RapidRIP quantifies the intracellular metabolome of 7 industrial strains of E. coli

McCloskey, Douglas; Xu, Julia; Schrübbers, Lars; Christensen, Hanne B.; Herrgård, Markus J.

Published in:
Metabolic Engineering

Link to article, DOI:
10.1016/j.ymben.2018.04.009

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Metabolic Engineering 47 (2018) 383–392

Contents lists available at ScienceDirect

Metabolic Engineering

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

RapidRIP quantifies the intracellular metabolome of 7 industrial strains of *E. coli*

Douglas McCloskey[a], Julia Xu[b], Lars Schrübbers[a], Hanne B. Christensen[a], Markus J. Herrgård[a,*]

[a] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark
[b] Department of Bioengineering, University of California - San Diego, La Jolla, CA 92093, USA

ABSTRACT

Fast metabolite quantification methods are required for high throughput screening of microbial strains obtained by combinatorial or evolutionary engineering approaches. In this study, a rapid RIP-LC-MS/MS (RapidRIP) method for high-throughput quantitative metabolomics was developed and validated that was capable of quantifying 102 metabolites from central, amino acid, energy, nucleotide, and cofactor metabolism in less than 5 minutes. The method was shown to have comparable sensitivity and resolving capability as compared to a full length RIP-LC-MS/MS method (FullRIP). The RapidRIP method was used to quantify the metabolome of seven industrial strains of *E. coli* revealing significant differences in glycolytic, pentose phosphate, TCA cycle, amino acid, and energy and cofactor metabolites were found. These differences translated to statistically and biologically significant differences in thermodynamics of biochemical reactions between strains that could have implications when choosing a host for bioprocessing.

1. Introduction

Traditional high-throughput screening approaches in metabolic engineering have primarily relied upon basic physiology markers of strain performance. Physiological markers often include basic input/output parameters such as growth rate, substrate uptake rate, product excretion rate, yield, and productivity. While useful, these markers give little information on the underlying strain physiology. Omics data types (e.g., transcriptomics, metabolomics, fluxomics, etc.,) could potentially provide a richer and deeper understanding of strain performance, which would allow more informed engineering decisions (McCloskey et al., 2013). However, -omics data types are often not utilized in routine screening of new production strain candidates because of high costs and low throughput (Hansen et al., 2017). Of the various -omics data types available, metabolomics provides the greatest potential to gain rich and deep insight on strain physiology at a lower cost and higher throughput (Fuhrer et al., 2011; Guder et al., 2017; Link et al., 2015).

Metabolomics methods used for the absolute quantification of intracellular metabolites are often on the order of 30 min (or 48 samples per day) (Bennette et al., 2011; Buescher et al., 2010; McCloskey et al., 2016a, 2015). Longer run-times are often required for complete chromatographic separation of biologically important isomers that can not be resolved by differences in MS fragmentation. Longer run-times are also needed to allow for enough instrument scan time to acquire enough points across each detected metabolite peak in order to accurately and reproducibly measure a high number of transitions in a given run (McCloskey et al., 2016a). With the advent of shorter columns with decreased particle sizes and mass spectrometers with faster scan rates, the potential to reduce method run-times to less than 5 min with minimal compromise to chromatographic resolution or quality of acquired peaks is now possible (Guder et al., 2017). However, current fast separation methods often compromise on the ability to resolve important biological isomers (e.g., glucose 6 phosphate and glucose 1 phosphate). In addition, fast separation methods often compromise on the separation of structurally similar compounds with interfering signals (e.g., AMP, ADP, and ATP) that may compromise accurate quantitation. The ability to separate these species is critical for quantitative modeling of biological systems (Almquist et al., 2014; Henry et al., 2006; Jamshidi and Palsson, 2008; Miskovic et al., 2017; Saa and Nielsen, 2016). An ultra high-throughput metabolomics method less than 5 min that is able to resolve of biologically important isomers and provide complete separation of structural similar compounds with interfering transitions has yet to be demonstrated.

Escherichia coli are often used as a model prokaryote for genomic and physiological studies (Archer et al., 2011; Arifin et al., 2014; Ishii et al., 2007; Vijayendran et al., 2007; Yoon et al., 2012), and as host strains for industrial bioprocesses (Chae et al., 2017; S. S.Y. Choi et al.,

* Corresponding author.
E-mail address: herrgard@biosustain.dtu.dk (M.J. Herrgård).

https://doi.org/10.1016/j.ymben.2018.04.009
Received 28 November 2017; Received in revised form 27 March 2018; Accepted 12 April 2018
Available online 25 April 2018

2016; S. Choi et al., 2016; Park et al., 2011, 2007; Yim et al., 2011) owing to the rich toolset available for gene and expression modification (Datsenko and Wanner, 2000; Jiang et al., 2015; Ronda et al., 2016; Wang et al., 2009) as well as accumulated historical knowledge base of biochemical understanding (Guo et al., 2013; Kanehisa et al., 2017; Keseler et al., 2013; Monk et al., 2017). Seven commonly used strains in biotechnology applications include K-12 MG1655, K-12 W3110, K-12 DH5a, BL21, C, Crooks, and W. It has been shown that particular strains are better suited for different applications (Monk et al., 2016). For example, DH5a is often preferred in cloning applications (Song et al., 2015; Taylor et al., 1993), while BL21 is often preferred for expression of recombinant proteins (Marisch et al., 2013a; Robichon et al., 2011). It has also been shown that choice of host strain for production of a given compound can significantly impact production titers (Na et al., 2013). Various comparisons between the seven industrial strains of *E. coli* under bioprocess conditions or as host strains for a particular metabolic engineering endeavor have been described (Chae et al., 2010; Fathi-Roudsari et al., 2016; Marisch et al., 2013a, 2013b; Monk et al., 2016). However, a comprehensive comparison of the metabolomes between the strains does not yet exist. An understanding of the starting metabolite levels of the strains would give researchers a key piece of information when deciding which *E. coli* strains to use as a host in their particular bioprocess.

In this study, a rapid reverse phase ion-pairing liquid chromatography triple quadrupole mass spectrometry (RapidRIP) for the rapid quantification of the intracellular metabolome for routine screening applications was developed and validated. It was shown that the method was capable of quantifying 102 metabolites in less than 5 min with comparable resolution, sensitivity, and reproducibility to a 33 min RIP-LC-MS/MS (FullRIP) method. The RapidRIP method was used to quantify the metabolome of seven industrial strains of *E. coli*. Significant differences in glycolytic, pentose phosphate, TCA cycle, amino acid, and energy and cofactor metabolites were found. These differences translated to strain-specific differences in reaction thermodynamics.

2. Material and methods

2.1. Biological material and growth conditions

Escherichia coli strains *E.coli* C (DSMZ 4860), *E. coli* Crooks (DSMZ 1576), *E. coli* DH5a (DSMZ 6897) *E. coli W* (DSMZ 1116), *E. coli W3110* (DSMZ 5911) were obtained from DSMZ-German Collection of Microorganism and Cell Cultures; *E. coli* BL21 (DE3) was purchased as competent cells from Agilent (Agilent Technologies), *E. coli* K-12 MG1655 (ATCC 700926). All cultures were grown in 25 mL of unlabeled or labeled glucose M9 minimal media (Sambrook and Russell, 2001) with trace elements (Fong et al., 2005) and sampled from a heat block in 50 mL autoclaved tubes that were maintained at 37 °C and aerated using magnetics.

Growth and sampling procedures for Pseudomonas, Mouse, and CHO are described in the Supplemental methods.

2.2. Materials and reagents

Uniformly labeled ^{13}C glucose was purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA). Unlabeled glucose and other media components were purchased from Sigma-Aldrich (St. Louis, MO). LC-MS reagents were purchased from Honeywell Burdick & Jackson® (Muskegon, MI), Fisher Scientific (Pittsburgh, PA) and Sigma-Aldrich (St. Louis, MO).

2.3. LC-MS/MS instrumentation and data processing

Metabolites were acquired and quantified on an AB SCIEX Qtrap® 5500 mass spectrometer (AB SCIEX, Framingham, MA) using a ACQUITY UPLC HSS T3 Column (100 Å, 1.8 µm, 2.1 mm X 30 mm) and processed using MultiQuant® 3.0.1 as described previously (McCloskey et al., 2015). The acquisition method used for all transitions is given in Table S10.

2.4. Metabolomics

Internal standards were generated as described previously (McCloskey et al., 2014a, 2014b). All samples and calibrators were spiked with the same amount of internal standard taken from the same batch of internal standards. Calibration curves were ran before and after all biological and analytical replicates. The consistency of quantification between calibration curves was checked by running a Quality Control sample that was composed of all biological replicates twice a day. Solvent blanks were injected every ninth sample to check for carryover. System suitability tests were injected daily to check instrument performance.

