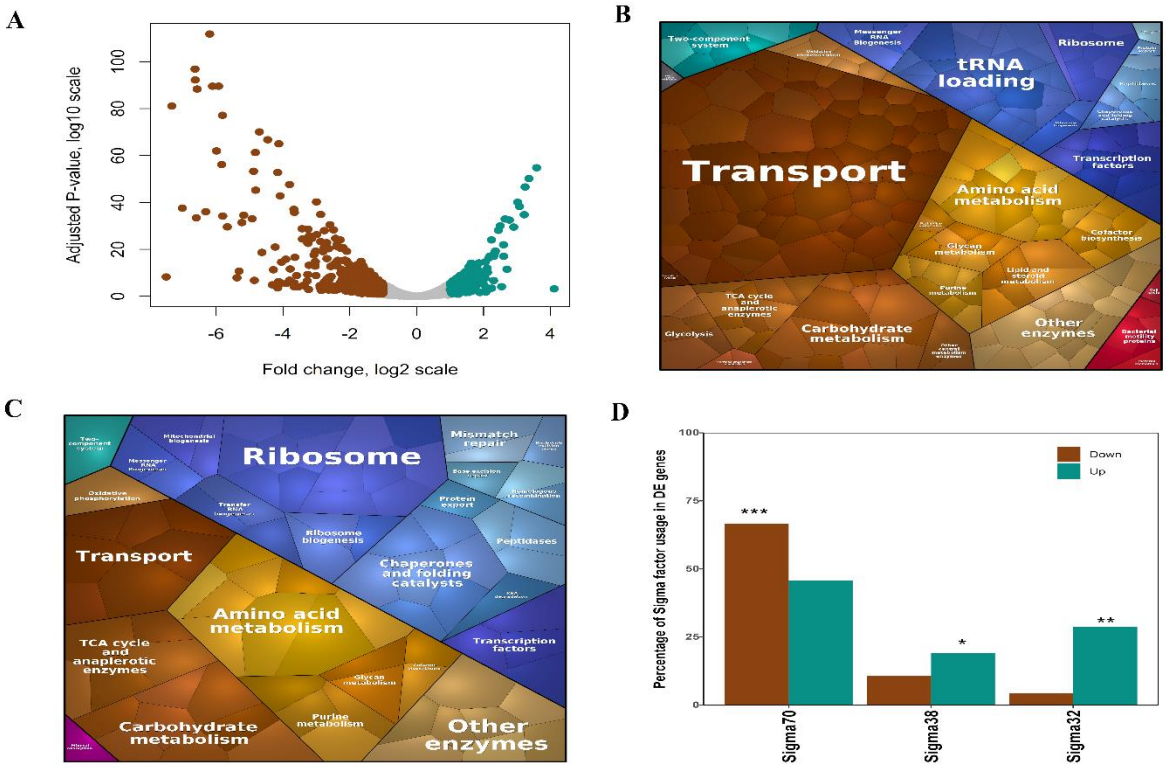

Figure 1. Severe perturbation of gene expression in Δ*crp* compared to WT. (A) Volcano plot of the DE genes Δ*crp* compared to WT depicted as Adj-P value (log10 scale) vs Fold change (log2 scale). The brown dots indicate downregulated genes and the cyan dots indicate upregulated genes. (B) The Voronoi treemaps showing the downregulated metabolic pathways enriched by KEGG classification. (C) The Voronoi treemaps showing the upregulated metabolic pathways enriched by KEGG classification. The genes enriched within each pathway are shown in Supplementary Figure S1. Note: The size of the hexagon within each category is directly proportional to the absolute fold change observed for the genes. The color of the hexagon denotes the specific pathways classified by KEGG. (D) Enrichment of genes under regulation of sigma factors; the brown bars and the cyan bars indicate the fraction of downregulated and upregulated genes in Δ*crp* vs to WT respectively. Significant increase or decrease is denoted by asterisks: one asterisk indicates $P < 0.05$, two asterisks indicate $P < 10^{-4}$ and three asterisks indicate $P < 10^{-10}$.

We identified the KEGG pathways that were significantly enriched among these DE genes and represented these pathways as Voronoi treemaps (Figure 1B-C & Supplementary Figure S1). Our data indicated significant downregulation of "Transporters" and "TCA cycle and anaplerotic enzymes", consistent with previous gene expression data. Among the transporters, the major transport gene *ptsG* responsible for glucose uptake, showed a ~6.5-fold reduction in its expression, consistent with the known regulatory effect of CRP. Alternatively, secondary glucose transporters (*mal* genes, *man* genes, and *lamB*), glycerol transporter (*glpF*) and outer membrane porin (*ompF*) were also downregulated (aFC ~3) as reported previously. Such reduction in *ompF* levels might indicate the altered osmolarity experienced by the organism

which can affect the protein translation capacity (51). Additionally, we observed downregulation of “tRNA loading” pathways and significant upregulation of “Ribosomes” and “Chaperone and folding catalysts” related pathways. These are associated with the general stress response of the cell required to maintain proper protein turnover and integrity (52). Owing to the huge investment of proteomic resources for synthesis of ribosomes, such increase in ribosome transcript levels might indicate an increase in unnecessary ribosomal proteins, resulting in reduced availability of resources for synthesis of metabolic proteins (53, 54). Further, we found upregulated genes enriched in “TCA cycle and anaplerotic enzymes” pathway to be primarily involved in glycolate and glyoxylate degradation (*glcDEF*, *aceA*, aFC > 2), which might indicate an increase in hedging mechanism related to alternate carbon metabolism. Overall, the considerable shifts in the transcriptome of these pathways, emphasize the metabolic dysregulation coupled to impaired protein biosynthesis machinery caused by the loss of a global regulator.

Additionally, of the large fraction of downregulated genes, ~66% ($P < 10^{-6}$) were found to be regulated by sigma 70, the major growth-related sigma factor associated with RNA Polymerase (Figure 1D), corroborating the association of CRP with sigma 70 reported previously (7, 55–58). Further, a significant fraction of upregulated genes was regulated by stress-related sigma factor, sigma 38 (~28%, $P < 10^{-2}$) and sigma 32 (~22%, $P < 10^{-13}$). Presumably, this suggested the reallocation of RNA Polymerase away from growth as a direct effect, and upregulation of stress related genes as an indirect effect due to the loss of a global regulator.

Adaptive evolution enables growth recovery via mutations in the *ptsG* intergenic region

To decipher how Δ*crp* copes with the perturbations in its global gene expression, we adaptively evolved five independent replicate populations of the mutant with multiple passages in batch culture during the mid-exponential phase (Supplementary Figure S2). We observed a rapid growth recovery within ~100 generations and the endpoint populations (EvoCrp) were further characterized in this study (Supplementary Figure S3).

Next, we performed whole genome resequencing (WGS) to identify the causal mutations in all the EvoCrp strains (Supplementary Table S3). Out of eleven unique mutations detected across all the EvoCrp strains, six were in the upstream promoter sequence of the *ptsG* gene. The intergenic mutations were mostly SNPs specifically in the operator sites of the negative regulators (repressors) namely Fis (59, 60), Mlc (61, 62), and ArcA (63) which repress *ptsG* gene

expression (Supplementary Figure S4A). Since ~70-80% of each of the EvoCrp populations showed a mutation in the Fis operator sites, we assessed whether Fis deletion can enable growth recovery by relieving *ptsG* gene expression. Interestingly, deletion of *fis* in *Δcrp* background did not cause any changes in growth rate (Supplementary Figure S5A). Next, we introduced the IG116 promoter mutation (as annotated in Supplementary Table S3) in the *Δcrp* background and characterized the binding affinity of Fis in IG116 *Δcrp* mutant strain using CHIP-qPCR (Supplementary Figure S5C & S6C). However, we did not observe any significant enrichment for Fis binding to the promoter in IG116 *Δcrp* strain as well as in WT strain (Supplementary Figure S6A-C), similar to a previous study in *E. coli* K-12 MG1655 (60). In addition, deletion of *mlc* in *Δcrp* background and characterization of the binding affinity of Mlc in IG116 *Δcrp* strain revealed similar results as in the case of Fis (Supplementary Figure S5A & S6D). This emphasized the fact that regulation of *ptsG* is not dependent on the interplay of the regulators Fis and Mlc. Instead, a detailed analysis of the WGS data indicated that most of the mutations in the *ptsG* promoter resulted in new transcriptional start sites (TSS) with "Pribnow" box-like consensus sequence (Supplementary Figure S4B). A similar phenomenon was observed in a recent study albeit in a different genetic background (20). We suggest, the *ptsG* promoter mutations enhance *ptsG* gene expression by augmenting the RNA Polymerase sigma 70 affinity as seen from our RNA-seq data of the evolved strains described below.

