Research in Microbiology 173 (2022) 103968

Contents lists available at ScienceDirect

Research in Microbiology

journal homepage: www.elsevier.com/locate/resmic

Brief Note

Proteomic profiling of the carbon-starved *Escherichia coli* reveals upregulation of stress−inducible pathways implicated in biological adhesion and methylglyoxal metabolism

Rakhan Aimbetov [a, *], Vasily Ogryzko [b]

[a] Nazarbayev University, 53 Kabanbay batyr avenue, 010000 Nur-Sultan, Kazakhstan
[b] UMR 8126, Institut Gustave Roussy, 114 rue Edouard Vaillant, 94805 Villejuif, France

ARTICLE INFO

Article history:
Received 22 March 2022
Accepted 15 June 2022
Available online 20 June 2022

Keywords:
Escherichia coli
Starvation
Mass spectrometry
Proteomics
Bacterial adhesion
Pyruvaldehyde (methylglyoxal)

ABSTRACT

Starvation in bacteria is a complex adaptive response to deprivation of nutrients that has been shown to implicate a number of stress networks that modulate pathogenicity and antibiotic resistance. Starvation in nature is qualitatively different from in-culture late stationary phase energy source depletion. To look into proteome-level alterations elicited by complete elimination of carbon source, we used *Escherichia coli* HT115-derived SLE1 strain cells and a combination of label-free and metabolic isotope labeling approaches. We isolated pathways differentially affected by carbon starvation and observed robust upregulation of proteins implicated in networks belonging to Gene Ontology terms *'Biological adhesion'* and *'Methylglyoxal metabolic process'*.

© 2022 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

Continuous development of mass spectra acquisition tools, in parallel with sophisticated bioinformatical data analysis solutions, has ensured the popularity of mass spectrometry as a powerful technology for protein identification and quantification. Stable isotope labeling by amino acids in culture (SILAC) is a preferred approach when it comes to quantitative evaluation of relative protein abundance based on metabolic incorporation of 'light' or 'heavy' versions of lysine and arginine into nascent polypeptides [1]. Label-free quantification (LFQ), on the other hand, is a cheaper and less laborious technique that has recently been getting more attention due to its virtual universality [2].

Bacterial starvation response is a complex phenotype of growth retardation, reduced viability, and overall switching of cellular metabolic machinery to a more robust energy-saving mode. Given the implication of stress-induced pathways in the modulation of pathogenicity [3] and antibiotic resistance [4], the study of proteome-level changes evoked by starvation possesses a significant theoretical and practical interest. In order to gather insight into the starved bacterial cell, we decided to assess the qualitative and quantitative impact of carbon starvation utilizing the combination of LFQ and SILAC methodologies.

2. Materials and methods

2.1. Strain and culture conditions

We used HT115-derived *Escherichia coli* strain SLE1 auxotrophic for arginine and lysine [5]. The cells were grown in M9 minimal medium supplemented with 0.3 mM of either $^{12}C_6$-lysine/$^{12}C_6^{14}N_4$-arginine ('light') or $^{13}C_6$-lysine/$^{13}C_6^{15}N_4$-arginine ('heavy') amino acids, at 37°C and 150 rpm. Carbon starvation was achieved by incubation of the cells in the medium devoid of amino acids and glucose for 48 h (Fig. 1A, Supplementary file).

2.2. Sample preparation

The cells were collected by centrifugation, washed once with cold PBS, and lysed in 1X LDS loading buffer (Novex). Following the estimation of protein concentration, equal quantities of protein, typically ≤400 μg, were processed in accordance with the FASP protocol [6]. Eluted peptides were desalted on Vivapure C18 micro spin columns (Sartorius Stedim Biotech), desiccated in SpeedVac, and dissolved in 10 μL of LC buffer A (0.1% formic acid in water).

