Published in final edited form as:

Anal Chem. 2014 October 07; 86(19): 9423–9427. doi:10.1021/ac502734e.

Quantifying Biomass Composition by Gas Chromatography/ Mass Spectrometry

Christopher P. Long and Maciek R. Antoniewicz*

Department of Chemical and Biomolecular Engineering, Metabolic Engineering and Systems Biology Laboratory, University of Delaware, Newark, DE 19716, USA

Abstract

We developed a set of methods for the quantification of four major components of microbial biomass using gas chromatography-mass spectrometry (GC/MS). Specifically, methods are described to quantify amino acids, RNA, fatty acids, and glycogen, which comprise an estimated 88% of the dry weight of *Escherichia coli*. Quantification is performed by isotope ratio analysis with fully ^{13}C -labeled biomass as internal standard, which is generated by growing *E. coli* on $[\text{U-}^{13}\text{C}]$ glucose. This convenient, reliable, and accurate single-platform (GC/MS) workflow for measuring biomass composition offers significant advantages over existing methods. We demonstrate the consistency, accuracy, precision, and utility of this procedure by applying it to three metabolically unique *E. coli* strains. The presented methods will have widespread applicability in systems microbiology and bioengineering.

Graphical abstract

Quantification of the various components of biomass is important for systems biology and bioengineering. The composition of an organism is a core feature of its phenotype, and provides insight into its underlying metabolic systems as well as differences between

*Corresponding author. mraanton@udel.edu. Phone: 302-831-8960. Fax: 302-831-1048.

Supporting Information

Example isotope ratio calculation; validation of RNA and glycogen measurements; validation of method consistency. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Author Contributions

CPL performed the work and wrote the manuscript. MRA oversaw the research.

The authors declare no competing financial interest.

environmental conditions, genotypes, and species. In the field of fluxomics, the resolution and accuracy of metabolic flux models derived from approaches such as flux balance analysis (FBA) and ^{13}C metabolic flux analysis (^{13}C -MFA) are known to be sensitive to biomass composition.^{1,2} Current methods for quantifying biomass are tedious and sometimes inaccurate, which can limit the performance of these fluxomic techniques.

The major components of microbial biomass are protein, RNA, lipids, and glycogen. For the model gram-negative microbe *Escherichia coli*, these four components have been reported to constitute 88% of the dry biomass.³ Current methods to quantify these major biomass components rely on a variety of enzymatic and spectroscopic based methods. Colorimetric assays are often employed to measure total protein content^{4–6}, while amino acid quantification requires hydrolysis followed by HPLC analysis⁷. The measurement of RNA typically requires purification followed by spectroscopic quantification.⁸ While colorimetric methods are also available for total lipid quantification⁹, the use of GC/FID and GC/MS is common for profiling fatty acids^{10–12}. Glycogen is often quantified using enzymatic hydrolysis followed by glucose analysis via HPLC or a colorimetric method.¹³

Here, we have developed a set of methods to quickly, accurately and precisely quantify 17 amino acids, all relevant fatty acids (5 demonstrated here, but easily extendable), RNA, and glycogen on a single, widely available analytical platform: GC/MS. This approach offers a simplified and convenient workflow, and removes reliance on enzymatic and spectroscopic calibrations. All quantifications are based on isotope ratio analysis using analyte-specific standards which are isotopically unique from the sample, giving a high degree of confidence in the results. A similar approach has been previously reported for quantifying metabolite pools.¹⁴ Here, “fully labeled” *E. coli* is used as the internal standard. To generate fully labeled *E. coli* in which all cellular carbon is ^{13}C , a large batch of *E. coli* is grown on $[\text{U-}^{13}\text{C}]\text{glucose}$. The components of this fully ^{13}C -labeled biomass are then quantified against known unlabeled standards. Once characterized, this fully labeled biomass can be used as an internal standard to quantify subsequent unlabeled biomass samples. This procedure adds flexibility and convenience to the workflow, and expands the potential application to organisms that are difficult to label fully.¹⁵ The methods described here are validated and then applied to three *E. coli* strains, confirming consistency, agreement with previously reported values, and demonstrating the practical importance of such measurements in systems biology.

Experimental Section

Materials and Methods

Chemicals—All chemicals were purchased from Sigma-Aldrich (St. Louis, MO). $[\text{U-}^{13}\text{C}]\text{Glucose}$ was purchased from Cambridge Isotope Laboratories (Andover, MA).

