Molecular BioSystems

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

Cite this: *Mol. BioSyst.*, 2013, **9**, 1743

Large-scale metabolome analysis and quantitative integration with genomics and proteomics data in *Mycoplasma pneumoniae*†

Tobias Maier,*[ab] Josep Marcos,[cd] Judith A. H. Wodke,[abe] Bernhard Paetzold,[ab] Manuel Liebeke,[f] Ricardo Gutiérrez-Gallego*[cd] and Luis Serrano*[abg]

Received 19th March 2013,
Accepted 27th March 2013

DOI: 10.1039/c3mb70113a

www.rsc.org/molecularbiosystems

Systems metabolomics, the identification and quantification of cellular metabolites and their integration with genomics and proteomics data, promises valuable functional insights into cellular biology. However, technical constraints, sample complexity issues and the lack of suitable complementary quantitative data sets prevented accomplishing such studies in the past. Here, we present an integrative metabolomics study of the genome-reduced bacterium *Mycoplasma pneumoniae*. We experimentally analysed its metabolome using a cross-platform approach. We explain intracellular metabolite homeostasis by quantitatively integrating our results with the cellular inventory of proteins, DNA and other macromolecules, as well as with available building blocks from the growth medium. We calculated *in vivo* catalytic parameters of glycolytic enzymes, making use of measured reaction velocities, as well as enzyme and metabolite pool sizes. A quantitative, inter-species comparison of absolute and relative metabolite abundances indicated that metabolic pathways are regulated as functional units, thereby simplifying adaptive responses. Our analysis demonstrates the potential for new scientific insight by integrating different types of large-scale experimental data from a single biological source.

Introduction

Recent advances in analytical techniques to monitor cellular macromolecules, such as mRNA or proteins on a large scale (transcriptomics and proteomics), are revolutionizing our understanding of living systems. The high-throughput analysis of cellular metabolites (metabolomics) is somewhat lagging behind these successes in adjacent fields, although both metabolomics and metabolic flux analysis provide valuable information on the phenotypic state of a biological system.[1–5] Several reasons are responsible for this disequilibrium: while mRNA and proteins are homogeneous molecules (*i.e.* they are made from a basic set of building blocks), cellular metabolites are chemically diverse and of vastly different cellular abundance. Therefore it is technically not feasible to address all compounds for comprehensive metabolomics profiling, using a single analytical approach.[6–9] Other than mRNA and proteins, metabolites are not directly encoded by the cellular DNA and their occurrence largely depends on transient cellular requirements. Additionally, metabolites are often either short-lived or unstable in a cellular context, thereby posing great challenges for sample preparation and processing prior to analysis.[10,11]

A further aspect limiting the analysis of a biological sample using metabolomics approaches is related to the complexity of the model system to be studied. For example, predictions of the metabolic inventory of homo sapiens based on an *in silico* metabolic network reconstruction and literature mining resulted in the functional annotation of 2766 different metabolites.[12] The metabolome complexity of prokaryotic organisms ranges from 240 predicted metabolites in the obligate insect pathogen *Buchnera aphidicola* to 1658 predicted metabolites in *Klebsiella pneumoniae*.[13]

[a] *EMBL/CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), Dr Aiguader 88, 08003 Barcelona, Spain. E-mail: tobias.maier@crg.es, luis.serrano@crg.es*
[b] *Universitat Pompeu Fabra (UPF), Barcelona, Spain*
[c] *Department of Experimental and Health Sciences, Pompeu Fabra University, Barcelona Biomedical Research Park, Dr Aiguader 88, 08003 Barcelona, Spain. E-mail: rgutierrez@imim.es*
[d] *Bio-analysis Group, Neuroscience Research Program, IMIM-Parc Salut Mar, Barcelona Biomedical Research Park, Dr Aiguader 88, 08003 Barcelona, Spain*
[e] *Theoretical Biophysics, Humboldt-Universität zu Berlin, Berlin, Germany*
[f] *Department of Surgery and Cancer, Imperial College London, Sir Alexander Fleming Building, London SW7 2AZ, UK*
[g] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluis Companys 23, 08010 Barcelona, Spain*

† Electronic supplementary information (ESI) available: Fig. S1–S7. See DOI: 10.1039/c3mb70113a

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

We used the bacterium *Mycoplasma pneumoniae* for a large-scale, experimental analysis of the cellular metabolome. *M. pneumoniae* represents an ideal model system to study the cellular metabolome on an organism-wide scale: it is of generally low complexity due to its reduced genome encoding only 689 open reading frames[14,15] yet, it can be grown in pure culture under laboratory conditions not depending on co-cultivation of host cells.[16] *M. pneumoniae* lacks a rigid cell wall and adheres to culture dishes for growth. Therefore, sample preparation for metabolite analysis, involving washing steps, rapid quenching of the metabolism and cell lysis, is greatly facilitated.[11,17] Furthermore, *M. pneumoniae* is an established model organism for system-wide studies; its genome is sequenced,[14,15] large-scale transcriptomics and proteomics data sets are published,[18–20] protein post-translational modifications are annotated[21] and a hand-curated metabolic map of the *M. pneumoniae* metabolism is available (Fig. S1, ESI†).[22] Due to the genome size reduction during evolution, *M. pneumoniae* possesses a streamlined metabolism with only few metabolic modules (pathways where metabolic intermediates are chemically related). It utilizes glucose as a main carbon source and ATP is predominantly generated by substrate level phosphorylation during glycolysis. The bacterium lacks a TCA cycle as well as a respiratory chain. As a consequence, *M. pneumoniae* needs to metabolize large amounts of glucose to produce the metabolic end-products lactic acid and acetic acid in order to generate the ATP necessary for growth and cellular functions. Accordingly, enzymes involved in glucose metabolism are highly abundant in the *M. pneumoniae* cytosol.[19]

Here we report organism-wide identification of the *M. pneumoniae* metabolome using a cross-platform approach by combining results obtained from gas chromatography coupled to mass spectrometry (GC-MS), liquid chromatography coupled to mass spectrometry (LC-MS), and nuclear magnetic resonance spectroscopy (NMR) measurements. We characterize the metabolome by comparing our results to studies in other bacteria and by integrating our data with a metabolic map[22] as well as with protein abundances of *M. pneumoniae*. We quantify intermediates of key metabolic pathways by GC-MS using external calibration curves and internal standards for cross-analysis referencing. Additionally, we integrate quantitative measurements of intracellular pools of amino acids and nucleotides with their respective abundances in the growth medium, the proteome and the genome, respectively. This allows us to quantitatively analyze cellular import mechanisms and cytoplasmic metabolite homeostasis as well as to elucidate how *M. pneumoniae* globally manages the supply and demand of these basic building blocks of the cell.

The *M. pneumoniae* metabolome

The genome-scale constraint-based reconstruction of the *M. pneumoniae* metabolic network contains 189 reactions with 216 different reactants.[22] We extracted a list of 150 metabolic intermediates based on this metabolic reconstruction, excluding reactants, such as tRNAs, protein-based metabolites and inorganic compounds not amenable to experimental analysis (Table S1, ESI†). A comparison of the thus extracted *M. pneumoniae in silico* metabolome with predicted and experimentally confirmed metabolites of *Escherichia coli*, *Bacillus subtilis* and *Saccharomyces cerevisiae* showed that 95% of all predicted *M. pneumoniae* metabolites form also part of the *in silico* metabolome of those organisms (6 of 8 metabolites not predicted by van der Werf and colleagues[23] are intermediates of Mycoplasma-specific phospholipid metabolism).

In an effort to experimentally characterize the *M. pneumoniae* metabolome, we identified 86 different cellular metabolites in an organism-wide screen, combining GC-MS, LC-MS and NMR measurements (Fig. 1a and Table S2, ESI†). These cross-platform approaches have been shown to be obligate in the analysis of complex metabolite samples.[24] Using three complementary techniques and both untargeted and directed identification approaches, we identified 53% of all predicted *M. pneumoniae* metabolites (Fig. 1b (ref. 22)), surpassing experimental coverage of predicted metabolites in other unicellular organisms[23,25–28] (Fig. 1b and Table S3, ESI†). 83% of experimentally detected metabolites in this study were previously identified in large-scale metabolic screens of other bacteria[23,25,26] (Fig. 1c).

In total, 73 *in silico* predicted metabolites of the *M. pneumoniae* metabolic network were not detected using either GC-MS, LC-MS or NMR analysis. 49 of those metabolites can technically be identified using corresponding commercially available pure compounds as standards.[23,26] However, only 16 of those predicted but not identified metabolites in *M. pneumoniae* were experimentally confirmed previously in *E. coli*, *B. subtilis* or *S. aureus*[23,25,26] (Fig. 1c). Those compounds predominantly map to the pentose phosphate pathway and the CoA-metabolism (Fig. 2). In fact, experiments with ^{13}C labelled glucose show that the outflux from glycolysis to the pentose phosphate pathway is very slow in *M. pneumoniae* (http://www.nature.com/msb/journal/v9/n1/full/msb20136.html). Thus, in addition to technical constraints of the used analysis methods, *M. pneumoniae* metabolites escaping experimental identification are most likely to be present below the limit of detection, either because they are unstable and rapidly turned over or of very low intracellular abundance.