Adaptive rewiring of gene expression of metabolic pathways in EvoCrp strains

A systems-level characterization at the transcriptome level was performed for the EvoCrp strains compared to Δ*crp* parent strain and statistically significant DE genes (aFC ≥ 2, adj-P < 0.05) were identified. Broadly, all EvoCrp strains compared to Δ*crp* showed fewer DE genes (ranging from ~75 to ~300 DE genes) as opposed to ~700 DE genes observed in Δ*crp* compared to WT (Figure 2A). A higher proportion of DE genes in EvoCrp showed an upregulation and a higher median expression level ($P < 10^{-6}$, Mann Whitney test, Figure 2B and Supplementary Figure S7 & S8) relative to Δ*crp*. To understand the adaptive response of EvoCrp towards growth recovery, we computed the correlation between the fold-change of gene expression in Δ*crp* vs WT and EvoCrp vs WT (Figure 2C & Supplementary Figure S9) and observed a strong positive correlation (Pearson correlation coefficient, $r = 0.74$, $P < 10^{-15}$). Also, the magnitude of difference of fold change between EvoCrp and WT was lesser compared to

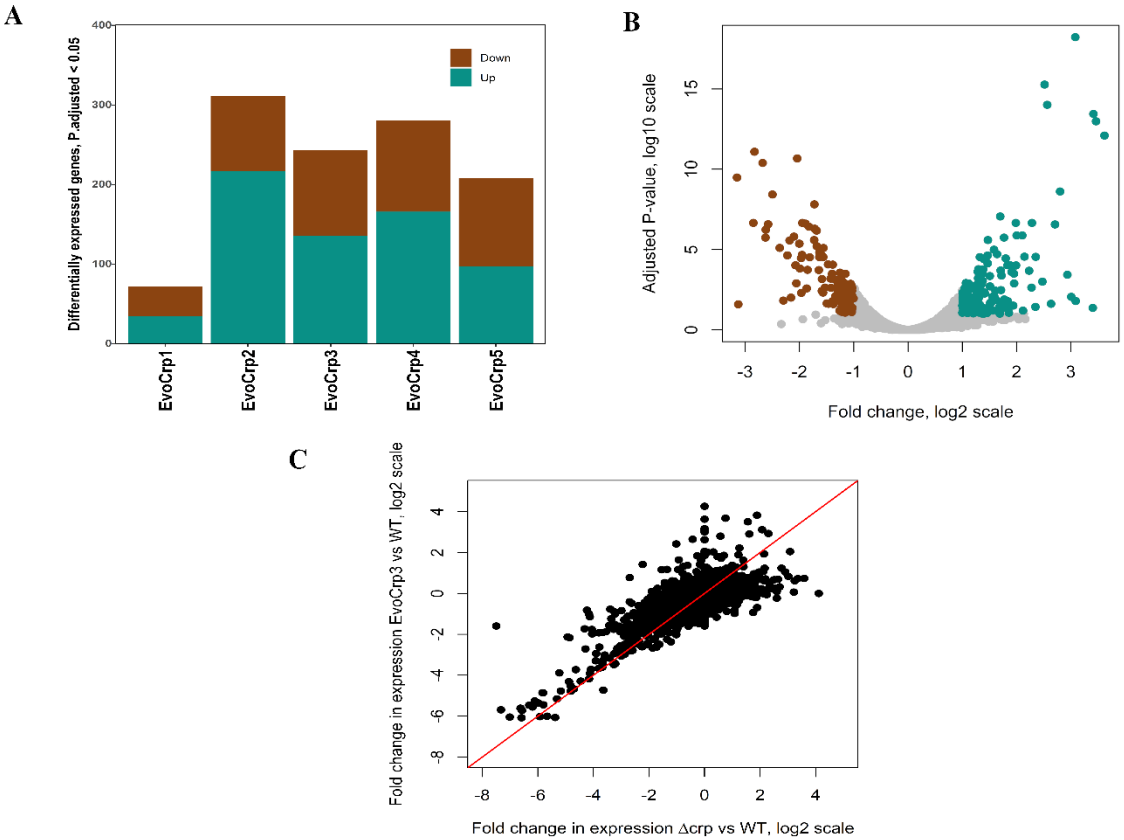

Figure 2. Transcriptome comparison of the EvoCrp compared to Δ*crp*. (A) Stacked plots showing the number of DE genes across all the EvoCrp strains. The downregulated genes are shown in brown and the upregulated genes are shown in cyan. (B) Volcano plot of the DE genes for one of the representative evolved strain, EvoCrp3 vs Δ*crp* depicted as Adj-P value (log10 scale) vs Fold change (log2 scale). See Supplementary Figure S7 for volcano plots of DE genes for other EvoCrp strains. (C) A correlation plot between EvoCrp3 vs WT and Δ*crp* vs WT (Pearson correlation coefficient, r = 0.74, $P < 10^{-15}$). See Supplementary Figure S9 for correlation plots of DE genes for other EvoCrp strains.

Δ*crp* vs WT ($P < 10^{-15}$, Mann Whitney test), thereby indicating a restoration in EvoCrp gene expression states towards WT levels.

Next, we investigated how EvoCrp strains were able to revert the perturbations in the metabolic pathways caused by the disruption of CRP. Overall, we observed that majority of the pathways that were affected in Δ*crp* were reverted to the WT state (Figure 3A-3B & Supplementary Figure S10-S14 for all other EvoCrp strains). For instance, across a majority of EvoCrp strains, we observed upregulation of "Transporters" and "tRNA loading", and significant downregulation of "Ribosome" or "Chaperone and folding catalyst" related pathways. This indicated restoration of protein turnover and folding machinery in EvoCrp strains. Notably, among transport genes, *ptsG* gene, showed an upregulation (aFC ~10) relative to its parent Δ*crp*, consistent with the enhanced promoter activity seen from mutation profiles. The restoration in

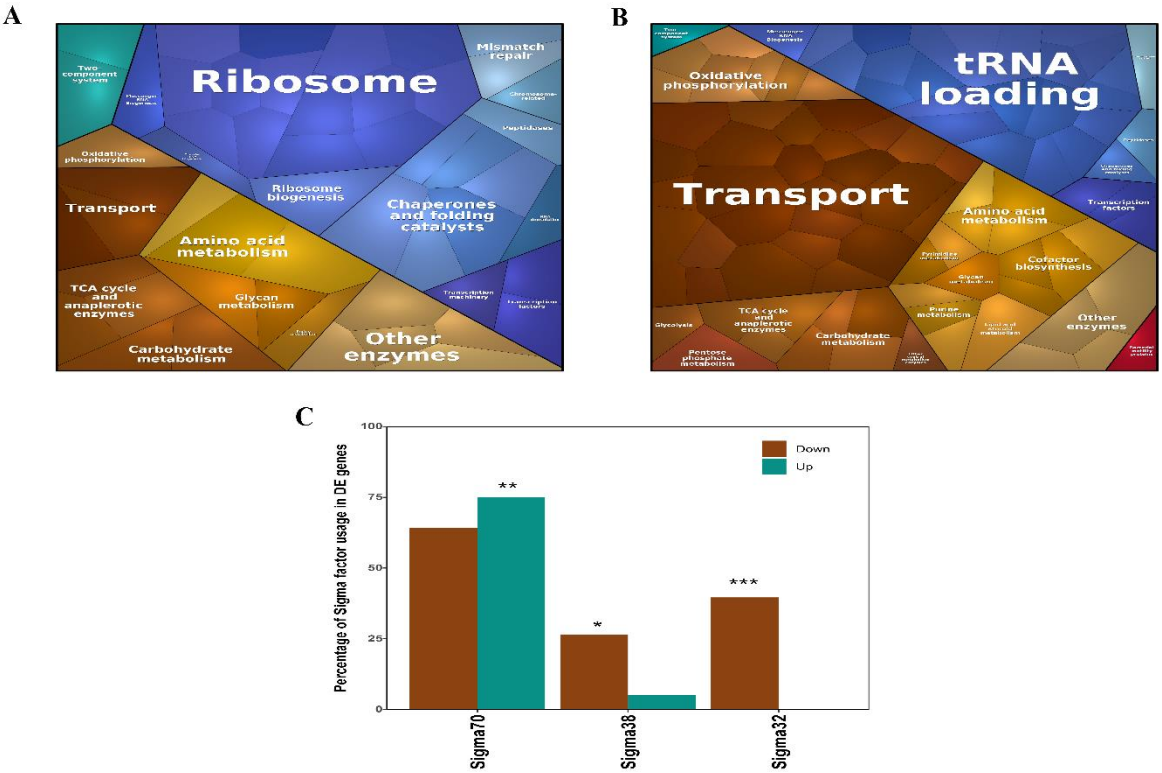

Figure 3. Enrichment analysis of the DE genes in EvoCrp3 compared to Δ*crp*. (A) The Voronoi maps showing the downregulated metabolic pathways enriched by KEGG classification. (B) The Voronoi treemaps showing the upregulated metabolic pathways enriched by KEGG classification. The genes enriched within each pathway are shown in Supplementary Figure S12. Note: The size of the hexagon within each category is directly proportional to the absolute fold change observed for each gene. The color of the hexagon denotes the specific pathways classified by KEGG. (C) Enrichment of genes under regulation of sigma factors; the brown bars and the cyan bars indicate the percentage of downregulated and upregulated genes in EvoCrp3 vs to Δ*crp* respectively. Significant increase or decrease is denoted by asterisks: one asterisk indicates $P < 0.05$, two asterisks indicate $P < 10^{-4}$ and three asterisks indicate $P < 10^{-10}$. See Supplementary Figure S15 for sigma factor usage of DE genes for other EvoCrp strains.

the gene expression in EvoCrp also showed a reallocation of RNA polymerase supported by the large fraction of the upregulated genes regulated by RNA polymerase sigma 70 (~76%, $P< 10^{-5}$) and the downregulated genes regulated by stress-related sigma factor, sigma 38 (~37%, $P < 10^{-3}$) and sigma 32 (~14%, $P <10^{-13}$) (Figure 3C & Supplementary Figure S15). Overall, in EvoCrp strains, we observed rewiring of metabolic pathways and protein biosynthesis machinery, which corroborated with reallocation of RNA polymerase favouring growth over stress related functions.