Strains and Cultures—*E. coli* BW25113 strains were obtained from the Keio Knockout Collection¹⁶. The parent strain (“wild-type”) was used in all validation studies and as the labeled biomass reference. Two knockouts, *pgi* (*phosphoglucose isomerase*) and *zwf* (*glucose-6-phosphate dehydrogenase*) were also analyzed. All cultures were grown aerobically in M9 minimal medium with 2 g/L of glucose¹⁷, either naturally labeled

(unlabeled) or [U-¹³C] (fully labeled). Cells were harvested at mid-exponential phase ($OD_{600} \approx 0.7$). For all procedures, samples containing the equivalent of 1 mL of a culture at $OD = 1.0$ (roughly 0.3 mg of dry weight) were used. All biomass samples were washed twice with glucose-free M9 medium prior to analysis. Dry weights were measured for all strains by filtration of 50 mL of a culture at a density of $OD = 1.0$ using a 0.2 μ m cellulose acetate filter (Sartorius 11107-47-N), followed by drying for several days at 80°C.

Chemical Standards—Unlabeled standards were prepared for all relevant analytes, and were added directly to the biomass pellet prior to the execution of the protocols described below. For amino acids, 40 μ L of a 2.5 mM per amino acid solution (Pierce 20088) was used. For RNA, a 1 mg/mL solution (ribonucleic acid from torula yeast, Sigma R6625) in water was prepared, of which 80 μ L was added to the sample. For glycogen, a 0.1 mg/mL solution (glycogen from bovine liver, Sigma G0885) in water was prepared, of which 100 μ L was added to the sample. For fatty acids, a solution of 0.3 mg/mL of myristic acid (C14:0), palmitic acid (C16:0), palmitoleic acid (C16:1), stearic acid (C18:0), and elaidic acid (C18:1) in hexane was prepared, of which 20 μ L was added.

GC/MS—GC/MS analysis was performed on an Agilent 7890B GC system equipped with a DB-5MS capillary column (30 m, 0.25 mm i.d., 0.25 μ m-phase thickness; Agilent J&W Scientific), connected to an Agilent 5977A Mass Spectrometer operating under ionization by electron impact (EI) at 70 eV. Helium flow was maintained at 1 mL/min. The source temperature was maintained at 230°C, the MS quad temperature at 150°C, the interface temperature at 280°C, and the inlet temperature at 250°C. For GC/MS analysis of amino acids, 1 μ L was injected at 1:40 split ratio. The column was started at 80°C for 2 min, increased to 280°C at 7°C/min, and held for 20 min. For GC/MS analysis of fatty acid methyl esters (FAME) and sugar derivatives, 1 μ L was injected splitless. The column was started at 80°C for 2 min, increased to 280°C at 10°C/min, and held for 12 min.

Amino Acid Analysis—The preparation and GC/MS analysis of biomass amino acids was performed as previously described by Antoniewicz et. al.¹⁸ Briefly, biomass pellets were hydrolyzed with 500 μ L of 6N HCl at 110°C for 24 h, then dried under air at 65°C. *Tert*-butyldimethylsilyl (TBDMS) derivatives of amino acids were prepared by adding 35 μ L of pyridine and 50 μ L of MTBSTFA + 1% TBDMCS (Sigma 375934) and incubating for 30 minutes at 60°C, and were then transferred to injection vials for GC/MS analysis. 17 of the 20 amino acids were detected and quantified by this method. The three not measured, arginine, cysteine, and tryptophan, are estimated to constitute 12% of total protein mass.³ Since intracellular amino acid pools are known to be very small relative to proteinogenic amino acids¹⁹, the measured signal will be dominated by the latter.

RNA and Glycogen Analysis—These two biopolymers were analyzed simultaneously, as each required hydrolysis followed by quantification of a sugar monomer (ribose and glucose, respectively). Biomass pellets were hydrolyzed by addition of 500 μ L of 6N HCl. Samples were immediately placed under air at 65°C for drying. The aldonitrile propionate derivatives of glucose and ribose were prepared by the method described by Antoniewicz et. al.²⁰ Briefly, 50 μ L of 2 wt% hydroxylamine hydrochloride in pyridine was added to the

dried sample, which was then incubated for 60 min at 90°C. Next, 100 μ L of propionic anhydride was added followed by incubation at 60°C for 30 min. The samples were then transferred to injection vials for analysis. Ribose eluted approximately 2 minutes before glucose, and the m/z 173 fragment was used for quantification, which contains last two carbon atoms of each sugar. We found that other cellular sugars, such as fructose and deoxyribose, do not co-elute with either ribose or glucose and therefore do not interfere with these measurements. It is important to note that we used the biopolymers RNA and glycogen as standards. The use of the monomers ribose and glucose as standards yielded inconsistent and inaccurate results, due to the observed kinetics of hydrolysis and subsequent degradation of the sugars under the acidic conditions.