Metabolomics samples were acquired from triplicate cultures (1 mL of cell broth at an OD600 ~ 1.0) using a previously described method (Douglas McCloskey et al., 2014a, 2014b). A pooled sample of the filtered medium that was re-sampled using the FSF filtration technique and processed in the same way as the biological triplicates was used as an analytical blank. Extracts obtained from triplicate cultures and re-filtered medium were analyzed in duplicate. The intracellular values reported, unless otherwise noted, are derived from the average of the biological triplicates (n = 6). Metabolites in the pooled filtered medium with a concentration greater than 80% of that found in the triplicate samples were not analyzed due to high background intereference. In addition, metabolites that were found to have a quantifiable variability (RSD > = 50%) in the Quality Control samples or any individual components with an RSD > = 80 were not used for analysis.

Missing values were imputed using a bootstrapping approach as coded in the R package Amelia II (Honaker et al., 2011) (version 1.7.4, 1000 imputations). Remaining missing values were approximated as ½ the lower limit of quantification for the metabolite normalized to the biomass of the sample. Prior to statistical analyses, metabolite concentrations were log transformed to generate an approximately normal distribution using the R package LMGene (Rocke et al., n.d.) (version 3.3, “mult” = ”TRUE”, “lowessnorm” = ”FALSE”). A Bonferroni-adjusted p-value cutoff of 0.01 as calculated from a Student's *t*-test was used to determine significance between metabolite concentration levels. The glog-normalized values were used for downstream statistical analyses.

2.5. Thermodynamics

In vivo free energy change of reaction were calculated as described previously (McCloskey et al., 2014a, 2014b) using a recent genome-scale reconstruction of *E. coli* (Orth et al., 2011). In short, free energy of formation values taken from the eQuilibrator data base (Flamholz et al., 2012) were adjusted to physiologically estimated values for the cytosolic, periplasmic, and extracellular space for temperature (37 °C), pH (7.5, 7.0, and 7.0, respectively), and ionic strength (0.2 M). The 95% confidence intervals of the metabolomics concentrations were used to calculate the 95% confidence intervals of the free energies of reaction. Free energy changes of reaction with non-overlapping confidence intervals between the strains were defined as statistically significant. Free energy changes of reaction with confidence intervals with different signs between the strains were defined as biologically significant. Free energy changes of reaction were only compared for those reactions with sufficient measured metabolomics and compound free energy of formation data coverage (50% and 99% of all reactants and products, respectively). Concentration values for phosphate, water, hydrogen, oxygen, and carbon dioxide were estimated as 1e-3, 55.0, 0.034e-3, 0.055e-3, and 1.4e-3M, respectively. Concentration confidence intervals for missing metabolites were estimated as 1.58e-3M and 1.58 e-6M. Free energy of formation confidence intervals not measured were

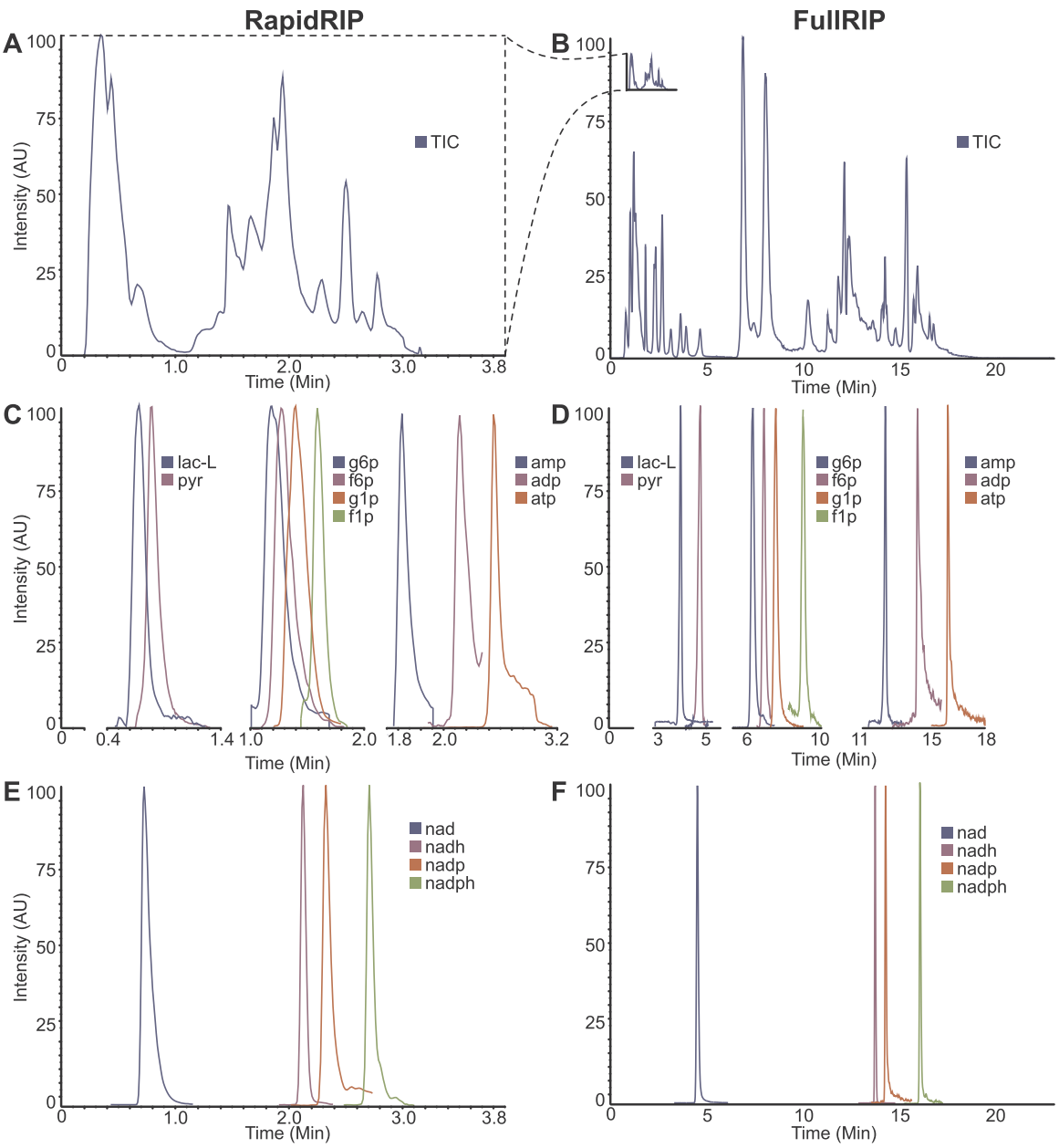

Fig. 1. RapidRIP vs. FullRIP chromatography. A comparison of chromatograms obtained from pools of standard mixes ran using the RapidRIP and FullRIP methods. The TIC (A and B), and XICs of lactate (lac-L), pyruvate (pyr), glucose 6-phosphate (g6p, fructose 6-phosphate (f6p), glucose 1-phosphate (g1p), fructose 1-phosphate (f1p), amp, adp, and atp (C and D), and nad(p)(h) (E and F) for the RapidRIP and FullRIP methods, respectively. Inset shows a scaled image of the TIC of the RapidRIP method projected onto the time axis of the FullRIP method.

estimated using the group contribution method (Flamholz et al., 2012).

Thermodynamic flux variability analysis (TFVA) was performed as previously described (Henry et al., 2007) using an open-source module for cobrapy called thermodynamics (https://github.com/dmccloskey/thermodynamics). Blocked reactions were those where the min and max flux was 0, essential reactions were those where the min and max flux were greater than 0, substitutable reactions were those where the min flux was 0 but the max flux was greater than 0, and constrained reactions were those where the min and max flux were the same.

3. Results and discussion

3.1. RapidRIP method development

A RapidRIP method was developed using the same mobile phase, temperature, and column chemistry to the FullRIP method (Fig. 1, Table 1). A shorter column with reduced particle size was used to maintain resolving power at a greatly reduced run-time (see Materials and Methods). A multitude of gradients, flow rates, and flow regimes were tested in order to optimize the method (Fig. S1). The optimized chromatographic parameters are given in the Material and Methods. Flow rate was modulated to allow for faster separation of later eluting compounds and faster column wash and equilibration, while maintaining the ability to retain early eluting compounds and separate hexose and pentose isomers. Percent mobile phase B was modulated to improve the separation of critical pairs, peak shape, and overall sensitivity (Fig. S1, Table S1). Interestingly, it was found that even within a greatly reduced method run time, substantial differences in separation between critical pairs could be obtained by careful tuning of the mobile phase gradient (Fig. S1, Table S1 and S2). Also, compound elution times

Table 1
RapidRIP chromatographic gradient. Chromatographic conditions are given in the Section 2.