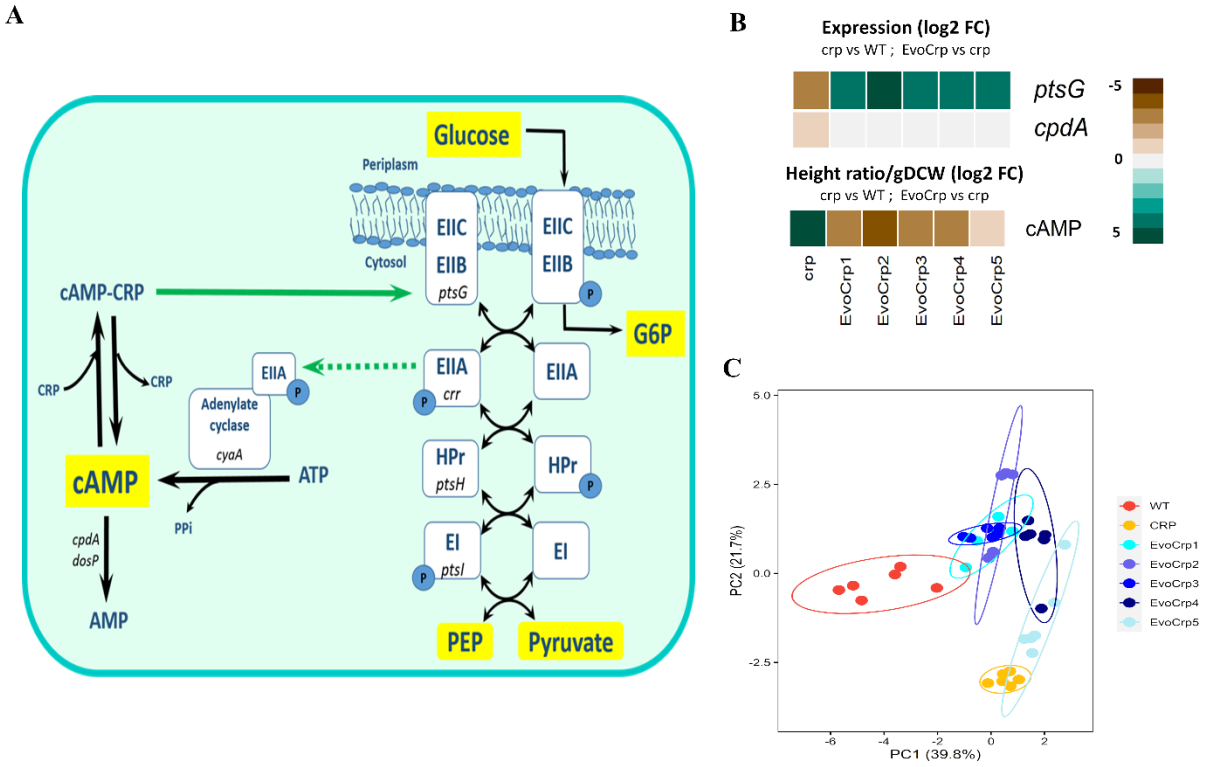

Figure 4. Fate of cAMP and global metabolome of the strains. (A) Network diagram of PEP-PTS system for glucose uptake with schematic representation of the regulation on cAMP by the enzymes of PEP-PTS. Key metabolites of the pathway are shown in yellow blocks. The green solid arrow indicates positive gene regulation from our study and the green dotted arrow indicates positive regulation of activity from literature evidence (64). (B) Expression profile of the DE genes of PEP-PTS with the metabolite levels of cAMP. Fold change values for gene expression and metabolite levels are obtained by comparing Δ*crp* vs WT and EvoCrp vs Δ*crp*. Expression values are obtained from average of two biological replicates expressed as log2 FC. Metabolite levels are obtained from average of three biological and two technical replicates expressed at height ratio/gDCW (log2 FC). (C) A multi-variate statistical analysis using PLS-DA for 24 significant metabolites found in our analysis. The six dots for each strain indicate the three biological and two technical replicates. The circles indicate 95% confidence intervals.

Evolutionary modulation of cAMP levels in absence of CRP

Glucose transport in *E. coli* is mediated via the *ptsG* component of the Phosphotransferase System (PTS) (64, 65). *E. coli* modulates its glucose uptake based on the limited pool size of the physiological signal molecule cAMP, an inducer of CRP activity (Figure 4A). We found that the deletion of CRP leads to a ~55-fold higher (adj-P < 10^{-9}) accumulation of cAMP compared to the WT (Figure 4B). Counterintuitively, we observed no change in *cyaA* gene expression, that generates cAMP from ATP (Figure 4A). Also, *cpdA* gene, responsible for degrading cAMP, had

lower expression in Δ*crp*. Thus, the absence of sequestration by CRP and reduced degradation resulted in the extremely high intracellular accumulation of cAMP in Δ*crp*. After evolution, we observed ~8-fold decrease in cAMP levels across all EvoCrp strains relative to Δ*crp* without any significant changes in *cyaA* and *cpdA* gene expression. This indicated an evolutionary restoration of cAMP levels that corroborated with optimal usage of ATP.

Integration analysis to decipher changes in transcriptome and metabolome of metabolic pathways

Next, we characterized other metabolites of central carbon metabolism perturbed due to loss of cAMP-CRP. We performed partial least squares discriminatory analysis (PLS-DA) on 24 metabolites that were found to be statistically significant (adj-P < 0.05, Student's t-test) in either EvoCrp vs Δ*crp,* EvoCrp vs WT or Δ*crp* vs WT strains. The two components of the PLS-DA scores plot (Figure 4C) which accounted for ~68% of the variance, showed restoration of the metabolite profiles in EvoCrp towards WT levels.

Integration of metabolite levels with its cognate gene expression profiles revealed perturbations majorly in key nodes of glycolysis and TCA cycle. In the glycolytic pathway (Figure 5A-B), 3-phosphoglycerate (3PG) was found to be higher in *Δcrp* (aFC ~ 3-fold, vs WT) and its level was reduced in EvoCrp strains (aFC ~ 2-fold, vs *Δcrp*). Further, we observed an increase in phosphoenolpyruvate (PEP) concentration in *Δcrp* compared to the WT (aFC ~ 8-fold), consistent with the reduced expression of *ptsG* gene. A recovery of PEP levels was observed across majority of the Evo*Crp* strains (aFC ~ 6-fold, vs *Δcrp*). PEP being a precursor of aromatic amino acids, a higher concentration in the *Δcrp* resulted in a concomitant increase in phenylalanine levels. Conversely, gene expression of aromatic amino acid biosynthesis genes namely *aroF* (aFC ~19), *tyrA* (aFC ~13) and *trpABCD* (aFC ~3) showed significant downregulation in *Δcrp* mutant. Such antagonism highlighted the known negative feedback regulation of amino acid biosynthesis (66). However, EvoCrp strains maintained similar levels of phenylalanine compared to *Δcrp,* despite lowered levels of PEP and upregulation of *aroF* gene expression (aFC >7). Pyruvate, an α-keto acid, is a precursor of the several branched-chain amino acids alanine, valine, leucine, and isoleucine (Figure 6A). We found a reduction in valine (aFC ~ 1.5-fold) and increase in leucine levels