We only identified seven additional metabolites experimentally which did not form part of the metabolic network (Table S3, ESI†), indicating comprehensive *in silico* reconstruction of the *M. pneumoniae* metabolism.[22] Those metabolites are: fumarate, succinate, ethanol, cytosine, cholesterol, ADP-glucose and *trans*-4-hydroxyproline. Their identification can largely be accounted for by their presence in the rich growth medium of *M. pneumoniae* and subsequent cellular import. More specifically, fumarate and succinate were detected in trace amounts by ^{1}H-NMR based metabolite detection. Succinate but not fumarate was also identified as a minor component of the growth medium. To exclude that either compound is a bona fide metabolic intermediate or product in *M. pneumoniae*, we compared our model organism to the related bacterium *Bacillus subtilis* by using protein BLAST analysis with the goal to identify enzyme orthologs in Mycoplasma involved in fumarate and succinate metabolism. We could assign functions related to fumarate

 This journal is © The Royal Society of Chemistry 2013

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

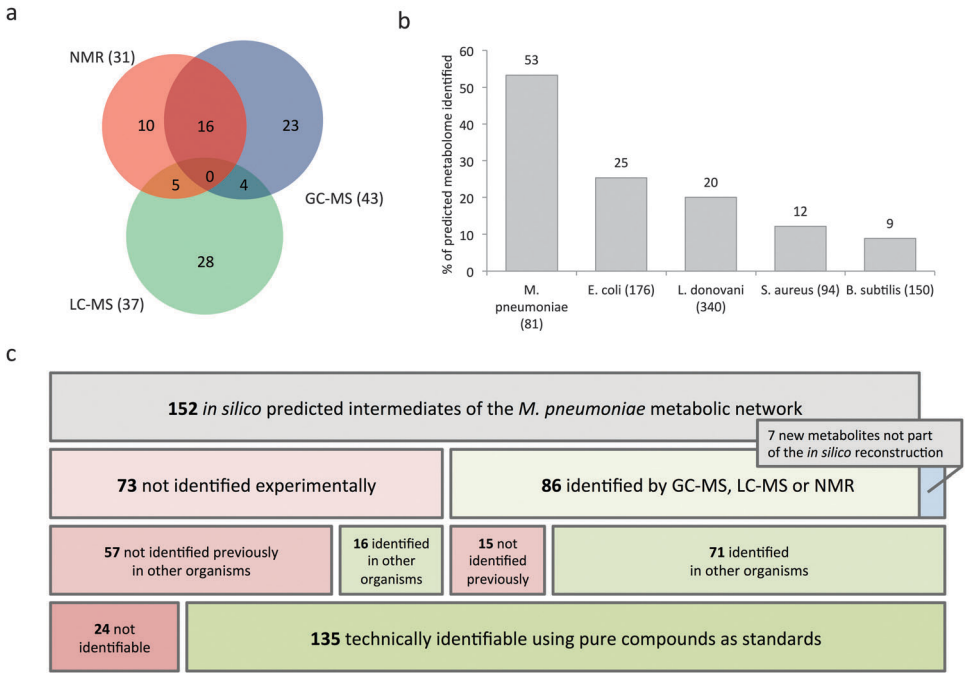

Fig. 1 The *M. pneumoniae* metabolome. (a) Specificity and overlap of different methodologies used to detect *M. pneumoniae* metabolites. (b) Experimentally confirmed metabolites compared to *in silico* predictions for selected organisms. Numbers of identified compounds are given in brackets after organism names. (c) Graphical representation of predicted, identified and technically identifiable metabolites in *M. pneumoniae*.

and succinate metabolism to two *M. pneumoniae* proteins (MPN576 is orthologous to the argininosuccinate synthase from *B. subtilis*, $E = 0.004$; MPN557 is orthologous to a protein of the succinate dehydrogenase of *B. subtilis*, $E = 0.03$). However, possible pathways allowing the synthesis of either metabolite remained incomplete.

Ethanol is a common metabolic end-product of fermentative metabolism. *M. pneumoniae* can in principle synthesize ethanol from pyruvate *via* acetaldehyde and by using a putative alcohol dehydrogenase (MPN564[29]). However, neither the metabolic intermediate acetaldehyde (this study) nor the enzymes catalyzing ethanol production (acetaldehyde dehydrogenase, alcohol dehydrogenase) could be experimentally detected by mass spectrometry.[19] Additionally, we could not measure ethanol secretion of *M. pneumoniae* cells during four-day growth in batch culture using a commercially available enzymatic assay (Fig. S2, ESI†), even though ethanol was detected in trace amounts in the growth medium using NMR.

Post-translational modification of proline residues in eukaryotic cells leads to the occurrence of *trans*-4-hydroxyproline in proteins. Bacteria, including *M. pneumoniae*, do not contain the required enzymes to modify proline accordingly. However, it was shown that *E. coli* can import hydroxyproline under suitable culture conditions.[30] We confirmed the presence of *trans*-4-hydroxyproline in the growth medium by GC-MS and NMR. However, only seven of 3564 peptides identified by mass spectrometry in a representative sample contained thus modified proline residues.

In contrast to most other bacteria, Mycoplasma species incorporate cholesterol into their cell membranes.[31] Since *M. pneumoniae* cannot synthesize sterols (Fig. S1, ESI† (ref. 22)) cholesterol is directly imported from the growth medium. We observed a sharp increase in intracellular cholesterol levels at later stages of batch culture growth, concomitant with a decrease of free cholesterol in the growth medium (Fig. S3, ESI†). We speculate that higher intracellular cholesterol levels influence cell membrane fluidity at later growth stages, possibly playing a role in intracellular pH homeostasis. However, since the exact lipid composition of the *M. pneumoniae* membrane has been shown to vary with the fatty acid composition of the medium,[32,33] cholesterol is not explicitly considered in the metabolic map.[22]

Cytosine and ADP-glucose – although detected experimentally – were found dispensable for the metabolic network. Experiments in defined growth medium showed that cytidine rather than cytosine can be directly imported by *M. pneumoniae*,[22] therefore cytosine is not part of the predicted list of metabolites. ADP-glucose was found functionally redundant with UDP-glucose, which forms part of the *in silico* network and was identified experimentally. Thus, for technical reasons ADP-glucose is not part of the list of predicted metabolites. Interestingly, common metabolites identified in other bacteria, such as the compatible solutes trehalose, betaine, taurine or ectoine, were either only present in trace amounts or not identified in this study, suggesting that *M. pneumoniae* must rely on other osmolytes, such as proline (1.16 mM cellular concentration), glycerol or glucose.

To obtain a topographic view of the *M. pneumoniae* metabolome, we projected the sets of predicted and experimentally identified metabolites onto a map of metabolic pathways (Fig. 2 and Fig. S1, ESI†). Experimentally confirmed metabolites

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

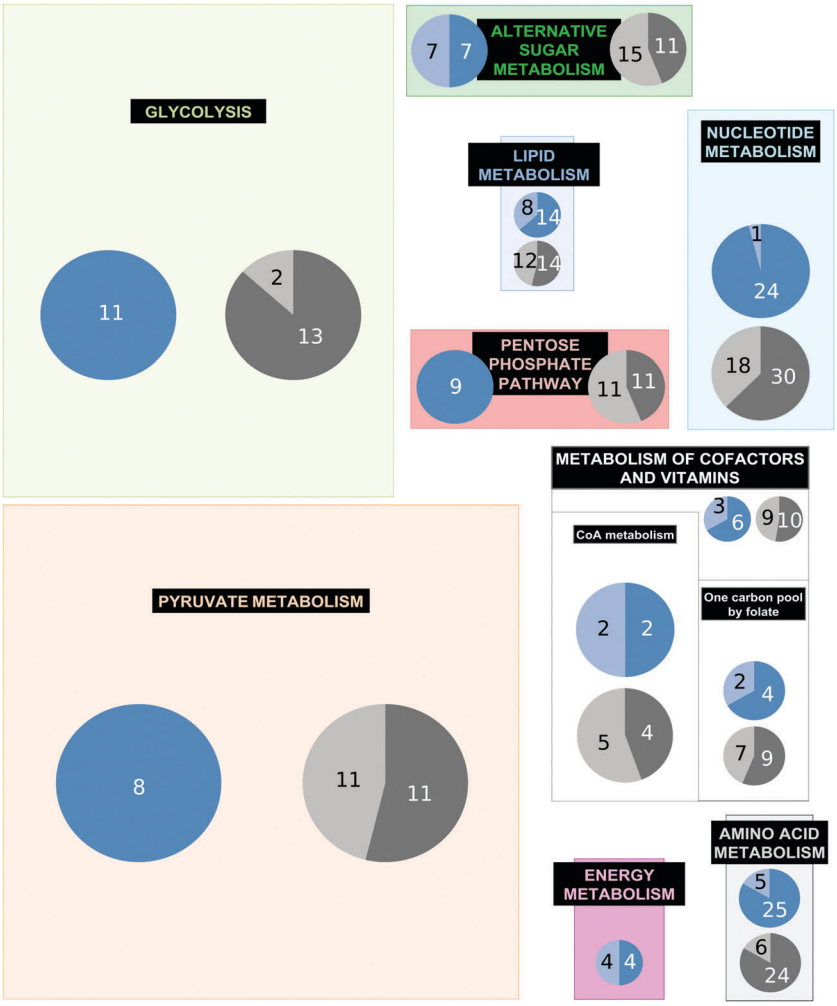

Fig. 2 Pathway-based integration of metabolic fluxes and protein abundances. Pathway boxes are sized according to the average protein abundance for the respective pathway. Blue pie plots detected (dark blue) *vs.* not detected (light blue) enzymes and grey pie plots display detected (dark grey) *vs.* not detected (light grey) metabolites of the respective pathway. Numbers represent absolute numbers of detected proteins and metabolites in each pathway, respectively.