* Corresponding author.
E-mail address: r.aimbetov@gmail.com (R. Aimbetov).

https://doi.org/10.1016/j.resmic.2022.103968
0923-2508/© 2022 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

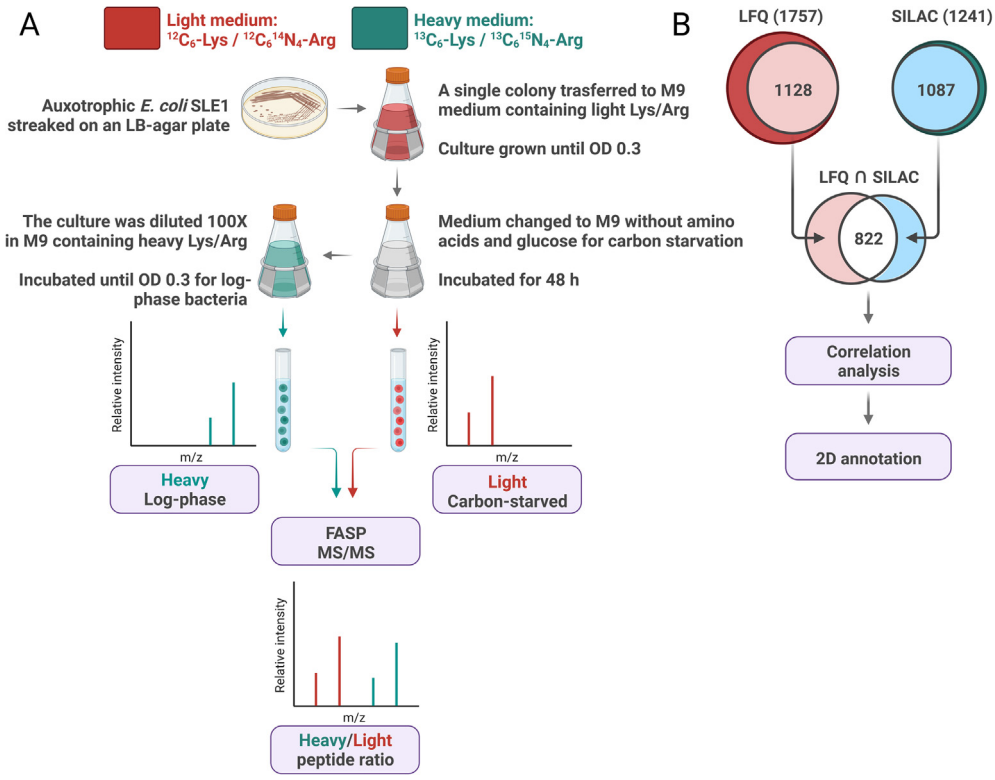

Fig. 1. A. The experimental workflow of the study. The *Escherichia coli* SLE1 strain cells were grown in M9 medium supplemented with 'light' lysine/arginine amino acids and glucose as a carbon source until the culture reached OD600 0.3. The cells then were transferred to and incubated in M9 devoid of nutrients for 48 h. On the day of the harvest, the portion of the starved culture was diluted 100 times in M9 containing 'heavy' lysine/arginine supplement, as well as glucose, and was allowed to grow until OD600 0.3. The two cultures were harvested simultaneously, and the cells were processed as described in Materials and methods. For label-free analysis, the pre-harvest starved cells were diluted into M9 containing 'light' amino acids and glucose as a source of carbon. **B. The statistical data analysis workflow of the study.** Within 1757 LFQ and 1241 SILAC protein groups, 1128 and 1087 had quantifiable ratios respectively, of which 822 were present in both datasets. These 822 groups were used for correlation calculation between the quantification techniques and 2D Gene Ontology annotation enrichment. The figure was created with BioRender.com.

2.3. Mass spectra acquisition

LC/MS analysis was performed on EASY-nLC 1000 (Thermo Scientific) paired with Q Exactive quadrupole-orbitrap hybrid mass spectrometer (Thermo Scientific). The peptide mixture was separated on EASY-Spray 15 cm × 75 μm 3 μm 100 Å C18 PepMap® reverse-phase column (Thermo Scientific) using 150 min three-step water-acetonitrile gradient (0–120 min, 5 → 35% LC buffer B (0.1% formic acid in acetonitrile); 120–140 min, 35 → 50%; 140–145 min, 50 → 90%; hold for 5 min) at 300 nL/min flowrate. The intensities of precursor ions were gauged in positive mode at scan range 400–2000 *m/z*, resolution 70,000, automatic gain control (AGC) target 1E6, maximum injection time 100 ms, followed by forwarding 10 most intense ions of a spectrum for MS2 fragmentation and measurement at resolution 17,500, AGC target 5E4, maximum injection time 100 ms, isolation window 2 *m/z* with 30 s dynamic exclusion.

