Anal. Chem. **2006,** *78,* 1272−1281

Microbial Metabolomics with Gas Chromatography/ Mass Spectrometry

Maud M. Koek,*,† Bas Muilwijk,† Mariët J. van der Werf,† and Thomas Hankemeier†,‡

Analytical Science Department, TNO Quality of Life, Utrechtseweg 48, 3704 HE, Zeist, The Netherlands, and LACDR Analytical Biosciences, Leiden University, Einsteinweg 55, 2333 CC, Leiden, The Netherlands

An analytical method was set up suitable for the analysis of microbial metabolomes, consisting of an oximation and silylation derivatization reaction and subsequent analysis by gas chromatography coupled to mass spectrometry. Microbial matrixes contain many compounds that potentially interfere with either the derivatization procedure or analysis, such as high concentrations of salts, complex media or buffer components, or extremely high substrate and product concentrations. The developed method was extensively validated using different microorganisms, i.e., *Bacillus subtilis*, *Propionibacterium freudenreichii*, and *Escherichia coli*. Many metabolite classes could be analyzed with the method: alcohols, aldehydes, amino acids, amines, fatty acids, (phospho-) organic acids, sugars, sugar acids, (acyl-) sugar amines, sugar phosphate, purines, pyrimidines, and aromatic compounds. The derivatization reaction proved to be efficient (>50% transferred to derivatized form) and repeatable (relative standard deviations <10%). Linearity for most metabolites was satisfactory with regression coefficients better than 0.996. Quantification limits were 40−500 pg on-column or 0.1−0.7 mmol/g of microbial cells (dry weight). Generally, intrabatch precision (repeatability) and interbatch precision (reproducibility) for the analysis of metabolites in cell extracts was better than 10 and 15%, respectively. Notwithstanding the nontargeted character of the method and complex microbial matrix, analytical performance for most metabolites fit the requirements for target analysis in bioanalysis. The suitability of the method was demonstrated by analysis of *E. coli* samples harvested at different growth phases.

Functional genomics techniques (transcriptomics, proteomics, and metabolomics) are becoming increasingly important in the life sciences. The aim of these techniques is to gain new insights and a better understanding of the biological functioning of a cell or organism.[1,2]

Metabolomics involves the nontargeted, holistic analysis of the changes in the complete set of metabolites (small organic compounds, MW < 1000) in the cell (the metabolome), body fluids, or tissue.[3] As the biochemical level of the metabolome is closest to that of the function of a cell (the phenotype), the study of the metabolome is key in understanding biological functioning.[1] By analyzing differences between metabolomes using biostatistics (multivariate data analysis; pattern recognition), metabolites relevant to a specific phenotypic characteristic can be identified. By using such a nontargeted, holistic approach instead of the traditional hypothesis-driven approach, metabolomics studies can lead to new insights in cellular behavior.[1,2]

The key issue in metabolomics is the translation of differences in the metabolomes into the phenotypic differences of the cells that these metabolome samples were derived from. Therefore, analytical methods used for metabolomics studies should be sensitive, quantitative, and robust.[1] The development of a generic analytical method fulfilling these requirements is very challenging, especially in view of the wide range of compound classes and large range of metabolite concentrations present in the samples. Another challenge applies especially to the analysis of microbial samples: large amounts and numbers of components derived from the growth medium and the buffer used for quenching (nutrients, salts, buffers, etc.) may be present, and their concentration may vary significantly from sample to sample, for instance, when comparing microorganisms grown on different growth media or harvested at different time points during growth. Due to the high concentrations, these matrix compounds can be a potential disturbance during derivatization or analysis and influence the performance of the complete analysis. Obviously, the precision and performance of the analytical method have to be good enough so that the variation due to sampling, sample workup, and analysis are smaller than the biological variations to be detected. Therefore, an elaborate method validation is needed to check the performance of the method for metabolites from different compound classes.

For the analysis of the complete metabolome, invasive sample preparation techniques (extraction of metabolites from the cell) coupled with hyphenated analysis techniques, such as gas chromatography (GC), capillary electrophoresis (CE), or liquid chromatography (LC) coupled with mass spectrometric detection (MS) are currently preferred.[4−6] However, none of the individual

* Corresponding author. E-mail: koek@voeding.tno.nl. Phone: +31 30 694 4309.

† TNO Quality of Life.

