Stable Isotope-Assisted Metabolomics for Network-Wide Metabolic Pathway Elucidation

Darren J. Creek,†‡ Achuthanunni Chokkathukalam,§ Andris Jankevics,§∥ Karl E. V. Burgess,† Rainer Breitling,§∥ and Michael P. Barrett*†

†Wellcome Trust Centre for Molecular Parasitology, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, G12 8TA, U.K.
‡Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Flemington Rd, Parkville, Victoria, 3010, Australia
§Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, G12 8QQ, U.K.
∥Groningen Bioinformatics Centre, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

ABSTRACT: The combination of high-resolution LC–MS-based untargeted metabolomics with stable isotope tracing provides a global overview of the cellular fate of precursor metabolites. This methodology enables detection of putative metabolites from biological samples and simultaneous quantification of the pattern and extent of isotope labeling. Labeling of Trypanosoma brucei cell cultures with 50% uniformly 13C-labeled glucose demonstrated incorporation of glucose-derived carbon into 187 of 588 putatively identified metabolites in diverse pathways including carbohydrate, nucleotide, lipid, and amino acid metabolism. Labeling patterns confirmed the metabolic pathways responsible for the biosynthesis of many detected metabolites, and labeling was detected in unexpected metabolites, including two higher sugar phosphates annotated as octulose phosphate and nonulose phosphate. This untargeted approach to stable isotope tracing facilitates the biochemical analysis of known pathways and yields rapid identification of previously unexplored areas of metabolism.

Recent advances in metabolomics technologies have enabled the analysis of metabolism on a global scale, allowing detection and semiquantification of hundreds of low molecular weight (<1000 Da) molecules in a single experiment.1,2 However, interpretation of these results remains dependent on our knowledge of active metabolic pathways, which is far from complete for many organisms.3–4 The ab initio identification of novel pathways remains a challenge. Stable isotope tracing is an established technique for determining the fate of individual metabolites, but it generally requires targeted analysis of predicted products.5,6 The common technology that underlies both stable isotope tracing and metabolomics (i.e., mass spectrometry) provides the opportunity to combine these two approaches, allowing network-wide investigation of metabolic pathways.

High-resolution mass spectrometry offers great potential for untargeted metabolomics, as high mass accuracy often (but not always) allows direct formula assignment for each detected mass (m/z).1,2,7,8 The complexity of liquid chromatography–mass spectrometry (LC–MS) data and the existence of many isobaric metabolites hinders unambiguous identification of all metabolite peaks, but recent advances in data processing allow automated identification and annotation of many metabolites with improved confidence.9,10

A remaining limitation regarding biological interpretation of metabolomics data is the dependence on predetermined...
metabolic pathways, often overlooking the potential involvement of novel pathways. Even in well-defined organisms, detected metabolites may not derive from the classical pathways, as metabolites may be imported from an exogenous source, produced biosynthetically by one or more pathways, or derived from different sources in different compartments within the same cell. For example, in our model organism *Trypanosoma brucei*, inositol can be synthesized in the Golgi apparatus for glycosylphosphatidylinositol anchor synthesis or imported for use in bulk inositol lipid synthesis.\(^1\)

The placement of metabolites in pathways is significantly improved by stable isotope tracing, which has recently elucidated some important descriptions of central metabolism.\(^6\)–\(^14\) For example, labeling patterns of aspartate and tricarboxylic acid (TCA) cycle intermediates in human lung cancer cells revealed anaplerosis via activation of pyruvate carboxylase,\(^15\) while in yeast a novel ribosegenesis pathway was confirmed by metabolite labeling.\(^2\) The major limitation with current isotope tracing approaches is the reliance on targeted analysis of labeled metabolites, making it impossible to trace precursor distribution into unexpected areas of metabolism. Nonetheless, these studies have already demonstrated that our classical view of metabolism in many organisms is overly simplistic, making an expansion of this approach beyond central metabolism highly desirable. An untargeted isotope tracing approach can be expected to reveal novel areas of metabolic networks that are essential to include in systems-based studies of metabolism.