Fatty Acid Analysis—All glassware was first rinsed with chloroform and dried to remove contaminating lipid residues. The biomass pellets were re-suspended in approximately 500 μ L of water, transferred to a glass Pyrex culture tube, and dried under air at 65°C. FAME derivatives were prepared by dissolving the dried biomass in 1 mL of methanol and 50 μ L concentrated sulfuric acid, and incubating for 2 h at 100°C. The mixture was then cooled to room temperature, and the FAME's were extracted by the addition of 1.5 mL water and 3 mL hexane. The upper organic phase was isolated and dried under nitrogen flow at 40°C. The dried FAME's were then re-dissolved in 100 μ L of hexane and transferred to glass GC vials for analysis. The molecular ions of all species of interest were quantified: C14:0 (M0 m/z 242, fully ^{13}C -labeled M14 m/z 256), C16:1 (m/z 268, 284), C16:0 (m/z 270, 286), C18:1 (m/z 296, 314), C18:0 (m/z 298, 316).

GC/MS Data Analysis—For metabolite quantification using isotope ratio analysis, all measured mass isotopomer distributions were first corrected for natural abundances by the method of Fernandez et. al.²¹ Additionally, the unlabeled (M0) content of fully labeled biomass was assessed (typically ~1–2%, due to presence of unlabeled inoculum) and corrected for. The total ion counts of the labeled species were calculated as the sum of the fully labeled (MN) and one-less (M(N-1)) isotopes. This was necessary to account for the introduction of ^{12}C atoms due to isotopic impurities in the [U- ^{13}C]glucose as well as the fixation of unlabeled CO_2 .²² The frequency of ^{12}C atoms from these sources were low enough such that significant amounts of M(N-2) isotopes were not observed, and thus this effect could be completely accounted for by the stated methods (see Supplemental Information for more details and an example isotope ratio calculation).

Validation of RNA and Glycogen Measurements—For RNA and glycogen quantification, care was taken to ensure that the target macromolecules were being measured, and not for example intracellular sugars such as ribose, glucose, or fructose phosphates. To validate this, a labeling switch experiment was performed. A culture of wild-type *E. coli* was grown on unlabeled glucose to OD 0.5, and then centrifuged, washed, and re-suspended in medium containing [U- ^{13}C]glucose as the only carbon source. Time-course data of labeling incorporation confirmed that low-turnover biomass components were indeed measured.

Validation of Method Consistency—The methods described here depend on the quantification of isotope ratios of ^{13}C -labeled and unlabeled species in a sample. In practice, an unlabeled biomass sample can be measured against a labeled reference, or vice versa. Since most chemical standards are unlabeled, it is straightforward to quantify fully labeled biomass. To measure unlabeled biomass samples, we first grew a large batch of fully labeled *E. coli* biomass, and aliquoted a large number of identical samples, each containing an equivalent of 1 mL of a culture at OD = 1.0. The aliquoted biomass pellets were stored at -80°C . This stock of fully labeled biomass was characterized, and then used as a reference with which to measure subsequent unlabeled biomass samples. In this case, the reference and sample biomass pellets were combined directly at the beginning of the workup. To confirm that both methods yielded consistent results, *E. coli* was grown in two parallel cultures, one on unlabeled glucose and one on $[\text{U-}^{13}\text{C}]$ glucose. The biomass compositions of both cultures were characterized by the two respective methods and we confirmed that both methods yielded consistent results.

Results and Discussion

Method Validation

First, we validated that the methods for RNA and glycogen quantification were indeed measuring these low-turnover components of biomass. This was determined by measuring time-course labeling profiles in an experiment where *E. coli* was first grown on unlabeled glucose followed by growth on fully ^{13}C -labeled glucose. The fractional labeling of both RNA and glycogen tracked well with the fraction of new biomass formed after the switch, over a period of three hours (Supporting Information). Intracellular metabolites, which turn over in minutes, would have become fully labeled in this time. Only a small fraction of the measured values, less than 15% for glucose and less than 5% for ribose, reflected fast-turnover metabolites. Therefore we concluded that the methods we applied were almost entirely measuring the targeted biomass components RNA and glycogen.