Time	%B	flow_rate
0	0	0.5
0.4	0	0.5
0.9	2	0.5
1.0	6	0.5
1.4	6	0.5
1.5	11	0.5
1.9	11	0.5
2.1	28	0.5
2.2	53	0.5
2.5	53	0.5
2.6	0	0.5
3.9	0	0.5
4.4	0	0.5

were evenly distributed and peak widths were minimized in order to maximize the number of transitions that could be analyzed in a given run.

3.2. RapidRIP method validation

The ability of the method to resolve biologically critical isomers was determined. Critical pairs included the hexose and pentose phosphates, mono-, di-, tri- adenine nucleotide phosphates, nad(p)(h), and citrate/isocitrate, ac(coa), and various organic acids and amino acids (Fig. S1, Tables S1 and S2). Baseline resolution (resolution greater than or equal to 1.0) was achieved for 75% of the representative critical pairs listed (9 of 12 pairs), while separation (retention time difference > 0.01 min) was achieved for all pairs. Importantly, compounds that were not baseline resolved have non-isotopic overlapping transitions, which allow them to be measured without isotopic interference from neighboring compounds.

Quantification accuracy and variation was assessed by measuring the average points across the peak, and the variation in peak height across multiple injects of a neat standard solution for five different acquisition methods (Table S3). 55, 94, 61, 96, 100% of compounds measured had on average 10 or more acquired points across the baseline for MRM2 + EPI, MRM2, MRM1 + EPI, MRM1, and MRMsub acquisition methods, respectively. 81, 94, 72, 98, 100% of compounds measured had a RSD of peak height of less than 30% for MRM2 + EPI, MRM2, MRM1 + EPI, MRM1, and MRMsub, respectively. These results indicated that the MRM2 and MRM1 acquisition methods were able to collect a sufficient number of points across the peak to allow for accurate quantification of almost all target transitions. In contrast, the drop in the number of points across the peak as well as the increased variation in peak height found for the MRM2 + EPI and MRM1 + EPI indicated that the acquisition methods that included an additional data-dependent enhanced product ion (EPI) scan were not suitable for accurate quantification. While the ability to collect an additional product ion spectra for further compound identification confirmation would be desirable, a tradeoff of decreased number of transitions would be required. Consequently, the MRM only methods were used in further method validation.

Method linearity and sensitivity was determined by running calibration curves for all compounds (Table S4). A total of 102 compounds were quantifiable. The limits of quantification ranged from less than 1 nanomolar (aromatic and phosphorylated metabolites) to above 100 micromolar (amino acids), with many metabolites spanning a detection range of 5 orders of magnitude.

The reproducibility of the method was determined by measuring pooled samples of representative sample matrices that included (Fig. 2, Table S5). 98, 100, 53, and 97 compounds were measured in Pseudomonas, *E. coli*, Mouse Plasma, and CHO, respectively. 90, 90, 53, and 93 compounds were found to be quantifiable in Pseudomonas, *E. coli*, Mouse Plasma, and CHO, respectively. 85%, 81%, 91%, and 88% of measured compounds had a peak height ratio RSD of less than 30% in Pseudomonas, *E. coli*, Mouse Plasma, and CHO, respectively. Metabolites with a RSD greater than 30% were generally found to be at the lower limits of detection(LLOD). All measured components had a retention time RSD of less than 10% in all sample matrices tested. Importantly, no carryover was found between any of the runs (data not shown).

3.3. RapidRIP vs. FullRIP

The RapidRIP method was compared to a previously published 33 min RIP-LC-MS/MS method. It was found that sufficient resolution between critical pairs could be maintained for a majority of the compounds analyzed between the RapidRIP and FullRIP methods (Fig. 1, Tables S2) to allow for comparable quantitation (Fig. 3). For example, the separation between the pentose isomers ribose 5-phosphate dn ribulose 5-phosphate was sufficient to achieve baseline separation. A notable compromise in resolution in the reduced method was the separation of glucose 6-phosphate and fructose 6-phosphate, which were not well separated compared to the FullRIP (Fig. 1.) However, the unique 199 product ion allows for the direct quantification of glucose 6-phosphate and indirect quantification of fructose 6-phosphate by subtraction of the calculated concentrations as determined from a calibration curve from the 169 and 199 ions.

It was found that both methods were comparable in linearity and sensitivity (Table S4). 54% of the compounds measured with RapidRIP had an LLOQ less than or equal to those measured with FullRIP; and 75% of the compounds measured with RapidRIP had an ULOQ greater than or equal to those measured with FullRIP. A decrease in sensitivity and LLOQ but increase in ULOQ was found primarily for early eluting compounds consisting of amino acids, nucleosides and nucleotides. An increase in resolution and LLOQ was found for many later eluting metabolites. The increase in resolution can be explained by the decreased peak width using the RapidRIP method. The overall increase in ULOQ may be attributed to an overall increase in ion suppression due to the number of components eluting from the column in a given period of time. This is most noticeable for the early eluting compounds noted above where a loss in sensitivity was also found.

The RapidRIP and FullRIP methods were found to have comparable quantitative accuracy (Fig. 3A-G). A correlation coefficient (Pearson's R) greater than or equal to 0.88 between glog normalized absolute metabolite concentrations (mmol*gDW-1) for all seven strains tested (Fig. 3A-G, Table S7) was found. Interestingly, the peak height ratio correlation coefficients between strains were much less than the absolute metabolite correlation coefficients (Table S7). This indicates that there was sufficient changes in analyte or internal standard peak heights from other components in the sample matrix that were well separated in the FullRIP method that are no longer well separated in the RapidRIP method to make the direct comparison between peak height ratios problematic. However, this also indicates that the use of a calibration curve was sufficient to compensate for the majority of these changes.

The RapidRIP and FullRIP methods were also found to have comparable quantitative precision (Fig. 3H). 61% and 58% of quantified metabolites for all strains had a % RSD less than or equal to 30 for RapidRIP and FullRIP methods, respectively. 7% and 4% of quantified metabolites for all strains had a % RSD greater than 60% RSD for RapidRIP and FullRIP, respectively. The distribution of %RSDs (Fig. 3H), and the percentages of metabolites below 30 and above 60%RSD between both methods indicates that the short method can retain the quantitative precision of the full method.

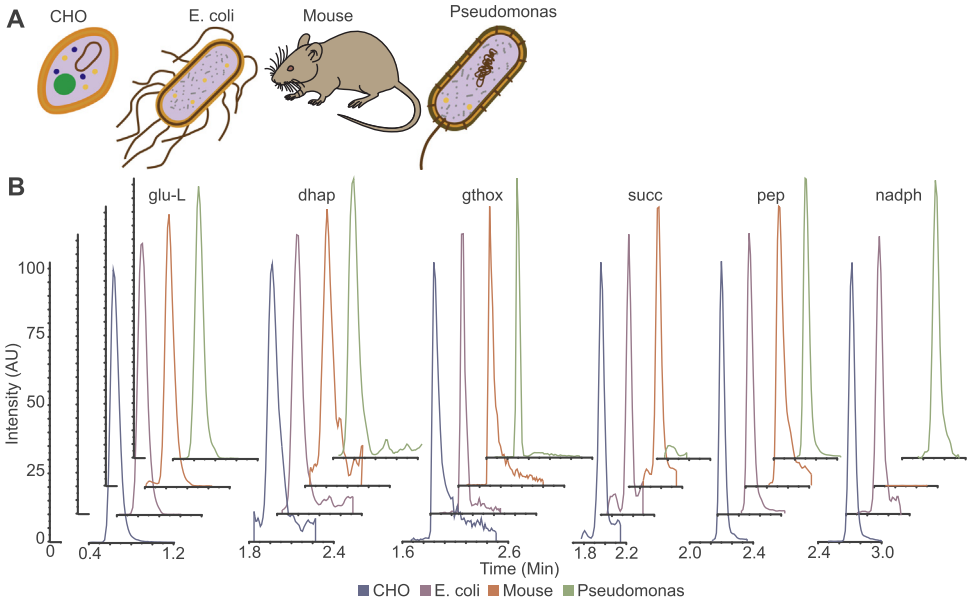

Fig. 2. Method robustness and reproducibility in complex biological matrices. Pooled samples of *Pseudomonas* sp. VLB120 (*Pseudomonas*), *E. coli* MG1655 K-12 (*E. coli*), mouse plasma (Mouse), and chinese hamster ovarian cells (CHO) were analyzed across multiple injections at periodic intervals. Representative compounds across the chromatogram for all sample matrices are shown. Representative compounds shown include L-glutamate (glu-L), dihydroxyacetone phosphate (dhap), glutathione oxidized (gthox), succinate (succ), phosphoenol pyruvate (pep), and reduced Nicotinamide adenine dinucleotide phosphate (nadph). All peaks were scaled to their maximum ion count for comparison. Peaks for succ in Pseudomonas and nadph in Mouse were not found due to an incomplete TCA cycle in Pseudomonas and lack of intracellular metabolites present in mouse plasma, respectively.