2.4. Discovery analysis

Raw mass spectrometric data were analyzed by Proteome Discoverer v.1.4.0.288 (Thermo Scientific). MS2 spectra were searched against the *E. coli* Swiss-Prot database using Mascot engine set for 10 ppm precursor mass and 0.02 Da fragment mass tolerances with 2 allowed missed cleavage sites. For labeled samples, amino acid modifications were as follows: $^{13}C_6$-lysine (+6.020129 Da) and $^{13}C_6^{15}N_4$-arginine (+10.008269 Da) SILAC labels, methionine oxidation (+15.994915 Da) as dynamic, cysteine carbamidomethylation (+57.021464 Da) as static. For unlabeled samples: methionine oxidation and asparagine/glutamine deamidation (+0.984016 Da) as dynamic, cysteine carbamidomethylation as static. False discovery rate was calculated using Percolator [7] with 0.01 strict and 0.05 relaxed target cut-off values.

2.5. Protein quantification

Label-free comparative protein quantitation was carried out in Sieve v.2.1.377 (Thermo Scientific). Total ion current alignment was done on a 5–120 min segment with 2 min retention time (RT) shift limit. Framing was performed on a 400–2000 *m/z* range with RT and *m/z* widths equal to 2.5 min and 10 ppm respectively, while the *'Frames from MS2 scans'* option was assigned a *TRUE* value. Protein IDs were imported from Proteome Discoverer. SILAC H/L ratios were determined using Proteome Discoverer's Precursor Ions Quantifier node with experimental bias normalization based on at least 20 protein counts.

2.6. Data

Data wrangling and analysis were performed in Jupyter Notebook using Python 3.6. Protein annotation and Gene Ontology term enrichment were done in Perseus 2.0.3.0 [8].

All raw data files with the accompanying result output have been uploaded to the ProteomeXchange Consortium repository (http://www.proteomexchange.org/) under the dataset identifier PXD003255 (https://doi.org/10.6019/pxd003255).