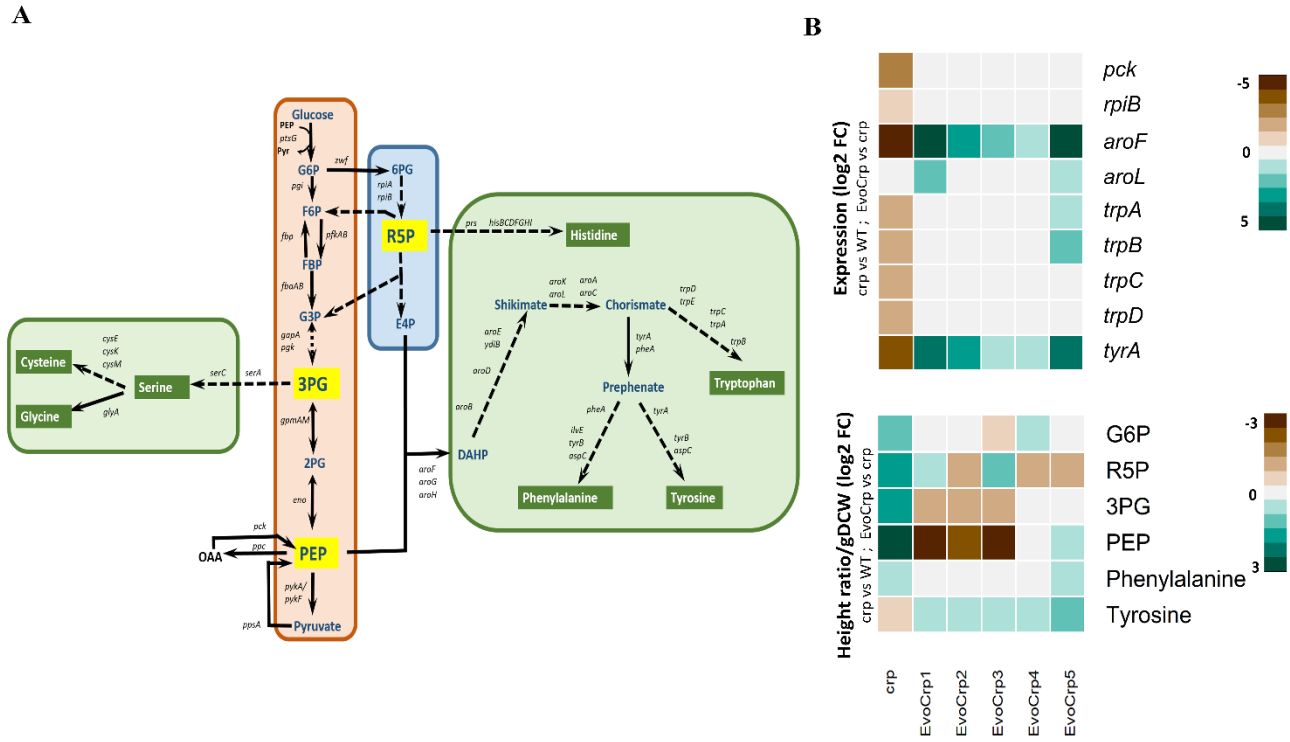

Figure 5. Integrated transcriptomics and metabolomics analysis at glycolytic pathway, Pentose Phosphate Pathway (PPP) node (A) A network diagram depicting glycolytic pathway, Pentose Phosphate Pathway (PPP) and amino acid biosynthetic pathways generating from the precursor metabolites of glycolytic and PPP pathway. (B) Expression profile of DE genes and metabolite levels significantly altered in the glycolytic and Pentose Phosphate Pathway (PPP) and amino acid biosynthetic pathways, were obtained by comparing Δ*crp* vs WT and EvoCrp vs Δ*crp*. Expression values are obtained from average of two biological replicates expressed as log2 FC. Metabolite levels are obtained from average of three biological and two technical replicates expressed as log2 FC of height ratio/gDCW.

(aFC ~ 2-fold) in *Δcrp* compared to WT. In contrast, we observed, reduction in valine (aFC ~ 2.3-fold, Figure 6D) and no significant changes in leucine in the EvoCrp strains compared to *Δcrp*.

We also determined the relative concentrations of citrate and other α-keto acids like alpha-ketoglutarate (αKG) and oxaloacetate (OAA), which are key intermediates of the TCA cycle (Figure 6B). In Δ*crp*, citrate level was lower (aFC ~1.75) compared to WT in agreement with the reduced gene expression of *gltA* (Figure 6C-D), whereas its levels were restored in EvoCrp

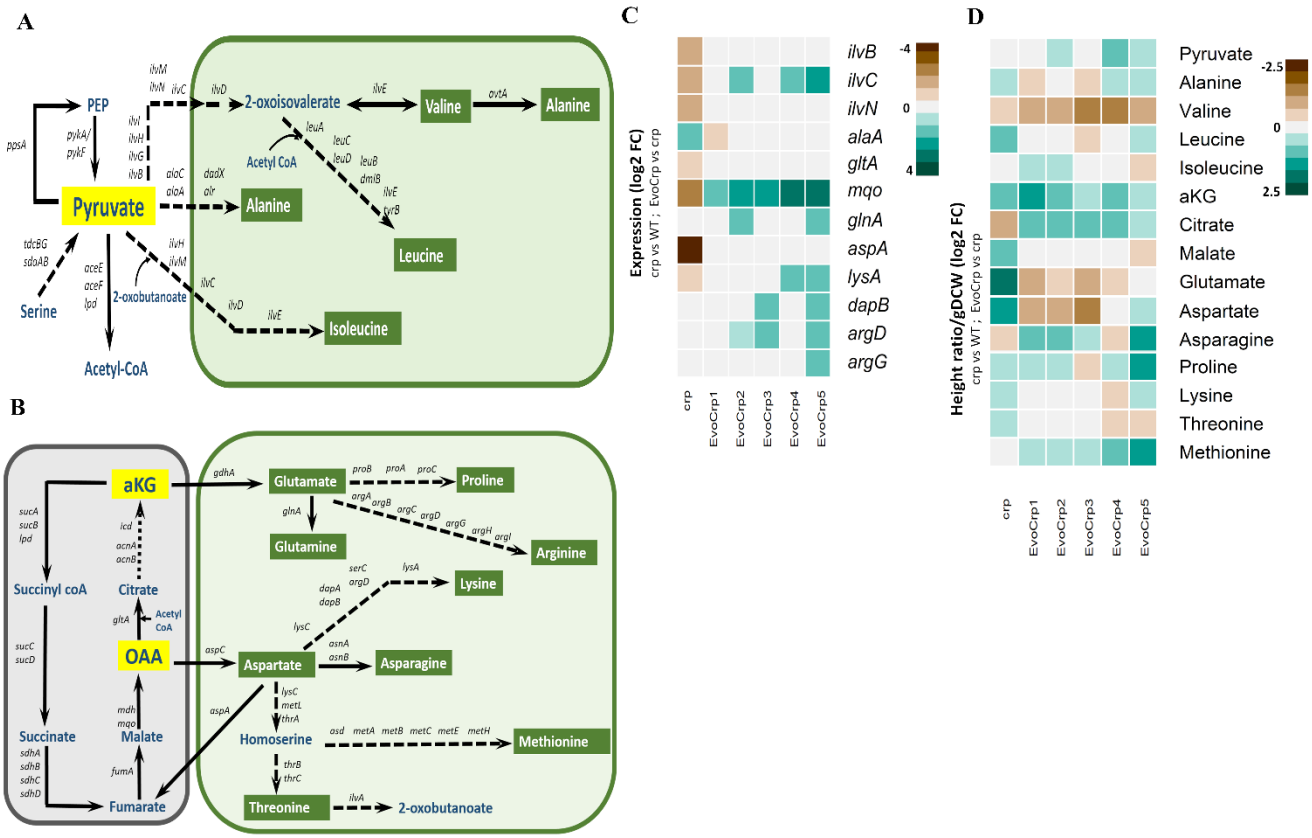

Figure 6. Integrated transcriptomics and metabolomics analysis at pyruvate and TCA cycle nodes (A) A network diagram depicting branch points from pyruvate node, and amino acid biosynthetic pathways generating from the precursor metabolite pyruvate. (B) A network diagram depicting branch points from TCA cycle and amino acid biosynthetic pathways generating from the precursor metabolites of TCA cycle. (C) Expression profile of DE genes of pyruvate node, TCA cycle and amino acid biosynthetic pathways, were obtained by comparing Δ*crp* vs WT and EvoCrp vs Δ*crp*. Expression values are obtained from average of two biological replicates expressed as log2 FC. (D) Metabolite levels perturbed at the pyruvate node, TCA cycle and, acid biosynthetic pathways, were obtained by comparing Δ*crp* vs WT and EvoCrp vs Δ*crp*. Metabolite levels are obtained from average of three biological and two technical replicates expressed as log2 FC of height ratio/gDCW.

strains. We observed a ~1.5-fold higher concentration of αKG in *Δcrp* compared to the WT as well as in EvoCrp compared to *Δcrp* (Figure 6D). αKG is a precursor for the amino acids glutamate, glutamine, arginine and proline. In *Δcrp* compared to WT, glutamate and proline levels were upregulated (aFC ~1.5 to 3) whereas, in the EvoCrp strains compared to *Δcrp*, we observed reduced levels of glutamate (aFC ~1.4 to 3) in contrast to higher levels of proline (aFC ~1.6). OAA, which serves as a precursor of amino acids like aspartate, asparagine, lysine, threonine, and methionine levels, can be interpreted from malate levels in the cell (67). The

higher OAA or malate (aFC ~2-fold) levels resulted in ~2-fold higher aspartate, ~1.3-fold higher levels of threonine and ~1.5-fold higher levels of lysine in Δ*crp* mutant compared to WT. Despite no changes in OAA/malate levels and ~1.75-fold reduction in aspartate concentration, we observed 1.6-fold higher asparagine and similar lysine levels in EvoCrp strains compared to *Δcrp*. Methionine, which showed no change in its concentration in *Δcrp* compared to the WT, was 1.6-fold high in all the EvoCrp strains compared to its parent *Δcrp*. To summarize, we observed accumulation of key proteinogenic amino acids like phenylalanine, alanine, aspartate, threonine, glutamate and lysine in *Δcrp* which coincided with the impaired protein biosynthesis machinery observed from gene expression analysis. Further, in the EvoCrp strains, we observed increased accumulation of costly amino acids (68) like phenylalanine, proline, lysine, asparagine and methionine, which were not restored to WT levels. Despite restoration of protein biosynthesis machinery, we observed the accumulation of these amino acids in EvoCrp suggestive of reallocation of ATP along with reduction in unnecessary cAMP synthesis.