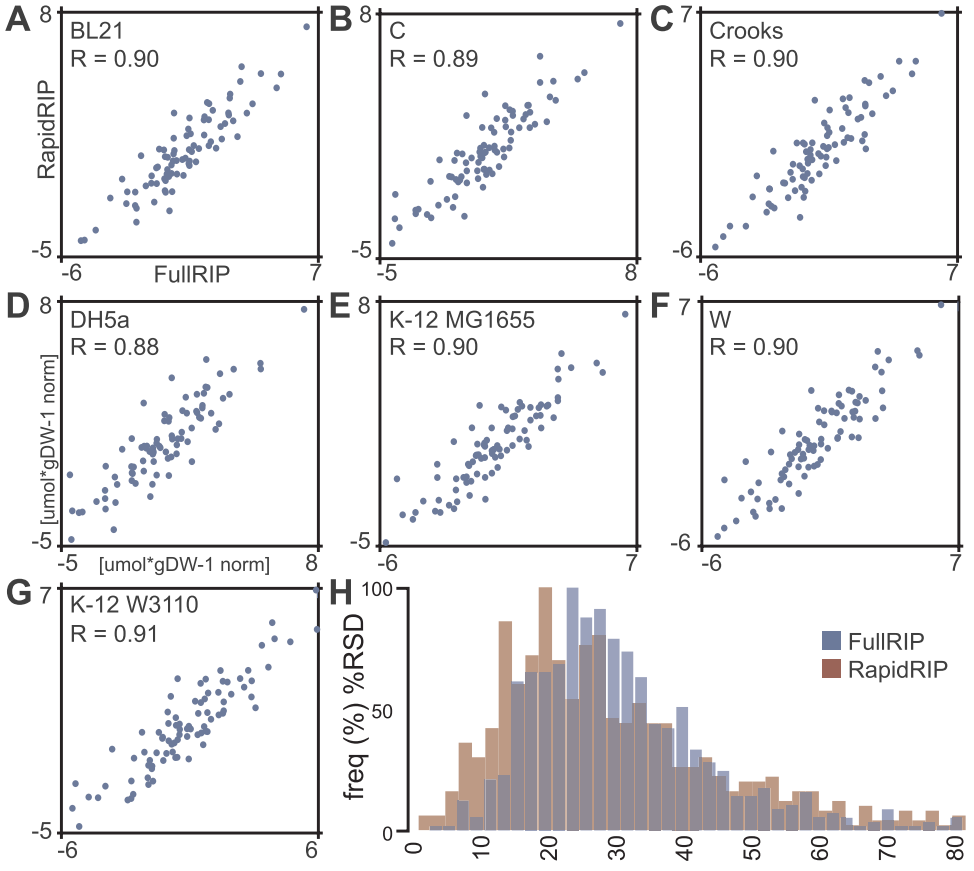

Fig. 3. RapidRIP vs. FullRIP quantitative accuracy and precision. A comparison of absolute quantitation of metabolites using the RapidRIP (y axis) and FullRIP (x axis) are shown for the 7 industrial strains of *E. coli*: A) *E. coli* BL21, B) *E. coli* C, C) *E. coli* Crooks, D) *E. coli* DH5a, E) *E. coli* K-12 *E. coli* MG1655, F) *E. coli* W, and G) *E. coli* K-12 W3110. The name of the strain and the Pearson correlation coefficient are annotated in the plots. The x and y axis are glog normalized absolute metabolite concentrations (mmol*gDW-1). For comparison, the glog normalized peak height ratios between analytes and internal standards are given in Table S7. H) Histogram of %RSD of RapidRIP and FullRIP.

3.4. Quantification of the intracellular metabolome of seven industrial strains of E. coli

The RapidRIP method was used to quantify the intracellular metabolome of 7 industrial strains of *E. coli* (Fig. 4, Table S6, see Material and Methods for strains). Consistent with previous literature, the most abundant metabolites in the strains consisted of reduced glutathione, L-glutamate, and ATP (Bennett et al., 2009; Taymaz-Nikerel et al., 2011). Hierarchical clustering (Fig. 4A) reveals that similar to gene expression levels, the levels of intracellular metabolites do not reflective genomic distances between the strains (Monk et al., 2016). The hierarchical clustering is also consistent with how the strains group by Partial Least Squares Differential Analysis (PLS-DA, Fig. 4C). The primary mode of separation between the strains involve differences in glycolytic, pentose phosphate, and TCA cycle intermediates, and the secondary mode of separation include metabolites involved in nucleotide metabolism (Fig. 4D). These differences are discussed in greater detail below.

Physiological ratios are ratios of individual metabolites that often reflect broad physiological states (e.g., energy depletion, redox imbalance, etc.,) that are tightly regulated by the cell. Physiological ratios,

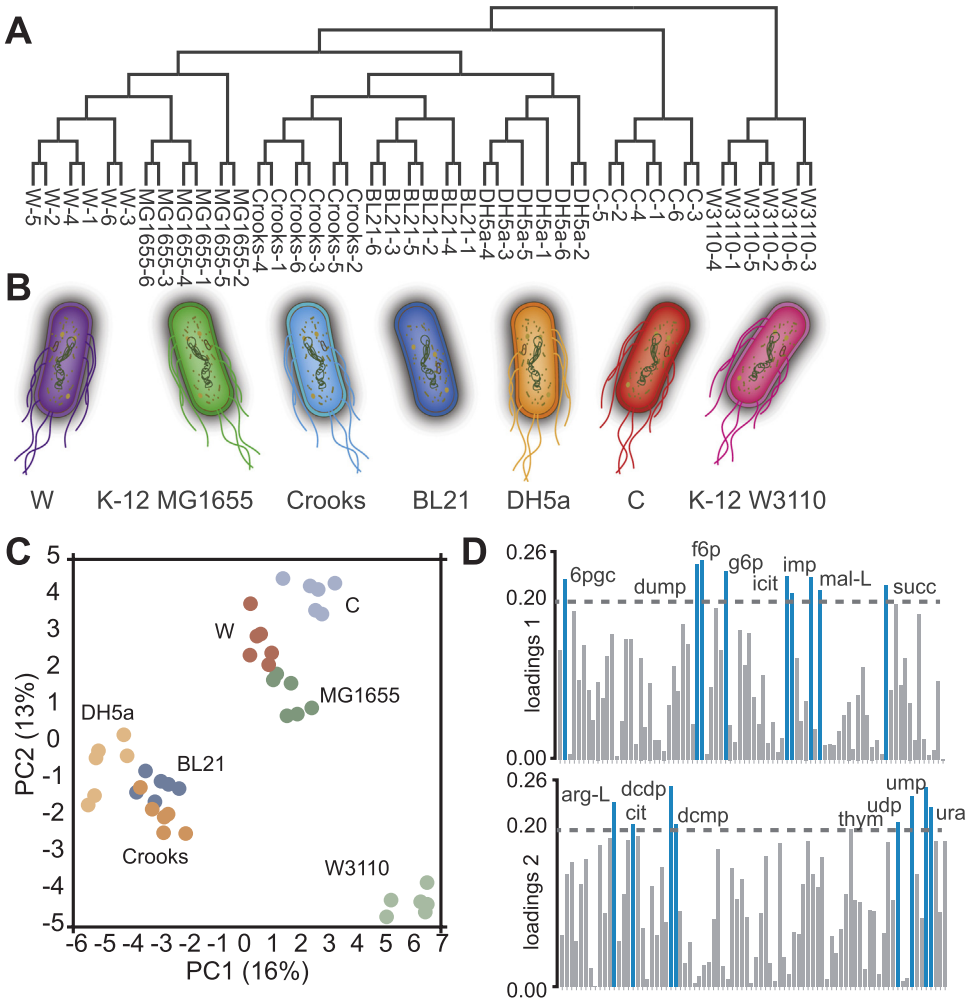

Fig. 4. Statistical analysis of the absolute metabolite concentrations for the 7 industrial strains of *E. coli*. A) A cluster analysis (Euclidean, Complete) of the 7 strains. B) The 7 strains are shown. C) PLS-DA scores plot of components 1 and 2. D) PLS-DA loadings plot of components 1 and 2. Metabolites with an absolute loadings greater than 0.2 are annotated and highlighted in light blue.

including the nitrogen charge, energy charge, glutathione ratio, and redox ratio, differed between the strains (see methods for ratio definitions, Fig. 5A). *E. coli* C and DH5a had the lowest nitrogen and energy charges due to elevated levels of akg and amp, respectively. Non-significant differences in the glutathione and redox ratio, but significant differences in the metabolites that compose those ratios were found. The conservation of these ratios, but not necessarily the levels of the individual components may indicate the importance of the glutathione and redox ratios for maintaining normal aerobic physiology.