Stable isotopes are becoming more commonly used in untargeted metabolomics studies to facilitate the quantification and identification of metabolites. Isotope labeled metabolite extracts can assist with an elemental formula assignment\(^16\),\(^17\) or can provide internal standards for quantitative LC–MS-based metabolomics by isotope dilution.\(^17\) Fully labeled extracts have also been applied to the differentiation of biogenic and exogenous features in LC–MS data.\(^18\),\(^19\) Stable isotopes have been applied to targeted metabolomics studies for stable isotope tracing and, in combination with quantification, for flux analysis.\(^5\),\(^12\),\(^20\) Expansion of these studies to the whole metabolic network requires accurate and efficient measurement of stable isotope distributions for all metabolites. Untargeted detection of isotopomer distributions has been demonstrated for gas chromatography–mass spectrometry (GC–MS); however, interpretation is difficult for unknown metabolites if they are not represented in spectral libraries.\(^21\),\(^22\) Here we introduce a method that uses high-resolution LC–MS to perform untargeted detection of isotopic patterns for stable isotope labeled metabolites in complex mixtures and thus enables identification of novel metabolic pathways without preconception of the fate of labeled precursors.

**EXPERIMENTAL SECTION**

**Parasite Culture, Metabolite Extraction and Sample Analysis.** Labeling of cellular metabolites was achieved by growth of the procyclic-form *Trypanosoma brucei* under standard cell culture conditions in SDM-79 medium,\(^23\) with 10 mM additional uniformly \(^13\)C-labeled glucose (~50% of total glucose labeled), for 5 days (10 doublings) to achieve close to steady-state labeling. A parallel culture was prepared under identical conditions, except with 10 mM additional unlabeled glucose to allow identification of unlabeled metabolites. A total of \(10^8\) parasites were quenched by rapid cooling to 0 \(^\circ\)C,\(^13\) growth medium was removed after centrifugation, and metabolites extracted with a monophasic chloroform/methanol/water (1:3:1) mixture, which extracts both polar and nonpolar metabolite species, as previously described.\(^24\) LC–MS analysis was performed with ZIC-HILIC hydrophilic interaction liquid chromatography (Merck Sequant), coupled to high-resolution Exactive Orbitrap mass spectrometry (Thermo, Hemel Hempstead, U.K.) operating in both positive and negative ionization modes, according to our published method.\(^24\) Samples were analyzed in triplicate.

**Metabolomics Data Processing.** Initial data processing of raw LC–MS data (mzXML format) was performed using the standard IDEOM (http://mzmatch.sourceforge.net/ideom.php) workflow as described previously.\(^9\) This workflow utilizes the XCMS Centwave algorithm for peak detection and mzMatch.R for alignment of samples, filtering, and metadata storage in peakML files.\(^26\) Only metabolites detected in the unlabeled cell extracts were retained for further analysis. Parameters used for noise filtering and (putative) metabolite identification are available in Supplementary Table 1 in the Supporting Information.

**Detection of Stable Isotope-Labeled Metabolites.** The peakML file after filtering and identification was scanned for labeled metabolites using the IDEOM Isotope Search and the PeakMLIsotope.UntargettedIsotope function of mzMatch-ISO (http://mzmatch.sourceforge.net/isotopes-targetted.php). The algorithm for detection of isotope labeling begins by first determining the number of possible labeled isotopomers a metabolite can have, according to its molecular formula or mass (e.g., \(\text{C}_4\text{H}_8\text{O}_3\) could have 1, 2, or 3 carbons labeled). Then the expected mass of each labeled isotopomer is calculated based on the known mass difference between the light and heavy isotopes. Peaks are assigned as isotopomers if their mass is within 4 ppm (the mass accuracy of the spectrometer) of the predicted mass and within a retention time window of \(\pm 0.2\) min (the average peak width) from the unlabeled peak, as isotopomers are expected to co-elute. This procedure is repeated for every possible isotopomer of each putatively annotated (unlabeled) metabolite to provide a comprehensive list of relative isotopomer intensities for each metabolite. If necessary, isotopomer signals that are missed in the initial processing, e.g., because of low intensities or irregular peak shapes, are gap-filled from raw data in a targeted manner by mzMatch-ISO. The accuracy and precision of the approach were validated by analysis of the natural \(^13\)C isotope \((\text{M} + 1.0033)\) in 60 identified metabolites in the unlabeled samples \((n = 3)\). The measured relative isotope abundance was \(1.1\% \pm 1.3\%\) lower than the theoretical isotope abundance \((\text{mean \pm standard deviation for 60 metabolites})\), and the average standard deviation of experimental isotope abundances was \(0.4\%\) \((\text{range 0.005−1.8\%})\).