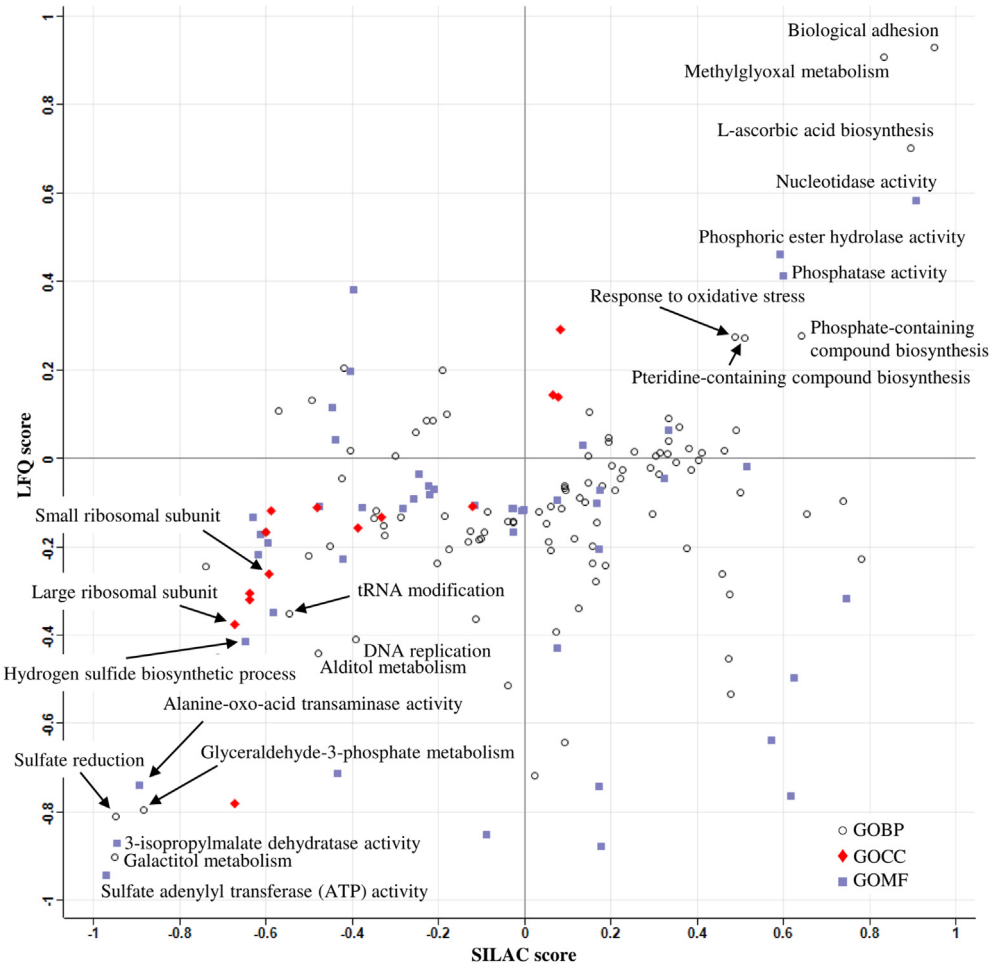

Fig. 2. Two-dimensional GO term annotation enrichment of SILAC- (x) and LFQ-quantified (y) proteins. *Scatter plot*, GO terms plotted against *x* and *y* according to their correlation *s*-score. *GOBP*, Gene Ontology biological process category. *GOCC*, Gene Ontology cellular compartment category. *GOMF*, Gene Ontology molecular function category.

3. Results

The goal of the present project was to quantitatively compare proteomes of carbon-starved and late log-phase exponentially growing cells. The combination of LFQ and SILAC permits the usage of these two methods as validators of each other and consequently allows the confident identification of differentially regulated proteins and networks while weeding out inevitable 'flukes' in detection.

In order to tailor experimental conditions to the selected bacterial strain, we performed direct plating experiments, i.e. a dilution of starved incubation culture was plated out onto LB-agar medium at defined time points with the subsequent enumeration of the colonies grown. As a result, we have established that 48 h is a threshold after which the cells start to lose their viability (Supplementary file).

The MS analysis of label-free samples resulted in 130,526 and 118,455 spectra for exponential and starved cells, which matched 8603 and 7769 high-confidence peptides assigned to 1602 and 1437 protein groups (1757 in total), respectively (Dataset S1). For labeled samples, 125,220 spectra, matched with 6037 high-confidence peptides, allowed to identify 1241 protein groups (Dataset S2). 1128 and 1087 protein groups identified by the label-free and SILAC methods respectively were quantifiable; 822 were present in both datasets (Fig. 1B, Data analysis file).

The abundance ratios of the protein groups common for both sets were $\log_2$- and *z*-transformed and plotted for correlation analysis. The Pearson's correlation coefficient between the ratios obtained using the two methods was equal to 0.63, $p = 3.45e\text{-}91$ (Data analysis file).

Gene Ontology (GO) is a comprehensive vocabulary of genes' and gene products' functional descriptions arranged into categories and terms [9]. In order to identify metabolic pathways affected by carbon starvation, we took advantage of Perseus' two-dimensional annotation feature and compared LFQ and SILAC datasets isolating GO terms with the highest enrichment and correlation score [10]. The 1D-annotated identifiers were partitioned into LFQ and SILAC data dimensions at a significance threshold $p < 0.1$ (Fig. 2). The corresponding *p*-values and Benjamini-Hochberg false discovery rates are provided in Table S1.

Shapiro–Wilk tests and qq plots showed that the LFQ and SILAC data are non-normally distributed (Data analysis file). However, given a large number of datapoints ($n = 822$), we could process the data following the Central Limit Theorem and build confidence intervals as if the data were distributed normally [11].