The intracellular concentration of 14 amino acids differed between the strains (Fig. 5B). *E, coli* DH5a was found to be a poor starting strain for L-tryptophan production. The intracellular concentrations of L-tryptophan were significantly lower than the other strains. *E, coli* MG1655, C, and W3110 were found to maintain significantly higher levels of D-aspartate. This is most likely due to utilization of the Phosphoenolpyruvate Carboxylase (PPC) in MG1655 and W3110 (McCloskey et al., 2016a, 2016b), and overall increased levels of TCA cycle intermediates in strain C. *E. coli* MG1655 and W3110 were also found to have the highest levels of L-arginine. It should be noted that when compared to the transcript levels or flux predictions of the amino acid producing pathways (Monk et al., 2016), little correlation between the absolute metabolite concentrations was found, which indicates the difficulty in predicting metabolite levels through indirect evidence as is consistent with previous works (Daran-Lapujade et al., 2007; Hackett et al., 2016).

Levels of central metabolism intermediates differed vastly between the strains (Fig. 6). *E. coli* MG1655, W, and W3110 were found to maintain significantly lower levels of glycolytic intermediates than other strains (i.e., g6p, f6p, fdp, 13dpg/23dpg, 2 pg/3 pg, and pep). However, the levels of pep in W3110 were found to be significantly higher than all other strains. *E. coli* MG1655 had the highest levels of ribulose and ribose 5- phosphate (ru5p-D and r5p) compared to the other strains. This most likely contributed to the higher levels of L-histidine, a downstream product of r5p, in *E. coli* MG1655.

Based on the data collected here, *E. coli* C would serve as a good candidate for aerobic succinate and malate production due to high endogenous levels of succinate. *E. coli* C had significantly higher DL-lactate, succinate, malate, and fumarate concentrations compared to the other six strains. In addition, the levels of TCA cycle intermediates citrate and 2-oxoglutarate were among the highest. *E. coli* C has indeed previously been successfully used as the base strain for succinate production from glucose under aerobic and anaerobic conditions (Balzer et al., 2013; Jantama et al., 2008; Lin et al., 2005a, 2005b; Sánchez et al., 2005). Interestingly, *E. coli* DH5a had the highest levels of citrate, aconitate, isocitrate, and 2-oxoglutarate. This could indicate altered regulation or utilization of the TCA cycle in *E. coli* DH5a.

3.5. Thermodynamic analysis of the seven industrial strains

Change in free energies of reaction (dGr) for each of the industrial

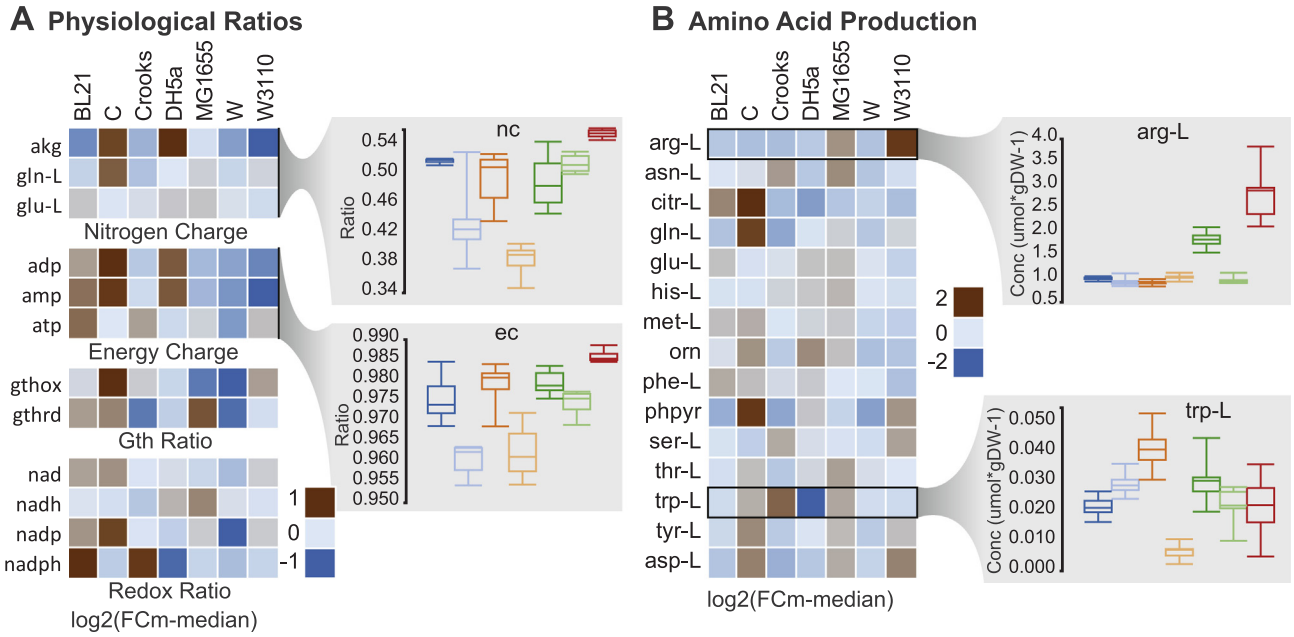

Fig. 5. Physiological ratios (A) and intracellular amino acid levels (B). Selected ratios and metabolite levels that significantly changed are highlighted in the inset. The nitrogen charge (nc) is defined as (gln-L + 0.5 *glu-L) * (gln-L + glu-L + akg)-1. The energy charge is defined as (atp + .5 * adp)* (atp + adp + amp)-1. The glutathione ratio (Gth) is defined as (gthrd*gthox-1). The redox ratio is defined as (nad(p)h * (nad(p) + nad(p)h)-1. Units are the log2 fold change of the median value of each strain to the mean of all 7 strains.