Data are either visualized in IDEOM within Excel, showing the relative intensity of labeled isotopomers (see Supplementary File 1 (Excel spreadsheet) in the Supporting Information), or in mzMatch-ISO-generated PDF files containing detailed results for each metabolite. These results include raw chromatograms representing the monoisotopic and corresponding isotopic peaks; a normalized plot showing the variability in labeling between replicates; a trend plot of the pattern of labeling in each sample group; and a plot quantifying the absolute labeling pattern of a selected isotopomer of interest (see Supplementary File 2 (pdf file) in the Supporting Information). Both of these data outputs were collated to generate a comprehensive list of labeled metabolites. The
procedure can be applied to $^{13}$C, $^{15}$N, $^2$H, or $^{18}$O isotope labels or user-defined mass differences, e.g., for chemical tags.27

**RESULTS AND DISCUSSION**

The method requires growth of the organism of interest in the presence of a predefined proportion (ideally 50%) of a stable isotope labeled nutrient. In our proof-of-concept example, we used procyclic-form *Trypanosoma brucei*, the protozoan parasite responsible for African sleeping sickness, which due to its reduced and well studied metabolism is a suitable model organism for validating our approach.8 Growth medium containing 50% uniformly $^{13}$C-labeled glucose as the main carbon source was used to generate steady-state labeling of the cellular metabolome. A parallel culture was incubated with unlabeled glucose at the same concentration. Quenching and metabolite extraction was achieved by a nonspecific metabolite extraction protocol, and LC–MS analysis utilized zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC) coupled to high-resolution mass spectrometry using Orbitrap technology.24 Data were analyzed using a combination of freely available software, XCMS,28 mzmatch.R,26 and IDEOM,9 with customized additions to detect and quantify isotopomer signals.

In our example, over 27,000 unique peak groups (defined by $m/z$ ratio and retention time) were detected in both positive and negative mode ionization. The vast majority of these peaks are noise or artifacts arising from the mass spectrometry, chromatography, sample preparation, and data processing (see Supplementary Figure 1 in the Supporting Information).10,24 A key component of this method is application of the default noise filters of mzMatch.R and IDEOM to obtain a list of monoisotopic peaks representing putative metabolites present in the label-free sample. A total of 82 metabolites were

---

**Figure 1.** (a) Schematic representation of the method. Red dots represent labeled metabolites. (b) Labeling pattern in succinate, aspartate, and orotate biosynthesis. In contrast to the expected predominant 2- and 4-carbon labeling from the TCA cycle, the majority of molecules contain three labeled carbons, consistent with a fermentative source of succinate. (c) Heavy-isotope labeling patterns in the new trypanosomal metabolites putatively annotated as octulose phosphate and nonulose phosphate. C represents the unlabeled control, and L represents the $^{13}$C-glucose labeled sample.
identified confidently by exact mass and retention time based on authentic metabolite standards (level 1 identification according to the Metabolomics Standards Initiative (MSI)29). A total of 506 putative annotations (levels 2 and 3, MSI) were made using exact mass and predicted retention times from the IDEOM database (see Supplementary File 1 (Excel spreadsheet) in the Supporting Information).9,24

The extent and pattern of labeling in each (putative) metabolite was determined by an automated search of all potential isotopomers according to accurate mass and retention time. Comparison of isotope distributions with unlabeled samples, and with theoretical natural isotope abundances, confirmed the presence of stable isotope labeling for 187 metabolites (Figure 1a). Visualization of labeled metabolites in the global metabolic network30 reveals incorporation of glucose-derived carbon into diverse areas of metabolism (Figure 2).

The isotope labeling patterns generated from 50% U-13C-labeled glucose allowed immediate identification of many active metabolic pathways within the cell. For example, the distinct labeling pattern for succinate (Figure 1b) of predominantly three labeled carbons is consistent with previous targeted studies of trypanosomes that demonstrated a fermentation pathway primarily responsible for succinate production, rather than a traditional TCA cycle source.31,32 The results also show a similar labeling pattern in aspartate and orotate, confirming biosynthesis of aspartate from phosphoenolpyruvate (via oxaloacetate) and subsequent de novo pyrimidine synthesis. The enzymes required for de novo pyrimidine synthesis have been identified and studied in bloodstream form trypanosomes,15,34 and here we provide the first direct evidence for de novo pyrimidine synthesis from glucose-derived aspartate in the procyclic life-cycle stage despite the high abundance of aspartate and pyrimidines available in the growth medium.