Second, we validated that the presented methods gave consistent results regardless of whether a ^{13}C -labeled or unlabeled internal standard was used for quantification, i.e. 1) using unlabeled chemical standards for quantifying a labeled biomass sample; and 2) using fully labeled biomass as reference material for the analysis of unlabeled biomass sample. This was shown by culturing *E. coli* in parallel on unlabeled and fully labeled glucose and applying both approaches. As expected, both approaches were found to yield identical biomass composition values (Supporting Information). This result shows that the methods described here are consistent and flexible, and can be adapted for the convenience of the user. For example, the use of a fully labeled reference biomass stock may be preferable if it is found that frozen biomass is simpler to store or more stable over time than unlabeled standard solutions. More significantly, it may enable the convenient analysis of species which are difficult to fully label, such as organisms that require complex medium for growth²³, for which labeled standards may otherwise be prohibitively expensive to generate.

Biomass Composition of Three *E. coli* Strains

Phenotypic differences between different genotypes, particularly gene knockouts, are of significant interest in systems biology.^{24,25} To demonstrate the utility of these methods, the biomass compositions of three *E. coli* strains were analyzed: wild-type (WT), and two knockout strains, *zwf* and *pgi*. These knockouts are missing the first reaction in the oxidative pentose phosphate pathway and glycolysis, respectively. As such, they have significantly altered metabolic states from the wild-type and each other. The overall compositional profiles of the three strains, shown in Figure 1, agree well with the established literature values from Neidhardt³. The total protein level was ~53% of dry weight for WT, which is comparable to Neidhardt's 55%. For total cell protein quantification, Neidhardt's values were assumed for the 3 unmeasured amino acids (other estimation approaches for these may be considered in the future, such as bioinformatic techniques to derive relative amino acid abundances from protein sequence data). There was slightly less protein in the knockouts, with 49% and 47% for *zwf* and *pgi*, respectively. There were a few differences in the amino acid profiles between the strains (Figure 2), such as elevated Glx (Glu+Gln) levels in WT compared to the knockouts. Neidhardt reported significantly higher levels of glycine, valine, isoleucine and lysine than what were measured in these *E. coli* strains. RNA was significantly reduced in the *pgi* strain, at 14% of dry weight compared to 24% in the wild-type and 21% for *zwf*. The total fatty acids were a consistent 5% of dry weight for all three strains, slightly less than the 7% of Neidhardt. The distribution of fatty acids is shown in Figure 3. There was significant variability in C18:1 levels between the three strains, with the *zwf* strain showed elevated levels while *pgi* showed decreased levels relative to the wild-type. *pgi* also had less C16:1 than the other two strains. The wild-type fatty acid levels, with the exception of C14:0 which is primarily associated with LPS, were less than the values reported by Neidhardt. This is likely due to differences in the *E. coli* strains characterized (K-12 in this study and B/r in Neidhardt).

Overall, the measured composition values for *E. coli* are in good agreement with those previously reported, while also demonstrating the importance of measuring the composition of novel phenotypes. This is often neglected in flux analysis and system biology studies.²⁶ Instead, Neidhardt's values, while summarized specifically for the *E. coli* B/r strain, are often assumed for other *E. coli* strains. Performing the biomass composition measurements presented here will therefore be useful for future systems microbiology applications, and in particular for advanced ¹³C metabolic flux analysis studies.²⁷

Conclusions

The methods presented here for biomass composition analysis are accurate and precise, as well as convenient and flexible. The use of a single analytical platform, GC/MS, and simple preparatory protocols means that a full biomass composition analysis (of the four major components) can be easily accomplished in many labs. The coverage of these techniques should be sufficient for most research endeavors in systems microbiology and engineering. The major components not explicitly measured here are DNA (3.1% of dry weight), lipopolysaccharide (3.4%), peptidoglycan (2.5%), and intracellular metabolites, cofactors,

and ions (3.5%).³ In most circumstances, the amounts of these components are expected to be relatively constant.^{1,28}

The choice of fully labeled *E. coli* biomass as internal standard gives flexibility with respect to shelf-life and cell culturing concerns. For extension to other organisms, it is expected that the hydrolysis conditions used here will result in no significant differences in lysis or hydrolysis kinetics that might affect accuracy. Thus, we expect equally good performance with other classes of microbes, e.g. gram-positive bacteria and eukaryotic cells, although this should be validated prior to application. If verified, the convenience of using [U-¹³C] labeled *E. coli* as internal standard, which is quickly and easily cultured, for analyzing more difficult organisms would offer a significant advantage.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NSF CAREER award (CBET-1054120).