Significantly up- and downregulated proteins were defined at $p < 0.01$ as elements of the LFQ ∩ SILAC intersection satisfying $x_i > \mu + (2.58 \times \text{standard deviation (SD)}/\sqrt{n})$ and $x_i < \mu - (2.58 \times \text{SD}/\sqrt{n})$, respectively. Additionally, the ratios had to be regulated in a biologically sensible way, i.e. at least 2-fold up- or

downregulated. As a result, the statistical analysis of the 822 protein groups present in both datasets led to the identification of 27 upregulated and 71 downregulated entries with 99% confidence (Data analysis file, Tables S2 and S3).

4. Discussion

For most microbes in their natural habitats, starvation is a norm while a state of satiety is a seldom event [12]. Consequently, reports describing microbial features in nutrient-enriched laboratory environments do not fully reflect the cellular molecular dynamics with respect to the cell's natural surrounding [13]. Given the implication of starvation response, on par with other frequent naturally occurring stresses, in the modulation of pathogenicity and antibiotic resistance in bacteria, the study of '-omic' changes secondary to prolonged periods of nutrients deficiency is a task of medical and environmental importance.

As a result of our investigation, we identified a number of networks impacted by carbon starvation (Table S1). Due to the word constraints, we cannot discuss all of them. However, we would like to pinpoint several, which, in our opinion, reflect the changes that have occurred in the most dramatic fashion.

Nutritional scarcity leads to rapid expression of survival-promoting stress modulators, most notably an integration host factor (IHF) [14], an alternative RNA polymerase sigma factor RpoS [15], and the guanosine tetraphosphate alarmone (ppGpp) [16].

One of the most prominent hallmarks of starvation in bacteria is a so-called 'stringent response', a state, regulated by ppGpp, of severe diminution of de novo protein synthesis and intensified turnover of pre-existing proteins [17]. In our analysis, conspicuous underrepresentation of *'Small ribosomal subunit'* and *'Large ribosomal subunit'* cellular compartment terms (GOCC) conveys an overall downshift in protein synthesis and downregulation of ribosomal proteins in carbon-starved cells.

Starved bacterial cells display a tendency towards grouping. The formation of cellular aggregates, such as biofilms, has been suggested as a protection mechanism against antibiotics [18]. The most overrepresented GO term in our study was *'Biological adhesion'*, which is in line with numerous reports of various stresses converging on cellular adhesion in bacteria.

The second most upregulated GO term in our research was *'Methylglyoxal metabolic process'*. As reviewed in [19], methylglyoxal is a toxic electrophile produced as a byproduct of glycolysis that attacks nucleophilic centers of biological molecules such as amino acids and nucleotides. Although posing a threat to long-term viability, the production of methylglyoxal may serve as a short-term adaptive strategy aimed at colonizing the bacterial environment under low glucose conditions. Moreover, the experimental elevation of methylglyoxal production has led to overexpression of RpoS and associated stress networks [20] highlighting the convergence of various stress modulators on common effectors and pathways.

5. Conclusion

We have performed a broad analysis of the bacterial proteome under carbon starvation identifying 27 upregulated and 71 downregulated protein groups with 99% confidence. The 2D-annotation analysis identified pathways differentially affected by carbon starvation such as Gene Ontology terms *'Biological adhesion'*, *'Methylglyoxal metabolic process'*, and various other stress-related networks. The raw data of the present study is publicly available for inquiries through the ProteomeXchange Consortium repository.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

RA would like to thank Dr. Ali Burak Özkaya for reading the manuscript and providing valuable input. This work was supported by Fondation ARC pour la Recherche sur le Cancer subvention no. PDF20130606592 to RA and VO.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.resmic.2022.103968.