‡ Leiden University.

(1) van der Werf, M. J.; Hankemeier, Th.; Jellema, R. H. *J. Ind. Microbiol. Biotechnol.* **2005**, *32*, 234−252.

(2) van der Werf, M. J. *Trends Biotechnol.* **2005**, *23*, 11−16.

(3) Fiehn, O. *Plant Mol. Biol.* **2002**, *48*, 155−71.

(4) Dunn, W. B.; Ellis, D. *Trends Anal. Chem.* **2005**, *24*, 285−294.

(5) Greef, J. v. d.; Stroobant, P.; Heijden, R. v. d. *Curr. Opin. Chem. Biol.* **2004**, 8, 559−565.

(6) Fiehn, O.; Weckwerth, W. *Eur. J. Biochem.* **2003**, *270*, 579−88.

10.1021/ac051683+ CCC: $33.50
Published on Web 01/12/2006

analytical methods will cover the full range of metabolites present in cells. Therefore, a generic metabolomics platform was set up in our laboratory to analyze ultimately all metabolites using a combination of GC/MS and LC/MS methods. In this paper, a GC/MS method is presented that is part of the platform, covering a large part of small (MW < 800) polar metabolites in cells.

GC/MS using electron impact (EI) ionization combines very high separation power and reproducible retention times with a versatile, sensitive, and selective mass detection. As the full scan response of the EI ionization mode for quadrupole instruments is approximately proportional to the amount of compound injected, i.e., more or less independently of the compound, all compounds suitable for GC analysis are detected nondiscriminatively. This makes the technique very suitable for comprehensive nontarget analysis, i.e., the analysis of a wide range of metabolites. Also the assignment of the identity of peaks detected with GC/MS using EI ionization via a database of mass spectra is straightforward, due to the extensive and reproducible fragmentation patterns obtained. If the MS spectrum is not present in the database, the fragmentation pattern can be used to obtain more information about the identity or compound class of a metabolite.

However, many metabolites contain polar functional groups that are thermally labile at the temperatures required for their separation or are not volatile at all. Therefore, derivatization of the compounds prior to GC analysis is necessary.

Several GC/MS-based analytical methods were reported for the analysis of a large number of compounds such as amino acids, sugars, and organic acids.[7−16] Most of these methods rely on derivatization with an oximation reagent, followed by silylation. The alkylsilyl reagents are the most versatile and universally applicable derivatizing agents for GC.[17,18] Nearly all functional groups present in metabolites that are problematic in gas chromatographic analysis, such as hydroxyl, amine, amide, phosphate, and thiol groups, can be converted to alkylsilyl derivates.[17,18] Direct silylation of sugars leads to a number of different peaks for every individual sugar compound, related to cyclic and open-chain structures. By introducing an oximation step prior to silylation, cyclization is inhibited, resulting in fewer peaks per sugar. In addition, α-ketoacids are protected against decarboxylation, and enolizable keto groups are fixed by oximation, facilitating the identification of the original molecular structures of metabolites.[17]

Almost all published GC/MS studies using derivatization with silylation or both oximation and silylation prior to analysis involve the measurement of plant metabolites[7,10−15] or urine samples.[9,16] So far, only one GC/MS-based method has been reported for the analysis of a wider range of metabolites in bacterial samples, but the validation of this method was rather limited.[8]

In this study, the performance and validation of a generic GC/MS method suitable for the analysis of microbial metabolomes is described. The application range of the GC/MS method was extensively tested and optimized. The method was applied to a set of samples of *Escherichia coli* harvested at different growth phases.

EXPERIMENTAL SECTION

Chemicals. Pyridine (Baker analyzed) was purchased from Mallinckrodt Baker BV (Deventer, The Netherlands). A solution of 56 mg of ethoxyamine hydrochloride (>99%, Acros Organics, Geel, Belgium) per milliliter of pyridine was used for oximation and *N*-methyl-*N*-trimethylsilyl trifluoroacetamide (MSTFA; Alltech, Breda, The Netherlands) was used for silylation.