mapped to all 11 predicted metabolic modules, suggesting at least basic activity of most routes during batch culture growth in rich medium. The low number of metabolic modules is indicative of the limited metabolic capacities of *M. pneumoniae*[22] and is a direct consequence of its genome size reduction. High experimental coverage of predicted metabolites was achieved for the central carbon metabolism and for intermediates of nucleotide and amino acid metabolism. Metabolites of the pentose phosphate pathway and of alternative carbon uptake routes, for example fructose-1-phosphate, mannitol-6-phosphate, ascorbate-6-phopshate, were experimentally observed less frequently (Fig. 2 and Fig. S1, ESI†). This finding is in agreement with the low expression levels of the corresponding importers and metabolizing enzymes (Fig. 2).

Summing up, we achieved unprecedented experimental coverage of the *in silico* predicted *M. pneumoniae* metabolome using three complimentary methods for metabolite detection. Integrating our data with the metabolic map allowed validating our metabolic reconstruction.[22] Identified metabolites not forming part of the metabolic map can be explained by the specific growth conditions of the *M. pneumoniae* laboratory strain.

To further characterize the metabolism dedicated to cellular information storage and processing, we quantified the building blocks of RNA, DNA and proteins and integrated their measured intracellular abundances with bound quantities, as well with their availability from the growth medium.

Amino acids

Metabolic map reconstruction,[22] genetic evidence[18] and experimental results on protein abundances[19] suggested that *M. pneumoniae* lacks almost all anabolic routes of metabolism, including those for amino acid synthesis. Hence, to a large degree they need to be imported from the surrounding medium for cell growth. We measured cytoplasmic free amino acids by GC-MS and connected thus determined quantities to the amino acid availability from the surrounding growth medium and to the quantitative amino acid composition of the *M. pneumoniae* proteome (Fig. 3a).

Intracellular amino acid concentrations range from 3.3 mM for glutamic acid to 0.1 mM for cysteine. Glutamic acid

 This journal is © The Royal Society of Chemistry 2013

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

accounts for 21.1% of all free intracellular amino acids (Fig. S4, ESI†). Generally, all amino acids were significantly enriched in the cytosol (on average 14 times) compared to the surrounding growth medium (Fig. S4, ESI†), suggesting active import mechanisms. Artificially increasing the concentrations of individual amino acids in the growth medium resulted in a concomitant increase of their intracellular amounts (Fig. 3b), thus indicating that extracellular amino acid pools directly influence intracellular amino acid concentrations. However, the relative abundance of cytosolic amino acids only partly reflects the available amino acid pools in the growth medium (Pearson correlation coefficient of relative amino acid abundances r_p = 0.16) or the quantitative amino acid distribution in the expressed *M. pneumoniae* proteome (r_p = 0.33)[19] (Fig. 3a). This suggests that *M. pneumoniae* imports different amino acids with different efficiencies.

To account for a potential outflux of methionine and arginine to homocysteine and ornithine, respectively (Fig. S1, ESI†), we performed label-chase experiments. We could not detect labelled downstream compounds for either methionine metabolism or arginine fermentation.

Interestingly, artificially increasing extracellular asparagine and glutamine levels additionally resulted in an intracellular increase of aspartate and glutamate, respectively. This *trans*-effect was not observed inversely, when artificially increasing aspartate and glutamate levels (Fig. 3b). We attribute this finding to the presence of an essential cellular glutamyl-t-RNA amidotransferase, encoded by the genes gatC,A,B (MPN236–MPN238). This heterotrimeric enzyme complex catalyses the conversion of misacetylated Glu-tRNA^{gln} to Gln-tRNA^{gln}, thereby compensating for the lack of Gln-tRNA-synthetase in *M. pneumoniae* (and other gram positive bacteria).[34] Glutamine or asparagine serve as substrates for the ATP consuming transamidation reaction, resulting in glutamate or aspartate release.

More than 25 million individual amino acid molecules are used to make up the average cellular proteome (Table S4, ESI†). The cellular pool of free amino acids contains ~467 000 molecules. Hence, for a single cell division (implying biomass duplication), in twenty hours during exponential growth, 354 individual amino acid molecules must be imported per second and cell and the intracellular amino acid pool is on average turned over

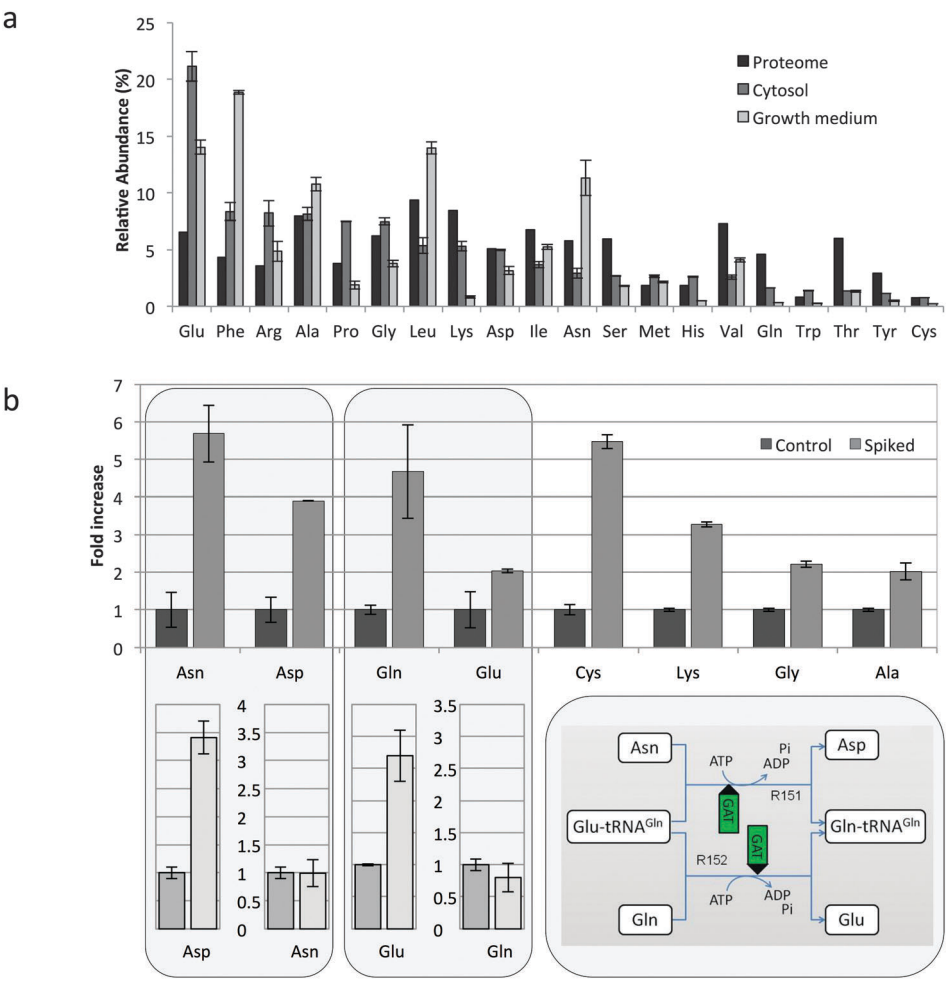

Fig. 3 Amino acid homeostasis in *M. pneumoniae*. (a) Relative amino acid abundance in the cytosol and the growth medium, as well as their quantitative occurrence in the proteome. (b) Extracellular increase of amino acids results in corresponding increase of intracellular pools. The *trans*-effect of extracellular increase of asparagine and glutamine shows cytosolic amidotransferase activity.

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

every 24 minutes (Table S4, ESI†). Integrating the measured cytosolic pools of individual amino acids with their respective contribution to the proteome and the doubling time during exponential growth under laboratory conditions revealed that the intracellular threonine-pool is turned over fastest in 5 minutes, while it takes around 71 minutes to turn over the large intracellular pool of glutamic acid (98 700 molecules). The available pools of free amino acids in the growth medium exceed the required amounts for biomass production during a typical 4 day batch culture growth experiment (150 cm^2 flask surface, ~3 mg protein synthesized, 100 ml growth medium), indicating that the supply of protein building blocks is not growth limiting (Fig. S4, ESI†).