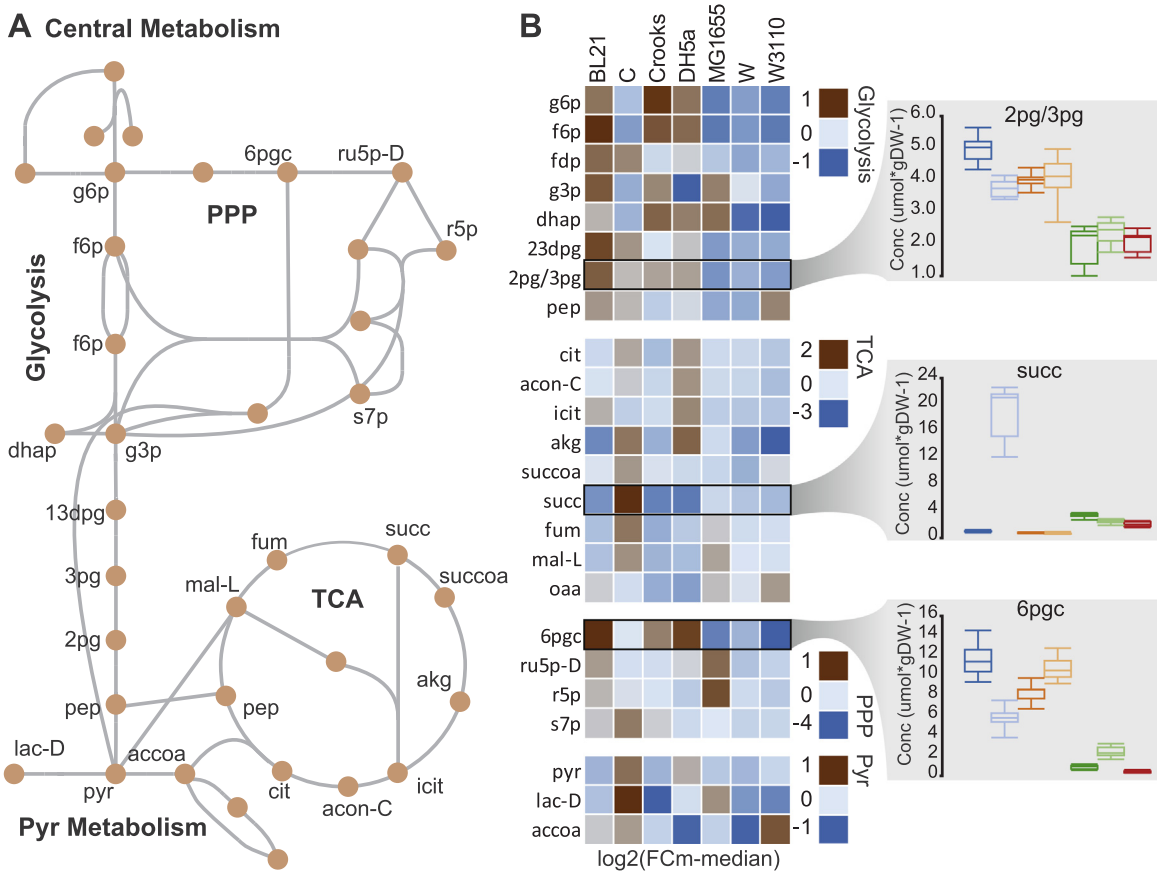

Fig. 6. Absolute metabolite concentrations of central metabolism intermediates for the 7 industrial strains of *E. coli*. A) Schematic of central carbohydrate metabolism. B) Heatmap of metabolite levels. Units are the log2 fold change of the median value of each strain to the mean of all 7 strains. Selected metabolites that changed significantly are shown in the inset.

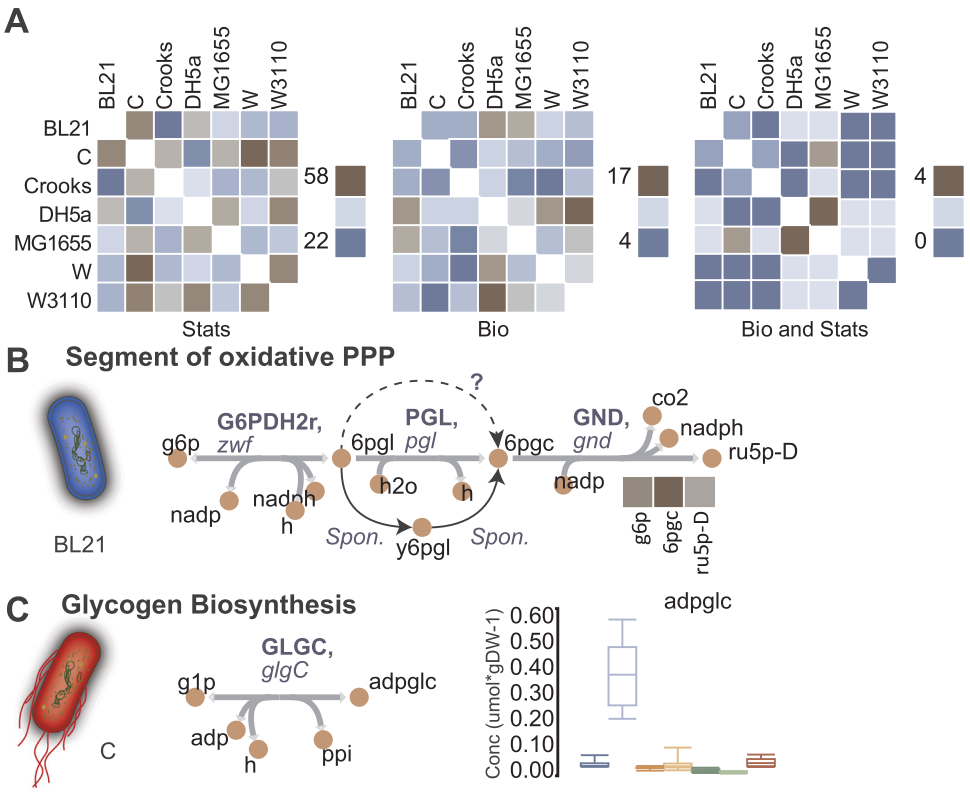

Fig. 7. Thermodynamic analysis of the 7 industrial strains of *E. coli*. A) Heatmap of pairwise counts of reaction that differed either statistically, biologically, or both between the strains. B) Schematic of the oxidative pentose phosphate pathway (oxPPP) with spontaneous and unknown (?) reactions shown for *E. coli* BL21. Note the high levels of oxPPP intermediates in *E. coli* BL21. C) Schematic of the first step in glycogen biosynthesis encoded by glucose 1-phosphate adenyltransferase (GLGC). Note the high level of ADP-glucose (adpglc) in *E. coli* C. The change in reaction free energy differed for GLGC differed biologically and statistically between *E. coli* C and all other strains except *E. coli* W and W3110.

strains were calculated and compared in order to identify statistically (i.e., non-overlapping confidence intervals) and biologically (i.e., a change in dGr sign indicating a change in reaction directionality) significant differences in reaction thermodynamics between the strains (Table S8). Of the 475 reactions with satisfactory measured metabolomics and free energy of formation coverage (see Methods), between 22 and 58 statistically different reactions were found between the strains, between 4 and 17 biologically different reactions were found between the strains, and between 0 and 4 statistically and biologically different reactions were found between the strains (Fig. 7, Table S9).

Thermodynamics analysis revealed differences in the thermodynamic potential to push carbon towards glycogen biosynthesis instead of more economically valuable carbon endpoints between the strains. Biologically and statistically significant differences were found between glucose 1-phosphate adenyltransfarase (GLGLC, EC 2.7.7.27) in *E. coli* C and all other strains except *E. coli* W and W3110 (Fig. 7). *E. coli* C had the highest levels of the glycogen precursor ADP-glucose, which leads to the thermodynamic differences. ADP-glucose (adpglc) is formed from glucose 1-phosphate via GLGC. Increased expression of *glgC* or mutations in *glgC* that increase GLGC activity have been shown to accumulate glycogen (Ballicora et al., 2003; Eydallin et al., 2007; Ghosh et al., 1992; Leung et al., 1986). No significant differences in the transcriptomic profiles of *glgC* in *E. coli* C were found and *E. coli* C is not known to harbor any unique mutation in the *glgC* gene, indicating that the elevated levels of adpglc could be due to kinetic factors.

Thermodynamics analysis also revealed that E. coli BL21 can be used for biotechnology without concern of a reduced flux through the oxidative pentose phosphate pathway (oxPPP). No biologically or statistically significant differences in change of reaction free energy were found between PGL in *E. coli* BL21 and the other seven strains (Fig. 7). *E. coli* BL21 had the highest levels of 6-phosphogluconate (6pgc). BL21 lacks the *pgl* gene that encodes the hydrolase6-phosphogluconolactonase (PGL) that converts 6-phosphoglucono-δ-lactone to (6pgl) to 6-phosphogluconate (6pgc) (Meier et al., 2012; Studier et al., 2009). In the absence of PGL, 6pgc is spontaneous converted to g-6-phosphogluconolactone (y6pgl), and then spontaneously broken down to 6pgc. However, BL21 appears to also have a PGL bypass (Meier et al., 2012) that utilizes an uncharacterized pathway to rapidly convert 6pgl to 6pgc. In either case, the results shown here are consistent with previous findings that 6pgl and 6pgc rapidly accumulate to high intracellular levels (Meier et al., 2012). Also consistent are the relative levels of the downstream metabolite ribulose 5-phosphate (ru5p-D) compared to MG1655 (Meier et al., 2012). Non-intuitively, the lack of PGL in the oxPPP does not appear to detrimentally affect the levels of Reduced Glutathione nor NADPH compared to the other strains (Fig. 5A). This indicates that utilization of the oxPPP through spontaneous conversion with or without utilizing an uncharacterized bypass is feasible thermodynamically.

4. Conclusion

In this study, a RapidRIP method capable of quantifying over 100 metabolites in less than 5 min was described. The method is capable of analyzing 327 samples per day (or 2289 samples per week). Including QC, calibrator, and carryover check samples, this amounts to approximately 1000 strains that can be screened per week on a single instrument. Metrics for sensitivity, resolution, linearity, accuracy, and precision was determined and compared to a FullRIP method. All metrics were found to be comparable with a few minor compromises as noted in the main text. The RapidRIP method significantly accelerates metabolomic characterization compared to existing methods. Combined with high-throughput cultivation methods this allows using metabolomics as a routine tool for characterizing the large-numbers of engineered or evolved microbial strains that modern cloning, genome editing and laboratory evolution approaches create.