Standards. Standards (e.g., Table 1) used for method optimization and for the determination of the application range were purchased from Sigma-Aldrich Chemie B.V. (Zwijndrecht, The Netherlands). The ^{2}H,^{15}N-labeled amino acid mix standard (20 different labeled amino acids) was purchased from Spectra Stable Isotopes. Stock solutions for determining the derivatization efficiencies were prepared in pyridine (∼1000 ng/μL); when metabolites were insoluble in pyridine, methanol/water (1:4 v/v) was used. Stock solutions of the various metabolites for spiking of cell extracts prior to lyophilization were prepared in an appropriate solvent (∼1000 ng/μL), preferably methanol/water (1:4 v/v).

Internal Quality Standards. Five different (deuterated) internal quality standards were used to monitor the performance of the GC/MS method during metabolomics studies. During method optimization, these standards were not always added. Phenylalanine-d_5 in methanol/water (1:4 v/v) was added prior to extraction. Leucine-d_3 and glucose-d_7 in methanol/water (1:4 v/v) were added prior to lyophilization. Alanine-d_4 and dicyclohexylphthalate in pyridine were added prior to derivatization. Stock solutions with a concentration of ∼1000 ng/μL were prepared. Cell extracts were spiked with an amount that resulted in a concentration of the compound of ∼10 ng/μL in the derivatized sample.

When disturbance from the naturally occurring metabolite was expected, alternative quality standards with comparable properties were used. For example, *E. coli* used in this study produces large amounts of phenylalanine, complicating the quantification of phenylalanine-d_5. In this case, alanine-d_4 was spiked before extraction and glutamic acid-d_3 was added prior to derivatization.

Microbial Samples. *Bacillus subtilis* strain 168 (ATCC 23857), *E. coli* NST 74 (ATCC 31884), and *Propionibacterium freudenreichii* VTD1 (ATCC 6207) were all obtained from the ATCC (Manassas, VA). *B. subtilis*, *E. coli,* and *P. freudenreichii* cells were grown under controlled conditions in a batch fermentor (Bioflow II, New Brunswick Scientific) at 30, 30, and 28 °C, respectively. The

(7) Fiehn, O.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. *Anal. Chem.* **2000**, *72*, 3573−3580.
(8) Strelkov, S.; Elstermann, M. v.; Schomburg, D. *Biol. Chem.* **2004**, *385*, 853−861.
(9) Kuhara, T. *J. Chromatogr., B* **2001**, *758*, 3−25.
(10) Chen, Z. L.; Landman, P.; Colmer, T. D.; Adams, M. A. *Anal. Biochem.* **1998**, 259, 203−211.
(11) Adams, M. A.; Chen, Z. L.; Landman, P.; Colmer, T. D. *Anal. Biochem.* **1999**, 266, 77−84.
(12) Katona, Z. F.; Sass, P.; Molnar-Perl, I. *J. Chromatogr., A* **1999**, *847*, 91−102.
(13) Roessner, U.; Wagner, C.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. *Plant J.* **2000**, *23*, 131−142.
(14) Gullberg, J.; Jonsson, P.; Nordstrom, A.; Sjostrom, M.; Moritz, T. *Anal. Biochem.* **2004**, *331*, 283−295.
(15) Jonsson, N.; Gullberg, J.; Nordstrom, A.; Kusano, M.; Kowalczyk, M.; Sjostrom, M.; Moritz, T. *Anal. Chem.* **2004**, *76*, 1738−1745.
(16) Paik, M. J.; Kim, K. R. *J. Chromatogr., A* **2004**, *1034*, 13−23.
(17) Blau, K., Halket, J., Eds. *Handbook of derivatives for chromatography*, 2nd ed.; John Wiley & Sons Ltd.: Chichester, U.K., 1993.
(18) Knapp, D. R. *Handbook of Analytical Derivatization Reactions,* 1st ed.; John Wiley & Sons: New York, 1979.

Table 1. Repeatability (RSD) of the Response and Derivatization Efficiency for Several Metabolites[a]

compound	*n*	RSD (%)	derivatization efficiency (%)
amino acids			
alanine	5	5	110
asparagine	6	7	30
aspartic acid	6	10	70
glutamic acid	6	9	50
glutamine	6	11	40
glycine	6	3	100
isoleucine	6	2	75
leucine	6	2	85
lysine	6	7	55
methionine	5	11	65
phenylalanine	5	5	80
proline	5	7	70
serine	5	7	80
threonine	6	3	70
tryptophan	6	12	25
valine	5	4	105
organic acids			
citric acid	6	6	75
fumaric acid	6	2	60
lactic acid	6	1	90
malic acid	6	3	60
oxaloacetaat	6	2	80
pyruvic acid	6	2	70
sugars			
2-deoxyglucose	6	4	80
fructose	6	2	95
glucose	5	4	85
ribose	6	3	95
xylitol	6	5	115
sugar phosphates			
fructose 6-phosphate	6	6	45
glucose 6-phosphate[b]	6	5−10	50−65
other			
5-fluorocytosin	5	6	25
glyceraldehyde 3-phosphate	5	7	30
glycerol 3-phosphate	6	4	30