Nucleobases

Nucleobases, adenine, cytosine, guanine, thymine and uracil, are versatile cellular building blocks. They are used to a different extent in DNA and in different types of RNA. They form the cellular ''energy currency'' in the form of nucleoside triphosphates (*e.g.* ATP) and they are part of other coenzymes, such as NAD or coenzyme A. In many microorganisms, they can also be used as cellular nitrogen sources.[35] We measured nucleobases and nucleosides in the *M. pneumoniae* cytosol as well as in the growth medium from samples taken at regular intervals during a four-day growth course (Fig. S5, ESI†). Methodological constraints in sample derivatization prior to GC-MS analysis led to substantial chemical conversion of nucleosides (and their phosphorylated derivatives) to their cognate bases for most of the nucleosides. The rate of conversion ranged from ~3% for adenosine to ~50% for thymine (Fig. S6, ESI†). Therefore, we reason that cumulative amounts of measured cytosolic nucleobases and nucleosides (NUBS) quantitatively represent the cellular inventory of free nucleobases and nucleosides, as well as intracellular pools of (labile) phosphorylated nucleosides and nucleobase containing coenzymes (Fig. 4).

We observed a continuous depletion of NUBS from the growth medium during the four-day growth course. However, final concentrations determined from medium samples after 96 hours of growth suggest that NUBS are not growth limiting (Fig. S5, ESI†). NUBS are on average 700 times more concentrated in the cytosol than in the surrounding growth medium. These findings indicate direct and active uptake of these building blocks by *M. pneumoniae*. Surprisingly, we found no *M. pneumoniae* protein, with significant homology to identified nucleobase transporters from *E. coli* and *B. subtilis* (Table S5, ESI†).[35]

Detected intracellular abundance changes along the four day growth course reflected the observed extracellular changes for all NUBS but for adenine/adenosine, which was determined to have an intracellular concentration of 5 mM at all measured time points, whereas it was significantly depleted from the medium during batch culture growth. On average, measured cellular concentrations for *M. pneumoniae* NUBS are 3 times lower than values determined for glucose fed *E. coli* during exponential growth.[36]

To gain an understanding of cellular nucleobase homeostasis and turnover, we quantified bound nucleobases in DNA, ribosomal RNA, mRNA and tRNA (Fig. 4). For this purpose, we used published data on the *M. pneumoniae* genome sequence, its ribosome content as well as mRNA and tRNA abundances.[19,22] The bacterial DNA and its rRNA are the dominant cellular nucleobase sinks and for all nucleobases (except U), the amount of bound material exceeded the freely available NUBS. The low amount of nucleobases bound in mRNA reflects the determined low cellular abundance of mRNA (on average 9.1 mRNA molecules per cell at any given time).[19] Generally, the ratio of

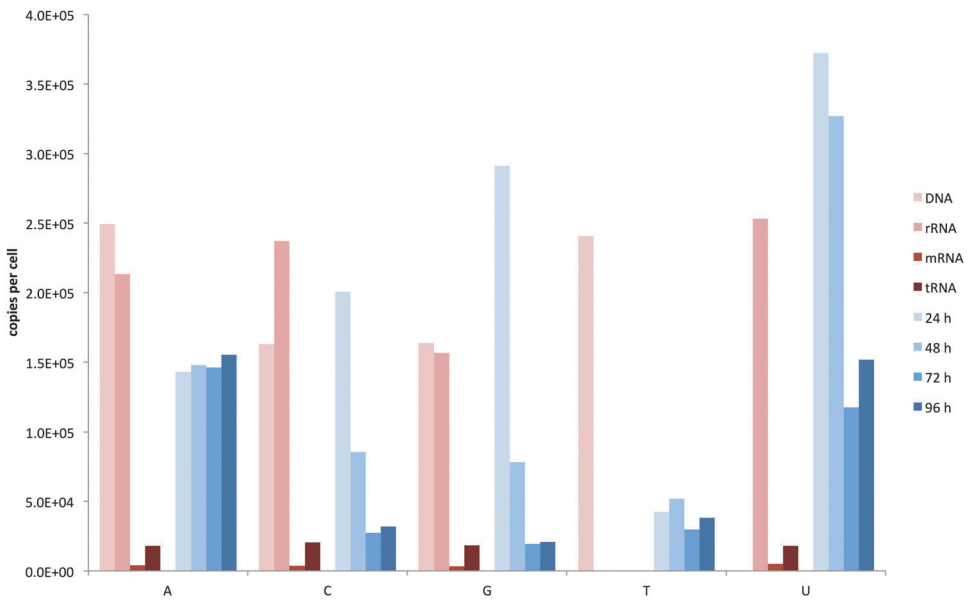

Fig. 4 NUBS (nucleobase + nucleoside) homeostasis. Cellular abundance of bound nucleobases (red bars) and free NUBS in the *M. pneumoniae* cytoplasm.

 This journal is © The Royal Society of Chemistry 2013

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

free to bound NUBS is higher compared to the corresponding values for free and bound amino acids in proteins, resulting in longer turnover times of intracellular nucleobase pools. These larger intracellular pool sizes reflect the high functional diversity of nucleobases, since they serve as building blocks for different classes of molecules, such as nucleotides, nucleoside triphosphates, and co-enzymes.

Integration of metabolites and protein abundances

In order to extract quantitative information on *in vivo* metabolic pathway usage, we integrated information on metabolic enzyme abundance[19] with metabolite data from this study (Fig. 2 and Fig. S1, ESI†). 78% of all proteins forming part of the metabolic network were identified and quantified (Fig. S1, ESI†). Interestingly, proteins not identified in a previously published full proteome analysis of *M. pneumoniae*[19] clustered to the pathways involved in phospholipid synthesis and the import of alternative carbon sources. These routes are dispensable for *M. pneumoniae* growth in rich medium, since phospholipids and glucose are provided directly from the growth medium. This result confirms findings in *Saccharomyces cerevisiae*, where it has been shown that under permissive conditions, non-essential proteins are more frequently found in lower abundance.[37] Strikingly, 65% of all protein molecules involved in the cellular metabolism (considering their abundance) are accounted for by only 19 different enzymes involved in glycolysis and pyruvate metabolism, confirming the importance of this central pathway used for energy generation. 20% of all metabolic enzymes are involved in nucleotide and amino acid metabolism, the basic building blocks of DNA, RNA and proteins (Fig. 2).

Interestingly, metabolite coverage of metabolic pathways does not match enzyme coverage, nor does it correlate stringently with enzyme abundance. For example, several intermediates of the highly active and abundant glycolysis and pyruvate metabolism escaped identification (1,3-diphosphoglycerate, 3-phosphoglycerate, 2-phosphoglycerate, acetyl-CoA, acetylphosphate), but we detected almost all intermediates of the amino acid and nucleotide metabolism (Fig. 2 and Fig. S1, ESI†). Only a few metabolites of the pentose phosphate pathway were detected, despite full enzyme coverage, suggesting the low activity of this pathway under laboratory growth conditions. Most intermediates of the folate metabolic pathway also escaped experimental validation. In addition to low pathway usage, resulting in low absolute levels, technical constraints, such as metabolite instability or limitations of the employed analysis methods could prevent their detection.

The finding that there is not a strict correlation between protein abundances and metabolite detection additionally suggests that in some pathways bottlenecks exist, resulting in accumulation of the upstream metabolite, while other enzymatic reaction steps occur at much higher rates, leading to substrate depletion below detection limits. In principle, this observation allows the estimation of *in vivo* kinetic parameters for participating enzymes when protein abundances, metabolite pool sizes and reaction velocities are available. To apply this hypothesis to our data, we determined *in vivo* turnover numbers (k_{cat}) of glycolytic enzymes at different stages during batch culture growth.

In vivo turnover of glycolysis

We measured cellular concentrations of key metabolites of sugar metabolic pathways. Intracellular pools of glycolysis intermediates are low, ranging from 994 molecules per cell (0.03 mM) for glyceraldehyde-3-phosphate (GAP) to 33 400 molecules per cell (1.1 mM) for fructose-1,6-bisphosphate (FBP, Fig. 5b). Based on measurements of extracellular glucose consumption in batch culture growth, we calculated an average glucose consumption rate of 8000–35 000 molecules per cell per second depending on the growth stage (Fig. S2, ESI†). We then integrated the glucose uptake speed ($v_{glucose}$) with measured protein abundances[19] and determined intracellular metabolite levels (Fig. 5a).

Two approximations allowed us to calculate *in vivo* k_{cat} values using available data: (1) since most imported glucose is directly converted into organic acids and only a small fraction is fed into branching pathways, such as the pentose phosphate pathway or the phospholipid pathway,[22] http://www.nature.com/msb/journal/v9/n1/full/msb20136.html (Fig. S2, ESI†), the experimentally measured glucose consumption during batch culture growth is approximately equal to the speeds of individual glycolytic reactions: $v_{glycolysis} = v_1 = v_2 = \cdots$. Reactions following FBP cleavage by fructose bisphosphate aldolase should be twice as fast, since consecutively, two C_3-compounds have to be processed instead of one C_6-sugar. (2) Where detectable, the intracellular pools of glycolytic intermediates remained largely unchanged during batch culture growth and were considerably larger than enzyme pools catalyzing the corresponding reactions. Therefore, the specific reaction speeds $v_{glycolysis}$ approximate V_{max}.