Gene expression analysis of electron transport chain coupled to ATP synthesis

Our transcriptomic data showed significant changes in the expression levels of the genes of the electron transport chain of *E. coli* that is responsible for the transfer of electrons to oxygen coupled to the generation of ATP (Figure 7A & 7C). We observed downregulation of *cyoABCDE* genes (aFC~2) and upregulation of *appABC* genes (aFC ~4.5) in Δ*crp* compared to the WT. The *cyoABCDE* genes are induced by high levels of oxygen and encodes for the terminal oxidase cytochrome *bo*, the major route for oxygen channelling in aerobic respiratory condition (69). On the contrary, *appABC* genes code for the terminal oxidase cytochrome *bd* and are expressed under low oxygenation levels (69). In EvoCrp compared to Δ*crp*, we observed restoration in these gene expression profiles. Additionally, the genes of ATP synthase complex, *atpABCDEFGH*, that are responsible for synthesis of ATP, were downregulated in the Δ*crp* compared to the WT. However, in EvoCrp strains relative to Δ*crp*, we observed an upregulation that indicated restoration of the ATP synthesis machinery. Apart from ETC genes, we observed restoration in gene expression of membrane-bound transhydrogenases (*pntAB*) (Figure 7B-7C), which was in agreement with the rapid growth recovery of EvoCrp strains. *pntAB* has been previously suggested to be necessary for anabolic NADPH formation whose expression profile levels have a strong correlation with the growth rate of the organism (70, 71). To summarize,

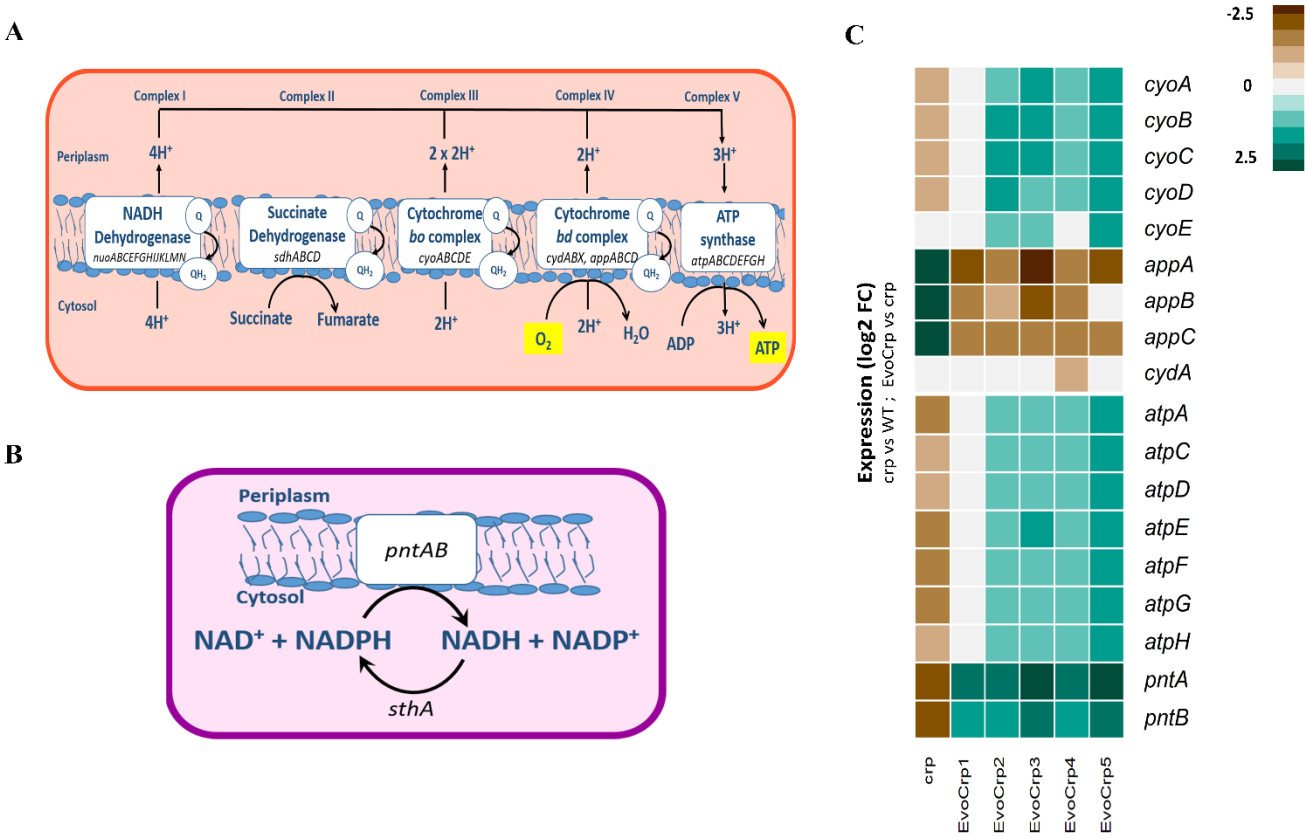

Figure 7. Gene expression data of oxidative phosphorylation and transhydrogenase genes. (A) A network diagram depicting the ETC and the membrane bound transhydrogenases. (B) Expression profile of DE genes of ETC and the transhydrogenase genes, obtained by comparing Δ*crp* vs WT and EvoCrp vs Δ*crp*. Expression values are obtained from average of two biological replicates expressed as log2 FC.

adaptive evolution resulted in restoration of oxidative phosphorylation and transhydrogenase expression profiles which are necessary for proper oxygen utilization and to balance the ATP and NADPH demand for anabolic processes in the evolved strains.

ME-model prediction reveals reallocation of proteome fractions

An inherent property of *E. coli* is to tightly coordinate the metabolism and protein economy in the cell towards an optimal resource allocation favouring growth fitness (3). To assess how EvoCrp strains were able to efficiently orchestrate the proteomic balance in the cell, we used a genome-scale ME-model (for Metabolism and Expression) (45). We predicted the "utilized ME" (uME) and "non-utilized ME" (nonME) proteome fractions specific to aerobic glucose metabolism in *E. coli* K-12 MG1655. Our data revealed that EvoCrp strains were able to efficiently upregulate their uME fraction (*ptsG* and oxidative phosphorylation genes) and

downregulate their nonME fraction (glyoxylate shunt, osmotic stress induced and amino acid degradation genes) to enhance their growth fitness. In contrast, we observed a reduction in uME fraction and increase in nonME fraction caused by deletion of CRP (Supplementary Figure S16). Also, lower fraction of TCA cycle genes and secondary glucose transporters (*man* and *mal* genes) were found to be consistent across all the EvoCrp strains, to mitigate the unnecessary proteome cost towards their synthesis. Overall, as an adaptive strategy to recover growth, the EvoCrp strains redirected the protein economy towards necessary processes.

Parallel populations converge to similar phenotypes at the end of ALE

Next, we characterized the growth rate, glucose uptake rate (GUR), acetate production rate (APR), oxygen uptake rate (OUR) and biomass yield which defines the phenotype of the strains (Supplementary Table S1). We observed a marked reduction in growth rate (~57%) in Δ*crp* strain compared to WT (Figure 8A). The GUR and OUR in *Δcrp* were both significantly reduced by ~56%, compared to the WT ($P < 0.05$, Student's t-test) (Figure 8B & 8C), which can be attributed to the reduction in gene expression of *ptsG* and oxidative phosphorylation, respectively. Similar trends were also observed in APR as well, despite no changes in its gene expression (Figure 8D). Further, we calculated pairwise correlation of growth rate with GUR (Pearson correlation coefficient, $r = 0.96$, $P < 10^{-3}$) and growth rate with OUR (Pearson correlation coefficient, $r = 0.92$). The strong correlation between them indicated that lowered GUR and lowered OUR might have resulted in an overall reduction of growth rate in the *Δcrp* strain (Supplementary Figure S17A-S17B). Conversely, we observed increase in growth rate (112%, $P < 0.05$; Figure 8A) as well as GUR (~130%, $P < 0.05$; Figure 8B) and APR (~115%, $P < 0.0\ 5$; Figure 8D) in all EvoCrp strains compared to the Δ*crp*. The OUR in EvoCrp showed variability in its increase compared to *Δcrp* ranging from ~99% to ~155% (Figure 8C). In summary, our data suggested that all the parallel populations converged to phenotypes similar to WT at the end of ALE within ~100 generations.