In order to demonstrate the applicability of RapidRIP for characterizing strain differences, od was used to quantify the metabolome of 7 industrial strains of *E. coli* during aerobic growth on glucose. Major differences in central, amino acid, nucleotide, energy, and redox metabolism metabolite levels were found. These differences translated to statistically significant differences in reaction thermodynamics; several of which were also found to be biologically significant through change

in reaction directionality. Knowledge of these differences provides researchers with valuable information when choosing which strain to use as host for a particular chemical production process. For example, researchers should not be dissuaded from using the BL21 strain based on theoretical predictions of compound yield derived from genomic information alone because the thermodynamics of the PPP in the absence of the *pgl* gene are such that flux and levels of NADPH are not significantly altered.

Acknowledgements

We would like to thank the Novo Nordisk Foundation for funding this work through the Novo Nordisk Foundation Center for Biosustainability. This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 686070.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2018.04.009.

References

Almquist, J., Cvijovic, M., Hatzimanikatis, V., Nielsen, J., Jirstrand, M., 2014. Kinetic models in industrial biotechnology - improving cell factory performance. Metab. Eng. 24, 38–60.

Archer, C.T., Kim, J.F., Jeong, H., Park, J.H., Vickers, C.E., Lee, S.Y., Nielsen, L.K., 2011. The genome sequence of E. coli W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of E. coli. BMC Genom. 12, 9.

Arifin, Y., Archer, C., Lim, S., Quek, L.-E., Sugiarto, H., Marcellin, E., Vickers, C.E., Krömer, J.O., Nielsen, L.K., 2014. Escherichia coli W shows fast, highly oxidative sucrose metabolism and low acetate formation. Appl. Microbiol. Biotechnol. 98, 9033–9044.

Ballicora, M.A., Iglesias, A.A., Preiss, J., 2003. ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. Microbiol. Mol. Biol. Rev. 67, 213–225 (table of contents).

Balzer, G.J., Thakker, C., Bennett, G.N., San, K.-Y., 2013. Metabolic engineering of Escherichia coli to minimize byproduct formate and improving succinate productivity through increasing NADH availability by heterologous expression of NAD (+)-dependent formate dehydrogenase. Metab. Eng. 20, 1–8.

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

Bennette, N.B., Eng, J.F., Dismukes, G.C., 2011. An LC-MS-based chemical and analytical method for targeted metabolite quantification in the model cyanobacterium Synechococcus sp. PCC 7002. Anal. Chem. 83, 3808–3816.

Buescher, J.M., Moco, S., Sauer, U., Zamboni, N., 2010. Ultrahigh performance liquid chromatography- tandem mass spectrometry method for fast and robust quantification of anionic and aromatic metabolites. Anal. Chem. 82, 4403–4412.

Chae, H.S., Kim, K.-H., Kim, S.C., Lee, P.C., 2010. Strain-dependent carotenoid productions in metabolically engineered Escherichia coli. Appl. Biochem. Biotechnol. 162, 2333–2344.

Chae, T.U., Ko, Y.-S., Hwang, K.-S., Lee, S.Y., 2017. Metabolic engineering of Escherichia coli for the production of four-, five- and six-carbon lactams. Metab. Eng. 41, 82–91.

Choi, S., Kim, H.U., Kim, T.Y., Lee, S.Y., 2016. Systematic engineering of TCA cycle for optimal production of a four-carbon platform chemical 4-hydroxybutyric acid in Escherichia coli. Metab. Eng. 38, 264–273.

Choi, S.Y., Park, S.J., Kim, W.J., Yang, J.E., Lee, H., Shin, J., Lee, S.Y., 2016. One-step fermentative production of poly(lactate-co-glycolate) from carbohydrates in Escherichia coli. Nat. Biotechnol. 34, 435–440.

Daran-Lapujade, P., Rossell, S., van Gulik, W.M., Luttik, M.A.H., de Groot, M.J.L., Slijper, M., Heck, A.J.R., Daran, J.-M., de Winde, J.H., Westerhoff, H.V., Pronk, J.T., Bakker, B.M., 2007. The fluxes through glycolytic enzymes in Saccharomyces cerevisiae are predominantly regulated at posttranscriptional levels. Proc. Natl. Acad. Sci. USA 104, 15753–15758.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. USA 97, 6640–6645.

Eydallin, G., Morán-Zorzano, M.T., Muñoz, F.J., Baroja-Fernández, E., Montero, M., Alonso-Casajús, N., Viale, A.M., Pozueta-Romero, J., 2007. An Escherichia coli mutant producing a truncated inactive form of GlgC synthesizes glycogen: further evidences for the occurrence of various important sources of ADPglucose in enterobacteria. FEBS Lett. 581, 4417–4422.

Fathi-Roudsari, M., Akhavian-Tehrani, A., Maghsoudi, N., 2016. Comparison of Three Escherichia coli Strains in Recombinant Production of Reteplase. Avicenna J. Med. Biotechnol. 8, 16–22.

Flamholz, A., Noor, E., Bar-Even, A., Milo, R., 2012. eQuilibrator–the biochemical thermodynamics calculator. Nucleic Acids Res. 40, D770–D775.

Fong, S.S., Burgard, A.P., Herring, C.D., Knight, E.M., Blattner, F.R., Maranas, C.D., Palsson, B.O., 2005. In silico design and adaptive evolution of Escherichia coli for production of lactic acid. Biotechnol. Bioeng. 91, 643–648.

Fuhrer, T., Heer, D., Begemann, B., Zamboni, N., 2011. High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. Anal. Chem. 83, 7074–7080.

Ghosh, P., Meyer, C., Remy, E., Peterson, D., Preiss, J., 1992. Cloning, expression, and nucleotide sequence of glgC gene from an allosteric mutant of Escherichia coli B. Arch. Biochem. Biophys. 296, 122–128.

Guder, J.C., Schramm, T., Sander, T., Link, H., 2017. Time-optimized isotope ratio LC-MS/MS for high-throughput quantification of primary metabolites. Anal. Chem. 89, 1624–1631.

Guo, A.C., Jewison, T., Wilson, M., Liu, Y., Knox, C., Djoumbou, Y., Lo, P., Mandal, R., Krishnamurthy, R., Wishart, D.S., 2013. ECMDB: the E. coli metabolome database. Nucleic Acids Res. 41, D625–D630.

Hackett, S.R., Zanotelli, V.R.T., Xu, W., Goya, J., Park, J.O., Perlman, D.H., Gibney, P.A., Botstein, D., Storey, J.D., Rabinowitz, J.D., 2016. Systems-level analysis of mechanisms regulating yeast metabolic flux. Science 354. http://dx.doi.org/10.1126/science.aaf2786.

Hansen, A.S.L., Lennen, R.M., Sonnenschein, N., Herrgård, M.J., 2017. Systems biology solutions for biochemical production challenges. Curr. Opin. Biotechnol. 45, 85–91.

Henry, C.S., Broadbelt, L.J., Hatzimanikatis, V., 2007. Thermodynamics-based metabolic flux analysis. Biophys. J. 92, 1792–1805.

Henry, C.S., Jankowski, M.D., Broadbelt, L.J., Hatzimanikatis, V., 2006. Genome-scale thermodynamic analysis of Escherichia coli metabolism. Biophys. J. 90, 1453–1461.

Honaker, J., King, G., Blackwell, M., 2011. Amelia II: a program for missing data. J. Stat. Softw. 45, 1–47.

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K., Hoque, A., Ho, P.Y., Kakazu, Y., Sugawara, K., Igarashi, S., Harada, S., Masuda, T., Sugiyama, N., Togashi, T., Hasegawa, M., Takai, Y., Yugi, K., Arakawa, K., Iwata, N., Toya, Y., Nakayama, Y., Nishioka, T., Shimizu, K., Mori, H., Tomita, M., 2007. Multiple high-throughput analyses monitor the response of E. coli to perturbations. Science 316, 593–597.

Jamshidi, N., Palsson, B.Ø., 2008. Formulating genome-scale kinetic models in the post-genome era. Mol. Syst. Biol. 4, 171.

Jantama, K., Haupt, M.J., Svoronos, S.A., Zhang, X., Moore, J.C., Shanmugam, K.T., Ingram, L.O., 2008. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate. Biotechnol. Bioeng. 99, 1140–1153.

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., Yang, S., 2015. Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. Appl. Environ. Microbiol. 81, 2506–2514.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., Morishima, K., 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45, D353–D361.

Keseler, I.M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., Fulcher, C., Huerta, A.M., Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Schröder, I., Shearer, A.G., Subhraveti, P., Travers, M., Weerasinghe, D., Weiss, V., Collado-Vides, J., Gunsalus, R.P., Paulsen, I., Karp, P.D., 2013. EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Res. 41, D605–D612.