Hence, adapting the formula

$$V_{max} = k_{cat}[E]$$

to

$$v_{glycolysis(t)} = k_{cat}[E]_{(t)}$$

delivered *in vivo* turnover numbers for phosphofructokinase, aldolase, phosphoglucose isomerase and pyruvate kinase at different timepoints t during batch culture growth. Thus determined k_{cat} values varied during batch culture growth for all enzymes, usually peaking at 72 h of growth (Fig. 5c). In general, the *in vivo* k_{cat} values represent lower boundaries for enzyme reaction rates under ideal conditions. Phosphofructokinase had the highest k_{cat} after 48 hours of growth (293 s^{-1}). Turnover numbers for several glycolytic enzymes could not be determined, because either available enzyme abundances exceeded substrate pools (*e.g.* for GAPDH) or substrate metabolite amounts were unknown.

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

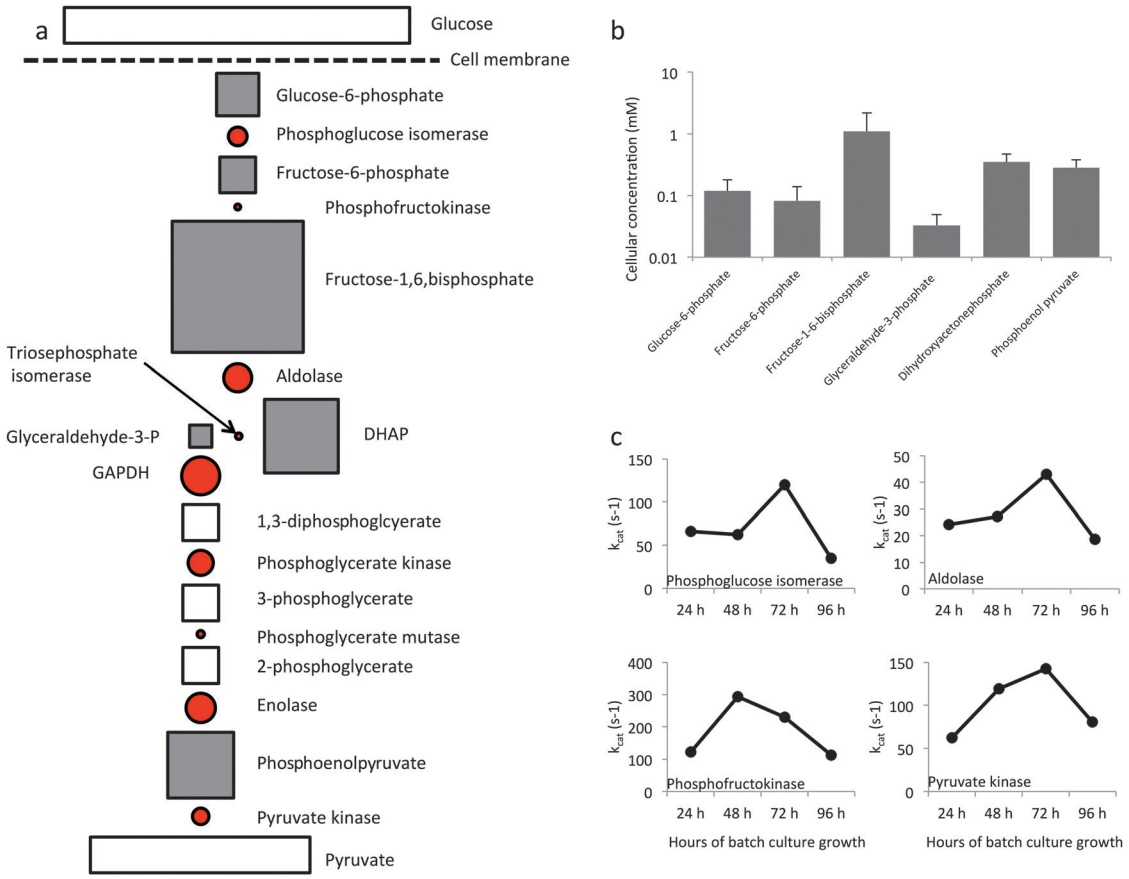

Fig. 5 Quantitative glycolysis in *M. pneumoniae*. (a) Areas of circles (enzymes) and squares (metabolic intermediates) represent intracellular molar amounts at 96 h of growth. For some metabolic intermediates, the cellular concentration could not be determined or the metabolite was not detected at all (open squares). (b) Measured cellular abundances of glycolytic intermediates. (c) *In vivo* k_{cat} values for selected enzymes with large substrate pools along batch culture growth.

Besides metabolite pool sizes and enzyme concentration, additional factors, such as allosteric- and direct regulation of enzyme activity (*e.g.* negative feedback inhibition), metabolite partitioning to branching pathways, reaction reversibilities and stochastic effects can influence the *in vivo* reaction rates. To obtain a more complete understanding of *in vivo* glycolysis and underlying regulatory mechanisms, dynamic modeling approaches are necessary, based on quantitative data, ideally on the sub-second scale and including time series following perturbations, such as enzyme inhibition or overexpression. Currently, technical limitations in sample preparation prevent such approaches for *M. pneumoniae* or any other species for that matter.

Discussion

We present here a large-scale study of the *M. pneumoniae* metabolome using a cross-platform approach. We identified more than 50% of predicted metabolic intermediates. Using additional, targeted tandem mass spectrometry approaches should allow a further increase in metabolome coverage. However, as a consequence of a streamlined metabolism, many intermediates of metabolic side routes in *M. pneumoniae*, such as the pentose phosphate pathway, are present below the detection limit and escape identification even using targeted approaches.

We absolutely quantified a set of key metabolites, such as glycolysis compounds, amino acids and nucleobases. Mapping detected and quantified metabolites onto the *in silico* metabolic network and integrating them with protein abundances resulted in a qualitative picture of the *M. pneumoniae* metabolic pathway activity under laboratory growth conditions and allowed the determination of *in vivo* turnover numbers of glycolytic enzymes. Integrating cytoplasmic metabolite pool sizes with available resources from the growth medium and their corresponding bound forms allowed conclusions on the homeostasis of central cellular building blocks and resulted in additional insights into cellular physiology and regulatory mechanisms.

Amino acid homeostasis of *M. pneumoniae* can serve as an illustrative example here. Integrating free amino acid pools from the growth medium and the cytosol with the quantitative occurrence of amino acids in proteins (Fig. 3a) suggested that amino acid import reactions, albeit active cellular processes (Fig. S4, ESI†), are not in tune with amino acid supply and demand under laboratory growth conditions. Possibly, intracellular amino acid pools reflect the proteome composition to a higher degree in *M. pneumoniae* cells isolated *in situ,* from lung

 This journal is © The Royal Society of Chemistry 2013

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

epithelial cells of patients. This notion suggests the existence of two different cellular modes for regulatory responses. The robust, ''hard-wired'' amino acid import regulation in *M. pneumoniae* only allows adaptation over many generations and probably involves mutational changes in upstream regions of genes encoding amino acid importers. In contrast, ''soft-wired'' regulatory mechanisms based on transcription factors, such as the heat shock response, result in immediate adaptation to environmental changes on both mRNA and protein level.[19] Furthermore, one can speculate that metabolic side routes, such as the uptake of alternative carbon and energy sources and the interconnectivity of metabolic pathways, increase in importance *in situ*, where no direct supply of high amounts of glucose is available.

Comparing measured intracellular metabolite concentrations from this study with corresponding absolute metabolite numbers from glucose-fed, exponentially growing *E. coli* cells[36] indicated that individual metabolite concentrations are organism-specific. However, relative metabolite abundances correlate moderately. For example, intracellular amino acids in *E. coli* were measured at a total concentration of 113 mM.[36] In *M. pneumoniae*, the cellular concentration of amino acids is only 15.5 mM. However, glutamate was determined to be the most abundant free amino acid in our study as well as in *E. coli* and log 2 transformed amino acid abundances correlate moderately (r_p = 0.51) (Fig. S7, ESI†). Comparing intracellular concentrations of glycolytic intermediates between the two organisms yields similar results. Total concentrations of compounds in glycolysis differed between these two organisms: 25.9 mM for *E. coli* and 2.0 mM for *M. pneumoniae*. However, FBP is in both *E. coli* (15 mM) and in *M. pneumoniae* (1.1 mM) the most abundant glycolytic intermediate. The log 2 transformed values for glycolytic compounds where abundances were determined for both organisms also correlate moderately (r_p = 0.42). A similar trend, although only based on three metabolites (G6P, F6P, FBP), can be observed when comparing intermediates glycolysis between *Lactococcus lactis*[38] and *M. pneumoniae* (Fig. S7, ESI†).