The biomass yield for each of the strains was determined by normalizing with their specific glucose uptake rate. While the WT and Δ*crp* strains showed similar biomass yields (0.41-0.42 gDCW/gGlucose), there was a consistent decrease in the biomass yield (~0.37 gDCW/gGlucose)

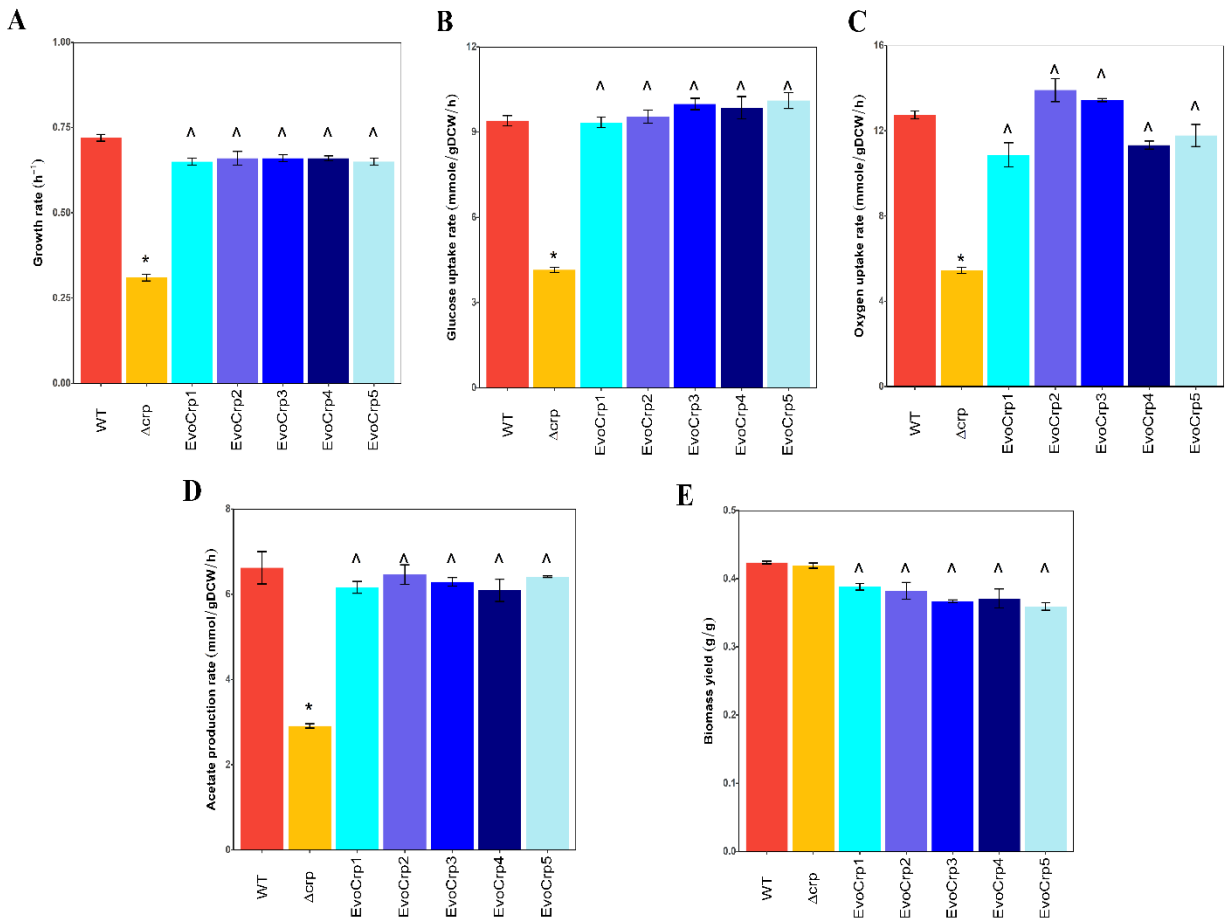

Figure 8. Phenomic studies of the strains. Bar-plot depicting the (A) growth rates, (B) glucose uptake rates, and (C) oxygen uptake rates (D) acetate production rates, and (E) biomass yields of the WT, Δ*crp* and the EvoCrp strains. The bars represent an average of growth rates obtained from three biological replicates. The error bars indicate the standard error across the replicates. Significance of decrease in Δ*crp* vs WT is shown by asterisk whereas significance of increase in EvoCrp vs Δcrp is shown by caret ($P < 0.05$, Student's t-test).

of EvoCrp strains that highlighted their inefficiency in optimally converting the carbon towards biomass (Figure 8E). Correlation of the measured biomass yield and growth rate showed (Pearson correlation coefficient, $r = -0.87$, $P < 0.05$, EvoCrp and Δ*crp*) the divergence of evolved strains away from Δ*crp* and WT (Supplementary Figure S17C). Since changes in biomass yields can be a consequence of alterations in ATP maintenance (ATPM, i.e., ATP used for processes other than growth and protein synthesis), we predicted ATPM yields for each of the strains (48, 72). The EvoCrp strains had a significantly higher ATPM compared to the WT as well as *Δcrp*, which was in agreement with the accumulation of costly amino acids (Supplementary Figure S17D). This phenomenon of lower energetic efficiency of carbon substrate utilization and

allocation towards growth revealed a rate-yield trade-off mechanism prevalent in ALE-adapted strains (73).

Discussion

We elucidate here the molecular events underlying impaired growth caused by deletion of CRP that delineates the coordination between the global regulator and several growth-rate dependent factors. In this study, we showed that adaptation involving mutations in *ptsG* promoter in response to loss of CRP takes place by three interdependent evolutionary strategies - higher rate of glucose import, restoration of protein biosynthesis machinery coupled with finely tuned proteome allocation and metabolic rewiring of ATP towards synthesis of costly amino acids along with reduction in wasteful cAMP synthesis. Specifically, this emphasized the indispensability of CRP regulation on genetic and metabolic machinery of the organism.

We illustrate that deletion of CRP caused significant decrease in the glucose uptake rate, which correlated strongly with decreased growth rate. During evolution, mutations occurred predominantly in the *ptsG* promoter that generated additional "Pribnow'-box like sequences to enhance the affinity of RNA polymerase towards *ptsG* gene. Notably, the mutations resulted in increased *ptsG* gene expression and thereby glucose uptake, emphasizing the role of *ptsG* promoter mutations in resolving the bottleneck caused by loss of CRP. Such phenomenon was also observed in a previous study, thereby implying genetic parallelism, despite the differences observed in genomic, transcriptomic and phenomic states across the two studies (20).

We additionally emphasize the regulatory control of CRP on protein biosynthesis machinery conducive for optimal growth. Biosynthesis of ribosomes required for protein translation, involves huge proteome investment. Recent studies have reported tight coordination of ribosomal protein expression with growth rate of the organism (3). For slow growing cells in stressed environmental conditions, the organism produces more ribosomes in order to hedge for unfavourable conditions (51). This in turn, limits the ribosomes available for synthesis for metabolic proteins manifested as increase in unnecessary ribosomal proteins and reduction in translational efficiency thereof (54).

Concurrently, our predicted system-wide proteome fraction of Δ*crp* mutant demonstrated a wasteful allocation of proteome towards stress or hedging functions. Evolution enabled the

strains to rebalance the proteome by increasing the growth-promoting proteome fraction and by mitigating the proteome cost associated with synthesis of expensive proteins, such as enzymes of TCA cycle. However, to meet the energetic demands of the faster growing strains, we suggest that the TCA cycle enzymes might have gained a higher *in-vivo* maximal rate, a phenomenon recurring in adaptive evolution (44, 74). These results emphasize the role of CRP in fine tuning the proteome allocation in response to its metabolic demands.

It is well known that growing cells maintain an optimal cAMP level necessary for proper carbon sensing and hence ATP optimization (75). We observed a high accumulation of cAMP in the absence of sequestration by CRP. On the contrary, evolved strains were able to restore their cAMP levels, thereby alleviating ATP wastage. However, we speculate that this surplus ATP pool conserved from reduced cAMP synthesis was invested towards synthesis of costly amino acids that was found to accumulate in the evolved strains. This might be partially responsible for the higher ATPM yield resulting in overall reduction in the biomass yields observed in the evolved strains.

In summary, our study shows that evolution enabled the strains to compensate for the lack of CRP and cAMP by reallocation of RNA polymerase and ATP albeit with inefficient utilization of carbon towards growth. Overall, we provide insights into the fundamental biological processes regulated by CRP that shape the growth physiology of *E. coli*.

Funding

This work was supported by DST fellowship/grant(SB/S3/CE/080/2015) and DBT fellowship/grant(BT/PF13713/BBE/117/83/2015) awarded to K.V.V., DBT/Wellcome Trust India Alliance Intermediate Fellowship/grant (grant number: IA/I/16/2/502711) awarded to A.S.N.S. and DST-WOSA grant (grant number:SR/WOS-A/ET-58/2017) awarded to S.S.

Data availability

The whole genome resequencing data presented this study are available at NCBI BioProject (https://www.ncbi.nlm.nih.gov/bioproject/) under accession no. PRJNA637503. The RNA sequencing data and the processed files from this study are available at NCBI Geo

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc) under accession no. GSE152214. The data obtained in this study will be accessible at the NIH Common Fund's NMDR (supported by NIH grant, U01-DK097430) website, the Metabolomics Workbench, https://www.metabolomicsworkbench.org.

Acknowledgements

We thank Jeffrey E. Barrick (Associate Professor, Department of Molecular Biosciences, The University of Texas at Austin) for his inputs in WGS and site-directed mutagenesis. We thank Edward Baidoo (Biochemist Research Scientist, Lawrence Berkley National Laboratory) for his valuable suggestions in metabolomics. We thank Awadhesh Pandit (Next Generation Genomics Facility, C-CAMP, NCBS) for providing genomics and transcriptomics services. We thank Dr. Mayuri Gandhi (Research Scientist, SAIF, IIT Bombay) for providing the LC-MS facility. We thank Prof. Pradeepkumar P. I. (IIT Bombay) for providing RT-PCR facility and Prof. Sarika Mehra (IIT Bombay) for providing vacuum concentrator for our use. We thank Colton Lloyd (Systems Biology Research Group, UCSD) for his inputs in ME-model simulations. We thank Reshma T. V. (NCBS) and Guhan K.A. (IIT Bombay) for their valuable technical inputs in ChIP-qPCR experiments. We also thank Dr. Pramod Rajaram S for his comments on the manuscript.