References

1. Pramanik J, Keasling JD. *Biotechnol. Bioeng.* 1997; 56:398–421. [PubMed: 18642243]
2. Pramanik J, Keasling JD. *Biotechnol. Bioeng.* 1998; 60:230–238. [PubMed: 10099424]
3. Neidhardt, FC. *Escherichia coli* and *Salmonella Typhimurium*. ASM Press; 1987.
4. Bradford M. *Anal. Biochem.* 1976; 72:248–254. [PubMed: 942051]
5. Peterson GL. *Anal. Biochem.* 1977; 83:346–356. [PubMed: 603028]
6. Smith P, Krohn R, Hermanson G, Mallia A, Gartner F, Provenzano M, Fujimoto E, Goeke N, Olson B, Klenk D. *Anal. Biochem.* 1985; 150:76–85. [PubMed: 3843705]
7. Rutherford SM, Gilani GS. *Curr. Protoc. Protein Sci.* 2009; 58:11.9.1–11.9.37.
8. Benthin S, Nielsen J, Villadsen J. *Biotechnol. Tech.* 1991; 5:39–42.
9. Wawrik B, Harriman BH. *J. Microbiol. Methods.* 2010; 80:262–266. [PubMed: 20093146]
10. Dobson G, Christie WW. *Eur. J. Lipid Sci. Technol.* 2002; 104:36–43.
11. Dodds ED, McCoy MR, Rea LD, Kennish JM. *Lipids.* 2005; 40:419–428. [PubMed: 16028722]
12. Li Y, Beisson F, Pollard M, Ohlrogge J. *Phytochemistry.* 2006; 67:904–915. [PubMed: 16600316]
13. McKinlay JB, Shachar-Hill Y, Zeikus JG, Vieille C. *Metab. Eng.* 2007; 9:177–192. [PubMed: 17197218]
14. Bennett BD, Yuan J, Kimball EH, Rabinowitz JD. *Nat. Protoc.* 2008; 3:1299–1311. [PubMed: 18714298]
15. Swarup A, Lu J, DeWoody KC, Antoniewicz MR. *Metab. Eng.* 2014; 24:173–180. [PubMed: 24909362]
16. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. *Mol. Syst. Biol.* 2006; 2:1–11.
17. Leighty RW, Antoniewicz MR. *Metab. Eng.* 2013; 20:49–55. [PubMed: 24021936]
18. Antoniewicz MR, Kelleher JK, Stephanopoulos G. 2007; 79:7554–7559.
19. Bajad SU, Lu W, Kimball EH, Yuan J, Peterson C, Rabinowitz JD. *J. Chromatogr. A.* 2006; 1125:76–88. [PubMed: 16759663]
20. Antoniewicz MR, Kelleher JK, Stephanopoulos G. *Anal. Chem.* 2011; 83:3211–3216. [PubMed: 21413777]

21. Fernandez, Ca, Des Rosiers, C., Previs, SF., David, F., Brunengraber, H. J. Mass Spectrom. 1996; 31:255–262. [PubMed: 8799277]
22. Leighty RW, Antoniewicz MR. Metab. Eng. 2012; 14:533–541. [PubMed: 22771935]
23. Ahn WS, Antoniewicz MR. Metab. Eng. 2013; 15:34–47. [PubMed: 23111062]
24. Long CP, Antoniewicz MR. Curr. Opin. Biotechnol. 2014; 28:127–133. [PubMed: 24686285]
25. He L, Xiao Y, Gebreselassie N, Zhang F, Antoniewicz MR, Tang YJ, Peng L. Biotechnol. Bioeng. 2013:1–11.
26. Crown SB, Antoniewicz MR. Metab. Eng. 2013; 20:42–48. [PubMed: 24025367]
27. Crown SB, Antoniewicz MR. Metab. Eng. 2013; 16:21–32. [PubMed: 23246523]
28. Stephanopoulos G, Aristidou A, Nielsen J. Metabolic Engineering: Principles and Methodologies. 1998:66.

Figure 1.
Biomass composition of three *E. coli* strains. Error bars indicate standard error of the mean (n=4).

Figure 2.
Amino acid distribution of three *E. coli* strains. The key is the same as Fig. 1; error bars indicate standard error of the mean (n=4).

Figure 3.
Fatty acid distribution of three *E. coli* strains. The key is the same as Fig. 1; error bars indicate standard error of the mean (n=4).