The discovery that relative abundances of metabolites of a certain class or pathway correlate moderately between organisms is surprising. Different growth environments and the use of different metabolic routes all together (*E. coli* can synthesize amino acids, and following glycolysis it uses respiration for ATP generation under oxic conditions) apparently influence absolute metabolite abundances but not their relative ratios. Intriguingly, a corresponding effect has been observed on the protein level, where relative abundances of proteins of certain functional classes in organisms with very different growth environments correlate well, while their absolute abundances can be different.[19] Both these findings suggest a certain rigidity of metabolic pathway architecture, implying that they are regulated as functional units rather than on the level of individual enzymatic reactions. This notion offers a model for explaining recent results showing how metabolic flux distributions evolved in bacteria.[39] Besides maximizing for biomass and energy yield under regular growth conditions, bacteria evolved their metabolic networks to minimize the necessary flux changes for adaptive responses (http://www.nature.com/msb/journal/v9/n1/full/msb20136.html). Regulating metabolic pathways as functional units, implying a more coarse-grained regulatory space, greatly simplifies metabolic flux adjustments.

Methods

Sample preparation

M. pneumoniae cells were grown in batch culture in suitable culture flasks in Hayflick rich growth medium. Generally, cells were grown as pre-culture for 96 hours, harvested and diluted into fresh growth medium and seeded into new culture flasks for experiments. Cells were grown for different time intervals, ranging from 24 to 96 hours. At indicated times, the growth medium was discarded and the cells were washed twice with ice-cold PBS containing 0.05% glucose. After complete removal of the wash buffer, the culture flask was placed on a bed of dry ice and −80 °C methanol was rapidly added for both quenching metabolism and lysing the cells. After cell scraping and collecting the samples, cell debris was spun down and the supernatant containing cellular metabolites was transferred to pre-cooled glass tubes containing internal standard as indicated below. Samples were immediately frozen in liquid nitrogen and lyophilized to dryness for 24 h–72 h.

Protein concentration and enzyme assays

Protein content was determined using the commercially available BCA kit (Thermo), essentially following the manufacturer's recommendations and as described in ref. 22. The determination of extracellular glucose, lactic acid, acetic acid and ethanol was carried out using commercially available kits (BioVision #K606 and #K607, Megazyme K-ACETRM, K-ETOH) as described in ref. 22.

GC-MS analysis

Different groups of compounds (glycolysis products, free amino acids, free bases) were targeted specifically using tailored protocols. Depending on the case, growth medium, total cell content, cell pellet, or cytoplasm was analyzed as described in each protocol. For compounds detected by GC/MS, the limit of quantification was estimated as the concentration of a metabolite resulting in a peak with a signal-to-noise ratio (S/N) of 10. If the lowest calibration level had a higher S/N ratio, then the concentration of the quantification limit was estimated as the lowest calibration level.

The LOQ for amino acids ranged from 700 to 2300 pg on the column, for the bases 10 pg on the column, for the nucleosides (adenosine and guanosine) 40 pg on the column, for cholesterol 200 pg on the column and for the glycolysis intermediates between 80 and 900 pg on the column.

Glycolysis compounds

For the analysis of glycolysis intermediates (only cell cytosol extract) samples, added with the internal standard (5 μL of methyltestosterone 10 ng μL^{-1}), were lyophilized and subsequently dried

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

in a vacuum oven (500 mbar, 50 °C) in the presence of diphosphorus pentoxide for at least 4 h. Then, aldehyde and ketone functional groups were converted into methyl oximes with 75 μL of methoxyamine hydrochloride in pyridine (2 g L^{-1}) at 40 °C for 90 min with intermediate mixing. Subsequently, hydroxyl groups were converted into trimethylsilyl groups with 100 μL MSTFA (*N*-methyl-*N*-trifluoroacetamide) at 40 °C for 50 min. Samples were transferred to glass inserts, spun for 5 min at 5000 rpm, and the supernatant transferred to a new vial for GC-MS analysis.[40] 4 μL of glycolytic products were analyzed (split ratio 1 : 10) using a HP-Ultra1 cross-linked methyl-silicone column, 16.5 m × 0.2 mm i.d., film thickness 0.11 μm (J&W Scientific, Folsom, CA, USA) in an Agilent 6890N gas chromatograph coupled to an Agilent 5973 mass selective detector. Helium was used as carrier gas at a constant pressure of 5 psi. The GC temperature is ramped as follows: initial 70 °C, held for 1 min, increased to 280 °C at 6 °C min^{-1}, and held for 1 min at 280 °C. For optimal sensitivity the acquisition, performed in SIM mode, was split into four time-segments with three characteristic ions per compound (plus the corresponding ^{13}C isotopes using 15 ms dwell times) in each. Four different time segments ranged from 6–16 min (13 ions: *m*/*z* 103.1, 117.1, 211.1, 232.2, **236.2**, 299.2, 315.2, 369.2, **371.3**, 384.3, **386.3**, 445.4, **448.4** for phospho-groups, PEP, GAP, DHAP, and G3P), 16–19.5 min (15 ions: *m*/*z* 103.1, **104.1**, 147.1, 160.1, **162.1**, 191.1, 205.1, **207.1**, 217.1, **220.1**, 307.2, **310.2**, 319.2, **323.2**, 409.3, fructose, I.S., glucose, and G1P), 19.5–28 min (11 ions: *m*/*z* 103.1, **104.1**, 299.2, 315.2, 357.2, **359.2**, 387.2, 459.3, **462.3**, 471.3, **475.3**: F6P, G6P, and R5P), and 28–32 min (9 ions: *m*/*z* 103.1, **104.1**, 299.2, 315.2, 357.2, **359.2**, 387.2, 459.3, **462.3**: FBP) (bolded are the ions corresponding to the 13C labelled compounds). Identification of the sample constituents was based on the theoretical fragmentation pattern expected for each compound. Quantification was performed with respect to the corresponding calibration curve standardised against the internal standard (methyltestosterone-MO-TMS). Samples and the calibration curves were analyzed at least in triplicate.

Amino acids

For the analysis of free amino acids cell extracts were employed. Prior to lyophilization nor-leucine was added as internal standard (50 μL, 25 nM). Lyophilised material was resuspended in 100 μL MTBSTFA (*N*-(*tert*-butyldimethylsilyl)-*N*-methyltrifluoroacetamide) and 100 μL acetonitrile and samples were incubated for 4 hours at 100 °C with intermediate mixing.[41] Then, samples were allowed to cool to room temperature and the content transferred to a micro insert. Samples were spun for 5 min at 5000 rpm, and the supernatant transferred to a new vial for GC-MS analysis. Typically, two μL of the derivatized amino acid products were analyzed (split ratio 10 : 1) using a Phenomenex Zebron ZB-5 cross-linked 5% phenyl polymethyl siloxane capillary column (15 m × 0.25 mm i.d., 0.25 μm film thickness) in an Agilent 6890N gas chromatograph (temperature program: initial temperature 100 °C, increased to 310 °C at 25 °C min^{-1}, and held at 310 °C for 3 min) coupled to an Agilent 5973 mass selective detector operated in scan mode (scan range from 50 to 650 Da). All amino acids could be analyzed directly from the scan-mode analysis and most were identified through the characteristic $[M]^{\bullet+}$, $[M\text{-}57]^{\bullet+}$, $[M\text{-}85]^{\bullet+}$, and $[M\text{-}159]^{\bullet+}$ ions. Occasionally a different fragment ion was employed for the quantification. In the case of arginine the reaction with MTBSTFA proceeds with cleavage of the guanidino group, producing a mixture of the *tert*-butyldimethylsilyl derivatives of carbodiimide and ornithine.[42] The latter is detected at *m*/*z* 474 (also the subsequent fragments at *m*/*z* 417 and *m*/*z* 286). Table S6 (ESI†) summarises the retention times together with the characteristic ions used for identification.

For the calibration curves a standard solution was prepared containing the amino acids at the concentrations above and below the concentrations found in the samples. The final curve contained six calibration points.

Nucleobases

Bases, nucleosides and cholesterol were analyzed from cell lysates or growth medium. In the case of growth medium 300 μL aliquots were lyophilized and subjected to analysis. For cell lysates samples were split in four 250 μL aliquots and analyzed as analytical quadruplicates of biological duplicates. A standard calibration curve containing all bases, nucleosides, and cholesterol ranging from 50 ng μL^{-1} to 1 μg μL^{-1} (6 points) was employed for quantitative purposes. Prior to lyophilization to each sample the internal standard was added (50 μL of a 0.5 mM norleucine solution). Following lyophilization samples were dried in a vacuum oven (500 mbar, 50 °C) in the presence of diphosphorus pentoxide for at least 4 h. Then, 100 μL of MSTFA (*N*-methyl-*N*-trimethylsilyl-trifluoroacetamide) were added to each sample (in the case of growth medium 200 μL) and samples were incubated for 30 min with intermediate mixing. Samples were allowed to cool to room temperature and the content transferred to a micro insert. Samples were spun for 5 min at 5000 rpm, and the supernatant transferred to a new vial for GC-MS analysis. Typically, one μL of derivatized products were analyzed (split ratio 1 : 150) using a Phenomenex Zebron ZB-5 cross-linked 5% phenyl polymethylsiloxane capillary column (15 m × 0.25 mm i.d., 0.25 μm film thickness) in an Agilent 6890N gas chromatograph (temperature program: initial temperature 70 °C, increased to 310 °C at 25 °C min^{-1} and held for 3 min at the final temperature) coupled to an Agilent 5973 mass selective detector operated in scan mode (scan range from *m*/*z* 50 to *m*/*z* 700). Bases and nucleosides were identified based on the molecular ion and the typical fragment after the loss of a methyl group (see Table S7, ESI†). Six point calibration curves were used for bases and nucleosides. Cholesterol was added to the standard base or nucleoside solutions at 10 ng μL^{-1}.