A.P. acknowledges Department of Science and Technology INSPIRE, Govt. of India for her fellowship (IF140914). M.S.I acknowledges Department of Biotechnology, Govt. of India for his fellowship (DBT/IIT-P/323).

References

1. Klumpp,S., Zhang,Z. and Hwa,T. (2009) Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, **139 (7)**, 1366–1375.

2. Berthoumieux,S., De Jong,H., Baptist,G., Pinel,C., Ranquet,C., Ropers,D. and Geiselmann,J. (2013) Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. *Mol. Syst. Biol.*, **9**, 634.

3. Scott,M., Gunderson,C.W., Mateescu,E.M., Zhang,Z. and Hwa,T. (2010) Interdependence of Cell Growth. *Science (80-.).*, **330 (6007)**, 1099–1102.

4. Gosset,G., Zhang,Z., Nayyar,S., Cuevas,W.A. and Saier,M.H. (2004) Transcriptome Analysis of Crp-Dependent Catabolite Control of Gene Expression in Escherichia coli. *J. Bacteriol.*, **186 (11)**, 3516–3524.

5. Nam,T.W., Park,Y.H., Jeong,H.J., Ryu,S. and Seok,Y.J. (2005) Glucose repression of the Escherichia coli sdhCDAB operon, revisited: Regulation by the CRP·cAMP complex. *Nucleic Acids Res.*, **33 (21)**, 6712–6722.

6. Castaño-Cerezo,S., Bernal,V., Blanco-Catalá,J., Iborra,J.L. and Cánovas,M. (2011) cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in Escherichia coli. *Mol. Microbiol.*, **82 (5)**, 1110–1128.

7. Kim,D., Seo,S.W., Gao,Y., Nam,H., Guzman,G.I., Cho,B.K. and Palsson,B.O. (2018) Systems assessment of transcriptional regulation on central carbon metabolism by Cra and CRP. *Nucleic Acids Res.*, **46 (6)**, 2901–2917.

8. Shimada,T., Fujita,N., Yamamoto,K. and Ishihama,A. (2011) Novel roles of camp receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS One*, **6 (6)**, e20081.

9. Franchini,A.G., Ihssen,J. and Egli,T. (2015) Effect of global regulators RpoS and cyclic-AMP/CRP on the catabolome and transcriptome of Escherichia coli K12 during carbon- and energy-limited growth. *PLoS One*, **10 (7)**, 1–25.

10. Basak,S. and Jiang,R. (2012) Enhancing E. coli Tolerance towards Oxidative Stress via Engineering Its Global Regulator cAMP Receptor Protein (CRP). *PLoS One*, **7 (12)**, e51179.

11. Martínez-Antonio,A. and Collado-Vides,J. (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr. Opin. Microbiol.*, **6 (5)**, 482–489.

12. LaCroix,R.A., Palsson,B.O. and Feist,A.M. (2017) A Model for Designing Adaptive Laboratory Evolution Experiments. *Appl. Environ. Microbiol.*, **83 (8)**, 1–14.

13. Notley-McRobb,L. and Ferenci,T. (2000) Experimental analysis of molecular events during mutational periodic selections in bacterial evolution. *Genetics*, **156**, 1493–1501.

14. Fong,S.S., Nanchen,A., Palsson,B.O. and Sauer,U. (2006) Latent pathway activation and increased pathway capacity enable Escherichia coli adaptation to loss of key metabolic enzymes. *J. Biol. Chem.*, **281 (12)**, 8024–8033.

15. Charusanti,P., Conrad,T.M., Knight,E.M., Venkataraman,K., Fong,N.L., Xie,B., Gao,Y. and Palsson,B.Ø. (2010) Genetic basis of growth adaptation of Escherichia coli after deletion of pgi, a major metabolic gene. *PLoS Genet.*, **6 (11)**, e1001186.

16. McCloskey,D., Xu,S., Sandberg,T.E., Brunk,E., Hefner,Y., Szubin,R., Feist,A.M. and Palsson,B.O. (2018) Evolution of gene knockout strains of E. coli reveal regulatory architectures governed by metabolism. *Nat. Commun.*, **9**, 3796.

17. Cooper,T.F., Remold,S.K., Lenski,R.E. and Schneider,D. (2008) Expression profiles reveal parallel evolution of epistatic interactions involving the CRP regulon in Escherichia coli. *PLoS Genet.*, **4 (2)**, e35.

18. Stoebel,D.M., Hokamp,K., Last,M.S. and Dorman,C.J. (2009) Compensatory evolution of gene regulation in response to stress by Escherichia coli lacking RpoS. *PLoS Genet.*, **5 (10)**, e1000671.

19. Srinivasan,R., Scolari,V.F., Lagomarsino,M.C. and Seshasayee,A.S.N. (2015) The genome-scale interplay amongst xenogene silencing, stress response and chromosome architecture in Escherichia coli. *Nucleic Acids Res.*, **43 (1)**, 295–308.

20. Lamrabet,O., Plumbridge,J., Martin,M., Lenski,R.E., Schneider,D. and Hindré,T. (2019) Plasticity of Promoter-Core Sequences Allows Bacteria to Compensate for the Loss of a Key Global Regulatory Gene. *Mol. Biol. Evol.*, **36 (6)**, 1121–1133.

21. Datsenko,K. a and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.*, **97 (12)**, 6640–6645.

22. Deatherage,D.E. and Barrick,J.E. (2014) Identification of mutations in laboratory evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol*, **1151**, 165–188.

23. Kim,J., Webb,A.M., Kershner,J.P., Blaskowski,S. and Copley,S.D. (2014) A versatile and highly efficient method for scarless genome editing in Escherichia coli and Salmonella enterica. *BMC Biotechnol.*, **14**, 84.

24. Uzzau,S., Figueroa-Bossi,N., Rubino,S. and Bossi,L. (2001) Epitope tagging of chromosomal genes in Salmonella. *Proc. Natl. Acad. Sci. U. S. A.*, **98 (26)**, 15264–15269.

25. Kahramanoglou,C., Seshasayee,A.S.N., Prieto,A.I., Ibberson,D., Schmidt,S., Zimmermann,J., Benes,V., Fraser,G.M. and Luscombe,N.M. (2011) Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli. *Nucleic Acids Res.*, **39**, 2073–2091.

26. Mukhopadhyay,A., Deplancke,B., Walhout,A.J.M. and Tissenbaum,H.A. (2008) Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in Caenorhabditis elegans. *Nat Protoc.*, **3 (4)**, 698–709.

27. Krishnamurthi,R., Ghosh,S., Khedkar,S. and Seshasayee,A.S.N. (2017) Repression of YdaS Toxin Is Mediated by Transcriptional Repressor RacR in the Cryptic rac Prophage of Escherichia coli K-12. *mSphere*, **2 (6)**, e00392-17.

28. Rio,D.C., Ares,M., Hannon,G.J. and Nilsen,T.W. (2010) Purification of RNA using TRIzol (TRI Reagent). *Cold Spring Harb. Protoc.*, **5**, 1–4.

29. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25 (14)**, 1754–1760.

30. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25 (16)**, 2078–2079.

31. Liao,Y., Smyth,G.K. and Shi,W. (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30 (7)**, 923–930.

32. Keseler,I.M., Mackie,A., Santos-Zavaleta,A., Billington,R., Bonavides-Martínez,C., Caspi,R., Fulcher,C., Gama-Castro,S., Kothari,A., Krummenacker,M., *et al.* (2017) The EcoCyc database: Reflecting new knowledge about Escherichia coli K-12. *Nucleic Acids Res.*, **45 (D1)**, D543–D550.

33. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26 (1)**, 139–140.

34. Minoru Kanehisa and Susumu Goto (2016) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.*, **28 (1)**, 27–30.

35. Liebermeister,W., Noor,E., Flamholz,A., Davidi,D., Bernhardt,J. and Milo,R. (2014) Visual account of protein investment in cellular functions. *Proc. Natl. Acad. Sci. U. S. A.*, **111 (23)**, 8488–8493.

36. Santos-Zavaleta,A., Salgado,H., Gama-Castro,S., Sánchez-Pérez,M., Gómez-Romero,L., Ledezma-Tejeida,D., García-Sotelo,J.S., Alquicira-Hernández,K., Muñiz-Rascado,L.J., Peña-Loredo,P., *et al.* (2019) RegulonDB v 10.5: Tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12. *Nucleic Acids Res.*, **47 (D1)**, D212–D220.

37. Fu,Y., Yoon,J.M., Jarboe,L. and Shanks,J. V. (2015) Metabolic flux analysis of Escherichia coli MG1655 under octanoic acid (C8) stress. *Appl. Microbiol. Biotechnol.*, **99 (10)**, 4397–4408.

38. Hollinshead,W.D., Rodriguez,S., Martin,H.G., Wang,G., Baidoo,E.E.K., Sale,K.L., Keasling,J.D., Mukhopadhyay,A. and Tang,Y.J. (2016) Examining Escherichia coli glycolytic pathways, catabolite repression, and metabolite channeling using Δpfk mutants. *Biotechnol. Biofuels*, **9**, 212.