References

[1] Hoedt E, Zhang G, Neubert TA. Stable isotope labeling by amino acids in cell culture (SILAC) for quantitative proteomics. Adv Exp Med Biol 2019;1140:531–9. https://doi.org/10.1007/978-3-030-15950-4_31.
[2] Megger DA, Bracht T, Meyer HE, Sitek B. Label-free quantification in clinical proteomics. Biochim Biophys Acta 2013;1834:1581–90. https://doi.org/10.1016/j.bbapap.2013.04.001.
[3] Fang FC, Libby SJ, Buchmeier NA, Loewen PC, Switala J, Harwood J, et al. The alternative sigma factor katF (rpoS) regulates Salmonella virulence. Proc Natl Acad Sci U S A 1992;89:11978–82. https://doi.org/10.1073/pnas.89.24.11978.
[4] Poole K. Stress responses as determinants of antimicrobial resistance in Gram-negative bacteria. Trends Microbiol 2012;20:227–34. https://doi.org/10.1016/j.tim.2012.02.004.
[5] Larance M, Bailly AP, Pourkarimi E, Hay RT, Buchanan G, Coulthurst S, et al. Stable-isotope labeling with amino acids in nematodes. Nat Methods 2011;8:849–51. https://doi.org/10.1038/nmeth.1679.
[6] Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Methods 2009;6:359–62. https://doi.org/10.1038/nmeth.1322.
[7] Brosch M, Yu L, Hubbard T, Choudhary J. Accurate and sensitive peptide identification with Mascot Percolator. J Proteome Res 2009;8:3176–81. https://doi.org/10.1021/pr800982s.
[8] Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods 2016;13:731–40. https://doi.org/10.1038/nmeth.3901.
[9] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–9. https://doi.org/10.1038/75556.
[10] Cox J, Mann M. 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. BMC Bioinf 2012;13(Suppl 16):S12. https://doi.org/10.1186/1471-2105-13-S16-S12.
[11] Pek J, Wong ACM, Wong OCY. Confidence intervals for the mean of non-normal distribution: transform or not to transform. Open J Stat 2017;7:405–21. https://doi.org/10.4236/ojs.2017.73029.
[12] Kolter R, Siegele DA, Tormo A. The stationary phase of the bacterial life cycle. Annu Rev Microbiol 1993;47:855–74. https://doi.org/10.1146/annurev.mi.47.100193.004231.
[13] Finkel SE. Long-term survival during stationary phase: evolution and the GASP phenotype. Nat Rev Microbiol 2006;4:113–20. https://doi.org/10.1038/nrmicro1340.
[14] Nystrom T. Glucose starvation stimulon of escherichia-coli - role of integration host factor in starvation survival and growth phase-dependent protein-synthesis. J Bacteriol 1995;177:5707–10. https://doi.org/10.1128/jb.177.19.5707-5710.1995.
[15] Battesti A, Majdalani N, Gottesman S. The RpoS-mediated general stress response in Escherichia coli. In: Gottesman S, Harwood CS, editors. Annu Rev Microbiol, 65. Palo Alto: Annual Reviews; 2011. p. 189–213. https://doi.org/10.1146/annurev-micro-090110-102946. vol. 65.
[16] Magnusson LU, Farewell A, Nystrom T. ppGpp: a global regulator in Escherichia coli. Trends Microbiol 2005;13:236–42. https://doi.org/10.1016/j.tim.2005.03.008.
[17] Kuroda A. A polyphosphate-Lon protease complex in the adaptation of Escherichia coli to amino acid starvation. Biosci Biotechnol Biochem 2006;70:325–31. https://doi.org/10.1271/bbb.70.325.

[18] Nguyen D, Joshi-Datar A, Lepine F, Bauerle E, Olakanmi O, Beer K, et al. Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. Science 2011;334:982–6. https://doi.org/10.1126/science.1211037.

[19] Ferguson GP, Totemeyer S, MacLean MJ, Booth IR. Methylglyoxal production in bacteria: suicide or survival? Arch Microbiol 1998;170:209–19. https://doi.org/10.1007/s002030050635.

[20] McCloskey D, Xu S, Sandberg TE, Brunk E, Hefner Y, Szubin R, et al. Adaptation to the coupling of glycolysis to toxic methylglyoxal production in tpiA deletion strains of Escherichia coli requires synchronized and counterintuitive genetic changes. Metab Eng 2018;48:82–93. https://doi.org/10.1016/j.ymben.2018.05.012.