Extensive labeling was observed in nucleotides and lipids (Figure 2 and Supplementary File 1 (Excel spreadsheet) in the Supporting Information). Nucleotide labeling patterns confirm both de novo pyrimidine synthesis (vide supra) and active purine salvage pathways incorporating ribose 5-phosphate (five labeled carbons) from the pentose phosphate pathway (PPP), while glycerophospholipid labeling patterns revealed incorporation of three carbons from glycerol 3-phosphate or dihydroxyacetone phosphate (products of glycolysis) and also two carbon lipid chain extensions from acetyl-CoA.35 While acetyl-CoA was not observed in this study (being incompatible with this chromatographic method), its presence can be inferred from the two-carbon incorporation into lipids and by the presence of two-carbon labeled acetylylsine and acetylcarnitine, although the relatively low level of label incorporation into these metabolites (15% and 21%, respectively) confirms previous reports that glucose is not the primary source of acetyl-CoA.36 A single LC column that optimally separates polar and nonpolar metabolites has yet to be successfully employed in metabolomics research. Notwithstanding, our nontargeted approach is clearly capable of identifying metabolites from a broad physicochemical spectrum and can be readily adapted to protocols involving multiple columns to separate metabolites if even more comprehensive coverage is required.

A key feature of the untargeted metabolomics approach is the discovery of unexpected labeling in metabolites and of labeling in unexpected metabolites. The most striking cases of unexpected labeled metabolites detected in our case study

Figure 2. KEGG global pathway map of identified and putatively annotated metabolites labeled by 13C-glucose (red) and unlabeled putative metabolites arising from alternative sources (gray). Inferred potentially active biosynthesis pathways are highlighted; black routes have not previously been known to be active in Trypanosoma brucei. Map generated using IPath tools.30

dx.doi.org/10.1021/ac3018795 | Anal. Chem. 2012, 84, 8442–8447
included glycerate and gluconate, potentially arising from dephosphorylation of glycolysis-derived phosphoglycerate and PPP-derived phosphogluconate, respectively, although specific enzymes for these reactions have not been identified in T. brucei. Importantly, prior knowledge of putative metabolite identities is not essential for detection of isotopomers with our method, and isotopomer profiling of unidentified peaks facilitates the discovery of novel metabolites. In this example, two labeled metabolites that were not present in online metabolism databases were detected with m/z retention time, and labeling patterns consistent with octulose phosphate and nonulose phosphate (Figure 1c). These metabolites are highly relevant for trypanosome biology as they indicate the presence of new links between the glycolytic and pentose phosphate pathways analogous to recent findings in yeast, and further work on these novel pathways may enable the discovery of new targets for trypanosomal drugs.

**CONCLUSIONS**

As we show, untargeted LC–MS-based metabolomics is able to successfully map the extent of metabolism of stable isotope-labeled glucose in procyclic-form trypanosomes. The results, obtained in a single experiment, are generally consistent with many years of prior research and genome annotations and also highlight new areas of metabolism that would not have been detected by a traditional targeted approach. The necessary computational analysis is straightforward and has been automated and included in the mzMatch.R and IDEOM applications (http://mzmatch.sourceforge.net/), making it easy to apply the same strategy to any organism to confirm the presence of metabolic pathways, discover new pathways, confirm metabolic activity under specific growth conditions or, by inclusion of multiple samples at specified time points with concurrent quantitative analysis, measure flux through a network.

**ASSOCIATED CONTENT**

1. Supporting Information

Supplementary Figure 1, distribution of sources of noise peaks removed by automated filtering; Supplementary Table 1, table of parameters applied for data processing with XCMS, mzMatch, and IDEOM; Supplementary File 1, Excel spreadsheet summary of identified and putatively annotated metabolites in procyclic-form *Trypanosoma brucei*; 13C-labeled metabolites are annotated in column J, detailed peak intensities for each metabolite and isotopomer are listed in the “RAWisotopomerPEAKS” sheet (All data from this study, and full functionality for data analysis and visualization in Excel, are available in the IDEOM file (http://puma.ibis.gla.ac.uk/untargeted/IDEOM_13Cglucose_PCF.zip), and Supplementary File 2, PDF file containing all mass traces and graphs for isotopomers of identified and annotated metabolites. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

*Corresponding Author*

*E-mail: Michael.barrett@glasgow.ac.uk.*

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

D.J.C. is supported by an NHMRC postdoctoral training fellowship. Metabolomics analysis was provided by the Scottish Metabolomics Facility. This work was partly supported by the Wellcome Trust through The Wellcome Trust Centre for Molecular Parasiotology, which is supported by core funding from the Wellcome Trust (Grant 085349).

**REFERENCES**

(9) Creek, D. J.; Jankevics, A.; Burgess, K. E. V.; Breitling, R.; Barrett, M. P. *Bioinformatics* 2012, 28, 1048–1049.