39. Wu,L., Mashego,M.R., van Dam,J.C., Proell,A.M., Vinke,J.L., Ras,C., van Winden,W.A., van Gulik,W.M. and Heijnen,J.J. (2005) Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly 13C-labeled cell extracts as internal standards. *Anal. Biochem.*, **336 (2)**, 164–171.

40. Teleki,A., Sánchez-Kopper,A. and Takors,R. (2015) Alkaline conditions in hydrophilic interaction liquid chromatography for intracellular metabolite quantification using tandem mass spectrometry. *Anal. Biochem.*, **475**, 4–13.

41. Bennett,B.D., Kimball,E.H., Gao,M., Osterhout,R., Van Dien,S.J. and Rabinowitz,J.D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.*, **5 (8)**, 593–599.

42. Zhang,R., Watson,D.G., Wang,L., Westrop,G.D., Coombs,G.H. and Zhang,T. (2014) Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based

untargeted metabolite profiling of Leishmania parasites. *J. Chromatogr. A*, **1362**, 168–179.

43. Xia,J., Psychogios,N., Young,N. and Wishart,D.S. (2009) MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.*, **37**, 652–660.

44. O'Brien,E.J., Lerman,J.A., Chang,R.L., Hyduke,D.R. and Palsson,B. (2013) Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.*, **9:693**.

45. Lloyd,C.J., Ebrahim,A., Yang,L., King,Z.A., Catoiu,E., O'Brien,E.J., Liu,J.K. and Palsson,B.O. (2018) COBRAme: A computational framework for genome-scale models of metabolism and gene expression. *PLoS Comput. Biol.*, **14 (7)**, e1006302.

46. LaCroix,R.A., Sandberg,T.E., O'Brien,E.J., Utrilla,J., Ebrahim,A., Guzman,G.I., Szubin,R., Palsson,B.O. and Feist,A.M. (2015) Use of Adaptive Laboratory Evolution To Discover Key Mutations Enabling Rapid Growth of Escherichia coli K-12 MG1655 on Glucose Minimal Medium. *Appl. Environ. Microbiol.*, **81 (1)**, 17–30.

47. Wagner,G.P., Kin,K. and Lynch,V.J. (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.*, **131 (4)**, 281–285.

48. Utrilla,J., O'Brien,E.J., Chen,K., McCloskey,D., Cheung,J., Wang,H., Armenta-medina,D., Feist,A.M. and Palsson,B.O. (2016) Global Rebalancing of Cellular Resources by Pleiotropic Point Mutations Illustrates a Multi-scale Mechanism of Adaptive Evolution. *Cell Syst.*, **2 (4)**, 260–271.

49. Orth,J.D., Conrad,T.M., Na,J., Lerman,J.A., Nam,H., Feist,A.M. and Palsson,B.Ø. (2011) A comprehensive genome-scale reconstruction of Escherichia coli metabolism — 2011. *Mol. Syst. Biol.*, **7**, 535.

50. Ebrahim,A., Lerman,J.A., Palsson,B.O. and Hyduke,D.R. (2013) COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.*, **7**, 74.

51. Dai,X., Zhu,M., Warren,M., Balakrishnan,R., Okano,H., Williamson,J.R., Fredrick,K. and Hwa,T. (2018) Slowdown of translational elongation in Escherichia coli under hyperosmotic stress. *MBio*, **9**, 1–9.

52. Jozefczuk,S., Klie,S., Catchpole,G., Szymanski,J., Cuadros-Inostroza,A., Steinhauser,D., Selbig,J.

and Willmitzer,L. (2010) Metabolomic and transcriptomic stress response of Escherichia coli. *Mol. Syst. Biol.*, **6**, 364.

53. Hui,S., Silverman,J.M., Chen,S.S., Erickson,D.W., Basan,M., Wang,J., Hwa,T. and Williamson,J.R. (2015) Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol. Syst. Biol.*, **11**, e784–e784.

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

55. Grainger,D.C., Hurd,D., Harrison,M., Holdstock,J. and Busby,S.J.W. (2005) Studies of the distribution of Escherichia coli cAMP-receptor protein and RNA polymerase along the E. coli chromosome. *Proc. Natl. Acad. Sci. U. S. A.*, **102 (49)**, 17693–17698.

56. Browning,D.F. and Busby,S.J.W. (2004) The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, **2 (1)**, 57–65.

57. Hollands,K., Busby,S.J.W. and Lloyd,G.S. (2007) New targets for the cyclic AMP receptor protein in the Escherichia coli K-12 genome. *FEMS Microbiol. Lett.*, **274 (1)**, 89–94.

58. Latif,H., Federowicz,S., Ebrahim,A., Tarasova,J., Szubin,R., Utrilla,J., Zengler,K. and Palsson,B.O. (2018) ChIP-exo interrogation of Crp, DNA, and RNAP holoenzyme interactions. *PLoS One*, **13 (5)**, e0197272.

59. Shin,D., Cho,N., Heu,S. and Ryu,S. (2003) Selective regulation of ptsG expression by Fis: Formation of either activating or repressing nucleoprotein complex in response to glucose. *J. Biol. Chem.*, **278 (17)**, 14776–14781.

60. Cho,B.K., Knight,E.M., Barrett,C.L. and Palsson,B. (2008) Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. *Genome Res.*, **18 (6)**, 900–910.

61. Plumbridge,J. and Upr,I.D.B.P. (1998) Expression of ptsG , the gene for the major glucose PTS transporter in Escherichia coli , is repressed by Mlc and induced by growth on glucose. **29 (4)**, 1053–1063.

62. Plumbridge,J. (2002) Regulation of gene expression in the PTS in Escherichia coli : the role and interactions of Mlc. **5**, 187–193.

63. Jeong,J.Y., Kim,Y.J., Cho,N., Shin,D., Nam,T.W., Ryu,S. and Seok,Y.J. (2004) Expression of ptsG

encoding the major glucose transporter is regulated by ArcA in Escherichia coli. *J. Biol. Chem.*, **279 (37)**, 38513–38518.

64. Deutscher,J., Francke,C., Postma,P.W., Francke,C. and Postma,P.W. (2006) How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiol. Mol. Biol. Rev*, **70 (4)**, 939–1031.

65. Postma,P.W., Lengeler,J.W. and Jacobson,G.R. (1993) Phosphoenolpyruvate : Carbohydrate Phosphotransferase Systems of Bacteria. **57 (3)**, 543–594.

66. Reznik,E., Christodoulou,D., Goldford,J.E., Briars,E., Sauer,U., Segrè,D. and Noor,E. (2017) Genome-Scale Architecture of Small Molecule Regulatory Networks and the Fundamental Trade-Off between Regulation and Enzymatic Activity. *Cell Rep.*, **20 (11)**, 2666–2677.

67. Yuan,J., Doucette,C.D., Fowler,W.U., Feng,X.J., Piazza,M., Rabitz,H.A., Wingreen,N.S. and Rabinowitz,J.D. (2009) Metabolomics-driven quantitative analysis of ammonia assimilation in E. coli. *Mol. Syst. Biol.*, **5 (302)**, 1–16.

68. Neidhardt,F.C., Ingraham,J.L. and Schaechter,M. (1991) Physiology of the bacterial cell — A molecular approach. *Trends Genet.*, **7 (10)**, 341.

69. Cotter,P.A. and Gunsalus,R.P. (1992) Contribution of the fnr and arcA gene products in coordinate regulation of cytochrome o and d oxidase (cyoABCDE and cydAB) genes in Escherichia coli. *FEMS Microbiol. Lett.*, **91**, 31–36.

70. Sauer,U., Canonaco,F., Heri,S., Perrenoud,A. and Fischer,E. (2004) The Soluble and Membrane-bound Transhydrogenases UdhA and PntAB Have Divergent Functions in NADPH Metabolism of Escherichia coli *. *J. Biol. Chem.*, **279**, 6613–6619.

71. Haverkorn van Rijsewijk,B.R.B., Kochanowski,K., Heinemann,M. and Sauer,U. (2016) Distinct transcriptional regulation of the two Escherichia coli transhydrogenases PntAB and UdhA. *Microbiol. (United Kingdom)*, **162**, 1672–1679.

72. Pirt,S.J. (1982) Maintenance energy: a general model for energy-limited and energy-sufficient growth. *Arch. Microbiol.*, **133**, 300–302.

73. Cheng,C., O'Brien,E.J., McCloskey,D., Utrilla,J., Olson,C., LaCroix,R.A., Sandberg,T.E., Feist,A.M., Palsson,B.O. and King,Z.A. (2019) Laboratory evolution reveals a two-dimensional

rate-yield tradeoff in microbial metabolism. *PLoS Comput. Biol.*, **15 (6)**, e1007066.

74. O'Brien,E.J., Utrilla,J. and Palsson,B.O. (2016) Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments. *PLoS Comput. Biol.*, **12 (6)**, 1–22.

75. You,C., Okano,H., Hui,S., Zhang,Z., Kim,M., Gunderson,C.W., Wang,Y.P., Lenz,P., Yan,D. and Hwa,T. (2013) Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, **500**, 301–306.