Transformation of nucleosides to bases was observed through the analysis of pure nucleoside standards (Fig. S6, ESI†). Approximately 25.4% of cytidine is converted to cytosine, 3.1% of adenosine to adenine, 6.2% of guanosine to guanine, 49.9% of thymidine to thymine, and 12.6% of uridine to uracil.

NMR measurements

Dried extracts were re-dissolved in 650 μL phosphate buffered H_2O/D_2O (10/90 v/v) containing 0.5 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid, d_6 sodium salt and 15 mM NaN_3, centrifuged

 This journal is © The Royal Society of Chemistry 2013

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

(5 min, 14 000*g*) and 600 μL supernatant transferred into NMR tubes. Spectra were acquired using an Avance 800 MHz NMR Spectrometer with triple resonance CryoProbe (Bruker Biospin, Coventry, UK) following a procedure described in ref. 43 and standard pulse sequences for 1D ^{1}H-NOESY, 2D ^{1}H–^{1}H COSY, ^{1}H–^{13}C HSQC experiments. Metabolite identification was done on accurate mass match and retention time match (standards were run under same chromatographic conditions) as well as comparing spectra of standard compounds and spectra available from online repositories (HMBD[44] and BMRB, http://bmrb.cerm.unifi.it).

LC-MS measurements

Dried extracts were re-dissolved in 400 μl 1 : 1 water–methanol (LC-MS grade), centrifuged and supernatants were transferred to HPLC glass-vials. Samples were analyzed by UPLC-MS with a HILIC mode and a reversed phase mode separation[45,46]; analysis was performed using an Acquity UPLC system (Waters Ltd. Elstree, U.K.) coupled to a Q-TOF Premier mass spectrometer (Waters MS Technologies, Ltd., Manchester, U.K.) operated in positive and negative ionisation mode (capillary voltage was 2.4 kV, sample cone was 35 V, desolvation temperature 350 °C, source temperature 120 °C, and desolvation gas flow 900 L h^{-1}). Data were collected in centroid mode with a scan range of 50–1000 *m/z*. In addition, samples were subjected to an ion-pairing mode HPLC-MS method[47] for the analysis of very polar metabolites like triphosphate nucleotides. The ion-pairing HPLC was coupled to a micrOTOF (Bruker, Rheinstetten) operating in negative ionization, full scan mode, and the mass range was set to 100–2000 Da at a resolution of 10 000 (at *m/z* 1033). Source parameters were as follows: drying gas (N2) flow rate, 8.0 l min^{-1}; drying gas temperature, 180 °C; nebulizer, 1.6 bar (160 kPa); capillary voltage, 4.5 kV; skimmer 1 voltage, −50 V; skimmer 2 voltage, −24 V. Accurate masses were extracted, and integration of designated peaks was performed by using QuantAnalysis (Bruker, Rheinstetten, Germany). Limit of detection for several tested compounds was between 20–200 fmol per injection. All the LC-MS data were evaluated for predicted *M. pneumoniae* metabolites,[22] including possible ions for common adducts in ESI mass spectrometry (*e.g.* $[M + H]^+$, $[M + NH_4]^+$, $[M + Na]^+$, or $[M + K]^+$ and in negative ion mode $[M - H]^-$).[48] As LC-MS profiling data often contain a huge amount of uninformative "noise"[49] we applied a filtering to our peak list, just peaks above a specified abundance threshold and with ions not on the predicted metabolite list were analyzed for further identification purposes by database searches in Metlin, massbank.jp.[50,51]

Label-chase experiments

To detect labeled and unlabeled metabolites involved in arginine fermentation and methionine metabolism, tandem LC-MS/MS analyses were carried out using a triple quadrupole (Quattro Premier XE) mass spectrometer provided with an orthogonal Z-spray–electrospray interface (Waters Associates, Milford, MA) coupled to an ultraperformance liquid chromatographic system, Acquity (Waters Associates), for the chromatographic separation. Drying gas, nebulizing gas, cone gas, and desolvation gas were nitrogen. Separations were performed on a WATERS ACQUITY UPLC-R BEH HILIC column (1.7 μm, 2.1 × 50 mm) using 10 mM CH_3COONH_4 (pH 4; A) 5% (A) in acetonitrile (B) as a solvent system and a gradient from 99% (B) to 30% (B) over 4 minutes at a flow rate of 0.3 mL min^{-1}. The total run time was 7.5 minutes. Data were acquired in positive MRM mode.

Glucose consumption calculations

Using the difference in external glucose in molecule numbers (ΔGlc) and the average number of cells (#cells) for each interval between two measured time points in seconds (Int_s) we calculated the glucose uptake rate per cell and second: ΔGlc per #cells per Int_s. The number of cells at each discrete time point has been determined based on the total protein measured and the finding that one *M. pneumoniae* cell contains 10 fg of protein,[22] *i.e.* dividing the total amount of protein in fg by 10 gives the number of cells at each discrete time point measured. To obtain the average cell number for one interval the number of cells measured at the beginning and at the end of the respective interval have been summed up and then divided by 2. The length of each interval in seconds has been calculated resting the time of growth at the beginning of the respective interval from that at the end.

Author contributions

TM designed the study, performed the experiments, analyzed the data, prepared figures and wrote the manuscript. JM processed samples, carried out the GC-MS experiments and analyzed data. JW analyzed data and prepared figures. BP performed label chase and minimal medium experiments. ML carried out the LC-MS and NMR analyses. RGG processed samples, carried out the GC-MS analysis, analyzed data, contributed to the study design and discussed results. LS contributed to the study design and discussed results. All the authors commented on the manuscript.

Acknowledgements

This work was supported by the European Research Council (ERC) advanced grant, the Fundación Marcelino Botin and the Spanish Ministry of Research and Innovation to the ICREA researcher LS and by a laCaixa fellowship for JW. The authors declare no competing financial interest.

References

1 W. B. Dunn, D. I. Broadhurst, H. J. Atherton, R. Goodacre and J. L. Griffin, *Chem. Soc. Rev.*, 2011, **40**, 387–426.
2 A. Cornish-Bowden and M. L. Cárdenas, *Nat. Biotechnol.*, 2000, **18**, 267–268.
3 O. Fiehn, *Plant Mol. Biol.*, 2002, **48**, 155–171.
4 J. M. Buescher, W. Liebermeister, M. Jules, M. Uhr, J. Muntel, E. Botella, B. Hessling, R. J. Kleijn, L. Le Chat, F. Lecointe, U. Mäder, P. Nicolas, S. Piersma, F. Rügheimer, D. Becher, P. Bessieres, E. Bidnenko, E. L. Denham,

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

E. Dervyn, K. M. Devine, G. Doherty, S. Drulhe, L. Felicori, M. J. Fogg, A. Goelzer, A. Hansen, C. R. Harwood, M. Hecker, S. Hubner, C. Hultschig, H. Jarmer, E. Klipp, A. Leduc, P. Lewis, F. Molina, P. Noirot, S. Peres, N. Pigeonneau, S. Pohl, S. Rasmussen, B. Rinn, M. Schaffer, J. Schnidder, B. Schwikowski, J. M. Van Dijl, P. Veiga, S. Walsh, A. J. Wilkinson, J. Stelling, S. Aymerich and U. Sauer, *Science*, 2012, **335**, 1099–1103.

5 J. K. Nicholson and J. C. Lindon, *Nature*, 2008, **455**, 1054–1056.

6 D. J. Creek, A. Jankevics, R. Breitling, D. G. Watson, M. P. Barrett and K. E. V. Burgess, *Anal. Chem.*, 2011, **83**, 8703–8710.

7 R. Goodacre, S. Vaidyanathan, W. B. Dunn, G. G. Harrigan and D. B. Kell, *Trends Biotechnol.*, 2004, **22**, 245–252.

8 F. Geier, E. Want, A. Leroi and J. Bundy, *Anal. Chem.*, 2011, **83**, 3730–3736.

9 L. M. Liberman, R. Sozzani and P. N. Benfey, *Curr. Opin. Plant Biol.*, 2012, **15**, 162–167.

10 A. Scalbert, L. Brennan, O. Fiehn, T. Hankemeier, B. S. Kristal, B. van Ommen, E. Pujos-Guillot, E. Verheij, D. Wishart and S. Wopereis, *Metabolomics*, 2009, **5**, 435–458.

11 W. M. van Gulik, *Curr. Opin. Biotechnol.*, 2010, **21**, 27–34.

12 N. C. Duarte, S. A. Becker, N. Jamshidi, I. Thiele, M. L. Mo, T. D. Vo, R. Srivas and B. Ø. Palsson, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1777–1782.

13 A. M. Feist, M. J. Herrgord, I. Thiele, J. L. Reed, B. Ø. O. Palsson and M. J. Herrgård, *Nat. Rev. Microbiol.*, 2009, **7**, 129–143.