Leung, P., Lee, Y.M., Greenberg, E., Esch, K., Boylan, S., Preiss, J., 1986. Cloning and expression of the Escherichia coli glgC gene from a mutant containing an ADPglucose pyrophosphorylase with altered allosteric properties. J. Bacteriol. 167, 82–88.

Lin, H., Bennett, G.N., San, K.-Y., 2005a. Chemostat culture characterization of Escherichia coli mutant strains metabolically engineered for aerobic succinate production: a study of the modified metabolic network based on metabolite profile, enzyme activity, and gene expression profile. Metab. Eng. 7, 337–352.

Lin, H., Bennett, G.N., San, K.-Y., 2005b. Metabolic engineering of aerobic succinate production systems in Escherichia coli to improve process productivity and achieve the maximum theoretical succinate yield. Metab. Eng. 7, 116–127.

Link, H., Fuhrer, T., Gerosa, L., Zamboni, N., Sauer, U., 2015. Real-time metabolome profiling of the metabolic switch between starvation and growth. Nat. Methods 12, 1091–1097.

Marisch, K., Bayer, K., Cserjan-Puschmann, M., Luchner, M., Striedner, G., 2013a. Evaluation of three industrial Escherichia coli strains in fed-batch cultivations during high-level SOD protein production. Microb. Cell Fact. 12, 58.

Marisch, K., Bayer, K., Scharl, T., Mairhofer, J., Krempl, P.M., Hummel, K., Razzazi-Fazeli, E., Striedner, G., 2013b. A comparative analysis of industrial Escherichia coli K-12 and B strains in high-glucose batch cultivations on process-, transcriptome- and proteome level. PLoS One 8, e70516.

McCloskey, D., Gangoiti, J.A., King, Z.A., Naviaux, R.K., Barshop, B.A., Palsson, B.O., Feist, A.M., 2014a. A model-driven quantitative metabolomics analysis of aerobic and anaerobic metabolism in E. coli K-12 MG1655 that is biochemically and thermodynamically consistent. Biotechnol. Bioeng. 111, 803–815.

McCloskey, D., Gangoiti, J.A., Palsson, B.O., Feist, A.M., 2015. A pH and solvent optimized reverse-phase ion-paring-LC–MS/MS method that leverages multiple scan-types for targeted absolute quantification of intracellular metabolites. Metabolomics 11, 1338–1350.

McCloskey, D., Palsson, B.Ø., Feist, A.M., 2013. Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli. Mol. Syst. Biol. 9, 661.

McCloskey, D., Utrilla, J., Naviaux, R.K., Palsson, B.O., Feist, A.M., 2014b. Fast Swinnex filtration (FSF): a fast and robust sampling and extraction method suitable for metabolomics analysis of cultures grown in complex media. Metabolomics 11, 198–209.

McCloskey, D., Young, J.D., Xu, S., Palsson, B.O., Feist, A.M., 2016a. MID max: LC-MS/MS method for measuring the precursor and product mass isotopomer distributions of

metabolic intermediates and cofactors for metabolic flux analysis applications. Anal. Chem. 88, 1362–1370.
McCloskey, D., Young, J.D., Xu, S., Palsson, B.O., Feist, A.M., 2016b. Modeling method for increased precision and scope of directly measurable fluxes at a genome-scale. Anal. Chem. 88, 3844–3852.
Meier, S., Jensen, P.R., Duus, J.Ø., 2012. Direct observation of metabolic differences in living Escherichia coli strains K-12 and BL21. Chembiochem 13, 308–310.
Miskovic, L., Alff-Tuomala, S., Soh, K.C., Barth, D., Salusjärvi, L., Pitkänen, J.-P., Ruohonen, L., Penttilä, M., Hatzimanikatis, V., 2017. A design-build-test cycle using modeling and experiments reveals interdependencies between upper glycolysis and xylose uptake in recombinantS. cerevisiaeand improves predictive capabilities of large-scale kinetic models. Biotechnol. Biofuels 10, 166.
Monk, J.M., Koza, A., Campodonico, M.A., Machado, D., Seoane, J.M., Palsson, B.O., Herrgård, M.J., Feist, A.M., 2016. Multi-omics quantification of species variation of escherichia coli links molecular features with strain phenotypes. Cell Syst. 3, 238–251 (e12).
Monk, J.M., Lloyd, C.J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A.M., Palsson, B.O., 2017. iML1515, a knowledgebase that computes Escherichia coli traits. Nat. Biotechnol. 35, 904–908.
Na, D., Yoo, S.M., Chung, H., Park, H., Park, J.H., Lee, S.Y., 2013. Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. Nat. Biotechnol. 31, 170–174.
Orth, J.D., Conrad, T.M., Na, J., Lerman, J.A., Nam, H., Feist, A.M., Palsson, B.Ø., 2011. A comprehensive genome-scale reconstruction of Escherichia coli metabolism-—2011. Mol. Syst. Biol. 7, 535.
Park, J.H., Jang, Y.-S., Lee, J.W., Lee, S.Y., 2011. Escherichia coli W as a new platform strain for the enhanced production of L-valine by systems metabolic engineering. Biotechnol. Bioeng. 108, 1140–1147.
Park, J.H., Lee, K.H., Kim, T.Y., Lee, S.Y., 2007. Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. Proc. Natl. Acad. Sci. USA 104, 7797–7802.
Robichon, C., Luo, J., Causey, T.B., Benner, J.S., Samuelson, J.C., 2011. Engineering Escherichia coli BL21(DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. Appl. Environ. Microbiol. 77, 4634–4646.
Rocke, D., Tillinghast, J., Durbin-Johnson, B., Wu, S.L., n.d. LMGene Software for Data Transformation and Identification of Differentially Expressed Genes in Gene Expression Arrays. R package version 2.4. 0.
Ronda, C., Pedersen, L.E., Sommer, M.O.A., Nielsen, A.T., 2016. CRMAGE: crispr optimized MAGE recombineering. Sci. Rep. 6, 19452.
Saa, P.A., Nielsen, L.K., 2016. Construction of feasible and accurate kinetic models of metabolism: a Bayesian approach. Sci. Rep. 6, 29635.
Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Coldspring-Harbour Laboratory Press, UK.
Sánchez, A.M., Bennett, G.N., San, K.-Y., 2005. Novel pathway engineering design of the anaerobic central metabolic pathway in Escherichia coli to increase succinate yield and productivity. Metab. Eng. 7, 229–239.
Song, Y., Lee, B.-R., Cho, S., Cho, Y.-B., Kim, S.-W., Kang, T.J., Kim, S.C., Cho, B.-K., 2015. Determination of single nucleotide variants in Escherichia coli DH5α by using short-read sequencing. FEMS Microbiol. Lett. 362. http://dx.doi.org/10.1093/femsle/fnv073.
Studier, F.W., Daegelen, P., Lenski, R.E., Maslov, S., Kim, J.F., 2009. Understanding the differences between genome sequences of Escherichia coli B strains REL606 and BL21(DE3) and comparison of the E. coli B and K-12 genomes. J. Mol. Biol. 394, 653–680.
Taylor, R.G., Walker, D.C., Mclnnes, R.R., 1993. E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. Nucleic Acids Res. 21, 1677–1678.
Taymaz-Nikerel, H., van Gulik, W.M., Heijnen, J.J., 2011. Escherichia coli responds with a rapid and large change in growth rate upon a shift from glucose-limited to glucose-excess conditions. Metab. Eng. 13, 307–318.
Vijayendran, C., Polen, T., Wendisch, V.F., Friehs, K., Niehaus, K., Flaschel, E., 2007. The plasticity of global proteome and genome expression analyzed in closely related W3110 and MG1655 strains of a well-studied model organism, Escherichia coli-K12. J. Biotechnol. 128, 747–761.
Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., Church, G.M., 2009. Programming cells by multiplex genome engineering and accelerated evolution. Nature 460, 894–898.
Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J.D., Osterhout, R.E., Stephen, R., Estadilla, J., Teisan, S., Schreyer, H.B., Andrae, S., Yang, T.H., Lee, S.Y., Burk, M.J., Van Dien, S., 2011. Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nat. Chem. Biol. 7, 445–452.
Yoon, S.H., Han, M.-J., Jeong, H., Lee, C.H., Xia, X.-X., Lee, D.-H., Shim, J.H., Lee, S.Y., Oh, T.K., Kim, J.F., 2012. Comparative multi-omics systems analysis of Escherichia coli strains B and K-12. Genome Biol. 13, R37.