14 T. Dandekar, M. Huynen, J. T. Regula, B. Ueberle, C. U. Zimmermann, M. A. Andrade, T. Doerks, L. Sánchez-Pulido, B. Snel, M. Suyama, Y. P. Yuan, R. Herrmann, P. Bork and L. Sanchez-Pulido, *Nucleic acids Res.*, 2000, **28**, 3278–3288.

15 R. Himmelreich, H. Hilbert, H. Plagens, E. Pirkl, B. C. Li and R. Herrmann, *Nucleic Acids Res.*, 1996, **24**, 4420–4449.

16 R. M. Chanock, W. D. James, H. H. Fox, H. C. Turner, M. A. Mufson and L. Hayflick, *Proc. Soc. Exp. Biol. Med.*, 1962, **110**, 884–889.

17 A. P. H. Danielsson, T. Moritz, H. Mulder and P. Spégel, *Anal. Biochem.*, 2010, **404**, 30–39.

18 M. Güell, V. van Noort, E. Yus, W.-H. Chen, J. Leigh-Bell, K. Michalodimitrakis, T. Yamada, M. Arumugam, T. Doerks, S. Kühner, M. Rode, M. Suyama, S. Schmidt, A.-C. Gavin, P. Bork and L. Serrano, *Science*, 2009, **326**, 1268–1271.

19 T. Maier, A. Schmidt, M. Güell, S. Kühner, A.-C. Gavin, R. Aebersold and L. Serrano, *Mol. Syst. Biol.*, 2011, **7**, 1–12.

20 S. Kühner, V. van Noort, M. J. Betts, A. Leo-Macias, C. Batisse, M. Rode, T. Yamada, T. Maier, S. Bader, P. Beltran-Alvarez, D. Castaño-Diez, W.-H. Chen, D. Devos, M. Güell, T. Norambuena, I. Racke, V. Rybin, A. Schmidt, E. Yus, R. Aebersold, R. Herrmann, B. Böttcher, A. S. Frangakis, R. B. Russell, L. Serrano, P. Bork and A.-C. Gavin, *Science*, 2009, **326**, 1235–1240.

21 V. van Noort, J. Seebacher, S. Bader, S. Mohammed, I. Vonkova, M. J. Betts, S. Kühner, R. Kumar, T. Maier, M. O'Flaherty, V. Rybin, A. Schmeisky, E. Yus, J. Stülke, L. Serrano, R. B. Russell, A. J. R. Heck, P. Bork and A.-C. Gavin, *Mol. Syst. Biol.*, 2012, **8**, 571.

22 E. Yus, T. Maier, K. Michalodimitrakis, V. van Noort, T. Yamada, W.-H. Chen, J. A. H. Wodke, M. Güell, S. Martínez, R. Bourgeois, S. Kühner, E. Raineri, I. Letunic, O. V. Kalinina, M. Rode, R. Herrmann, R. Gutiérrez-Gallego, R. B. Russell, A.-C. Gavin, P. Bork and L. Serrano, *Science*, 2009, **326**, 1263–1268.

23 M. J. van der Werf, K. M. Overkamp, B. Muilwijk, L. Coulier and T. Hankemeier, *Anal. Biochem.*, 2007, **370**, 17–25.

24 M. Bu, D. Czernik, J. C. Ewald, U. Sauer, J. M. Büscher and N. Zamboni, *Metab., Clin. Exp.*, 2009, **81**, 2135–2143.

25 M. Liebeke, K. Dörries, D. Zühlke, J. Bernhardt, S. Fuchs, J. Pané-Farré, S. Engelmann, U. Völker, R. Bode, T. Dandekar, U. Lindequist, M. Hecker and M. Lalk, *Mol. BioSyst.*, 2011, **7**, 1241–1253.

26 T. Soga, Y. Ohashi, Y. Ueno, H. Naraoka, M. Tomita and T. Nishioka, *J. Proteome Res.*, 2003, **2**, 488–494.

27 M. J. van der Werf, K. M. Overkamp, B. Muilwijk, M. M. Koek, B. J. C. van der Werff-van der Vat, R. H. Jellema, L. Coulier and T. Hankemeier, *Mol. BioSyst.*, 2008, **4**, 315–327.

28 R. t'Kindt, R. A. Scheltema, A. Jankevics, K. Brunker, S. Rijal, J.-C. Dujardin, R. Breitling, D. G. Watson, G. H. Coombs and S. Decuypere, *PLoS Neglected Trop. Dis.*, 2010, **4**, e904.

29 J. Weiner III, *Nucleic Acids Res.*, 2003, **31**, 6306–6320.

30 D. D. Buechter, D. N. Paolella, B. S. Leslie, M. S. Brown, K. A. Mehos and E. A. Gruskin, *J. Biol. Chem.*, 2003, **278**, 645–650.

31 J. K. Johnson and N. L. Somerson, *Appl. Environ. Microbiol.*, 1980, **40**, 391–399.

32 J. D. Pollack, N. L. Somerson and L. B. Senterfit, *Infect. Immun.*, 1970, **2**, 326–339.

33 R. N. McElhaney and M. E. Tourtellotte, *Science*, 1969, **164**(3878), 433–434.

34 A. W. Curnow, K. W. Hong, R. Yuan, S. I. Kim, O. Martins, W. Winkler, T. M. Henkin and D. Söll, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 11819–11826.

35 H. de Koning and G. Diallinas, *Mol. Membr. Biol.*, 2000, **17**, 75–94.

36 B. D. Bennett, E. H. Kimball, M. Gao, R. Osterhout, S. J. Van, J. D. Rabinowitz and S. J. Van Dien, *Nat. Chem. Biol.*, 2009, **5**, 593–599.

37 R. Costenoble, P. Picotti, L. Reiter, R. Stallmach, M. Heinemann, U. Sauer and R. Aebersold, *Mol. Syst. Biol.*, 2011, **7**, 464.

38 M. Papagianni, N. Avramidis and G. Filiousis, *Microb. Cell Fact.*, 2007, **6**, 16.

39 R. Schuetz, N. Zamboni, M. Zampieri, M. Heinemann and U. Sauer, *Science*, 2012, **336**, 601–604.

40 E. C. Y. Chan, K. K. Pasikanti and J. K. Nicholson, *Nat. Protocols*, 2011, **6**, 1483–1499.

41 T. G. Sobolevsky, A. I. Revelsky, B. Miller, V. Oriedo, E. S. Chernetsova and I. A. Revelsky, *J. Sep. Sci.*, 2003, **26**, 1474–1478.

42 G. Corso, M. Esposito and M. Gallo, *Biol. Mass Spectrom.*, 1993, **22**, 698–702.

 This journal is © The Royal Society of Chemistry 2013

Published on 27 March 2013. Downloaded by Max-Planck-Institut für marine Mikrobiologie on 14/11/2013 08:48:37.

43 O. Beckonert, H. C. Keun, T. M. D. Ebbels, J. Bundy, E. Holmes, J. C. Lindon and J. K. Nicholson, *Nat. Protocols*, 2007, **2**, 2692–2703.

44 D. S. Wishart, C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, N. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J. A. Cruz, E. Lim, C. A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuniga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazyrova, R. Shaykhutdinov, L. Li, H. J. Vogel and I. Forsythe, *Nucleic Acids Res.*, 2009, **37**, D603–D610.

45 K. Spagou, I. D. Wilson, P. Masson, G. Theodoridis, N. Raikos, M. Coen, E. Holmes, J. C. Lindon, R. S. Plumb, J. K. Nicholson and E. J. Want, *Anal. Chem.*, 2011, **83**, 382–390.

46 E. J. Want, I. D. Wilson, H. Gika, G. Theodoridis, R. S. Plumb, J. Shockcor, E. Holmes and J. K. Nicholson, *Nat. Protocols*, 2010, **5**, 1005–1018.

47 M. Liebeke, H. Meyer, S. Donat, K. Ohlsen and M. Lalk, *Chem. Biol.*, 2010, **17**, 820–830.

48 H. Tong, D. Bell, K. Tabei and M. M. Siegel, *J. Am. Soc. Mass Spectrom.*, 1999, **10**(11), 1174–1187.

49 A. Jankevics, M. E. Merlo, M. Vries, R. J. Vonk, E. Takano and R. Breitling, *Metabolomics*, 2011, (suppl. 1), 29–36.

50 C. A. Smith, G. O'Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan and G. Siuzdak, *Ther. Drug Monit.*, 2005, **27**, 747–751.

51 H. Horai, M. Arita, S. Kanaya, Y. Nihei, T. Ikeda, K. Suwa, Y. Ojima, K. Tanaka, S. Tanaka, K. Aoshima, Y. Oda, Y. Kakazu, M. Kusano, T. Tohge, F. Matsuda, Y. Sawada, M. Y. Hirai, H. Nakanishi, K. Ikeda, N. Akimoto, T. Maoka, H. Takahashi, T. Ara, N. Sakurai, H. Suzuki, D. Shibata, S. Neumann, T. Iida, K. Tanaka, K. Funatsu, F. Matsuura, T. Soga, R. Taguchi, K. Saito and T. Nishioka, *J. Mass Spectrom.*, 2010, **45**, 703–714.