[a] For all compounds, an amount between 9 and 16 ng was finally injected into the GC/MS. [b] The RSD and efficiency was determined from five different series of standards, measured during one year with different GC/MS instruments with the same or comparable setup.

fermentors contained 2 L of mineral salts medium D[19] containing 50 mg/L L-tryptophan and 10 g/L glucose, MMT12 medium,[20] or SLB medium[21] at pH 6.8, 6.5 and 6.8, respectively.

In the case of *B. subtilis* and *E. coli*, the oxygen tension was maintained at 30% by automatic increase of the stirring speed in the fermentor, while with *P. freudenreichii*, the headspace of the fermentor was flushed with nitrogen (0.05 L/min). Samples from *B. subtilis* and *P. freudenreichii* bioreactors were taken at the midlogarithmic phase. Samples from *E. coli* bioreactors were taken at different time points during growth.

Quenching and Extraction. Samples (~0.5 g of dry weight) were taken as quickly as possible from the fermentor and immediately quenched, to halt cellular metabolism, in −45 °C in methanol as described previously.[22] Prior to extraction, an internal standard (phenylalanine-d_5 or alanine-d_4) was added and a sample was taken for biomass determination. The biomass content of the samples was established by determining the dry weight of the sample. The intracellular metabolites were extracted from the cell suspensions by chloroform extraction at −45 °C as described by Ruijter and Visser.[23] In brief, chloroform was added to the methanol/water mixture to break the cell walls and denaturate the enzymes. Subsequently, the samples were shaken to extract the metabolites and centrifuged to separate the water/methanol and chloroform phases. The water/methanol phase containing the extracted metabolites was used for further sample workup.

Derivatization. Cell extracts (methanol/water, 50:50 v/v) or standard solutions were lyophilized at −37 °C in autosampler vials. The dry extracts were derivatized with 10 μL of a 56 mg/mL ethoxyamine hydrochloride solution in pyridine and 20 μL of pyridine for 90 min at 40 °C. Subsequently, the extracts were silylated for 50 min at 40 °C with 70 μL of MSTFA.

GC/MS Analysis. The derivatized extracts were analyzed with an Agilent 6890 gas chromatograph coupled with an Agilent 5973 mass selective detector. The 1-μL aliquots of the extracts were injected into a DB5-MS capillary column (30 m × 250 μm i.d., 0.25-μm film thickness; J&W Scientific, Folson, CA) using PTV injection (Gerstel CIS4 injector) in the splitless mode. The temperature of the PTV was 70 °C during injection, and 0.6 min after injection, the temperature was raised to 300 °C at a rate of 2 °C/s and held at 300 °C for 20 min. The initial GC oven temperature was 70 °C, 5 min after injection the GC oven temperature was increased with 5 °C/min to 320 °C and held for 5 min at 320 °C. Helium was used as a carrier gas and pressure programmed such that the helium flow was kept constant at a flow rate of 1.7 mL/min. Detection was achieved using MS detection in electron impact mode and full scan monitoring mode (m/z 15−800). The temperature of the ion source was set at 250 °C and that of the quadrupole at 200 °C.

Calculation of Derivatization Efficiency. To determine the efficiency of the derivatization, i.e., the percentage of the amount of a compound that is transferred into its derivatized form, the derivatized compounds were quantified in a semiquantitative manner by assuming that the response for a compound in the total ion chromatogram was proportional to the amount of compound injected.[24] Prerequisites for this assumption are that the quadrupole mass spectrometer is properly tuned and that (almost) all fragment ions produced during EI ionization are acquired during the scan of the mass spectrometer. In addition, no discrimination of the analyte during the GC analysis, i.e., injection and separation, and sample pretreatment may occur. By comparing the response of the derivatized compounds with reference compounds of a known concentration, the amount of injected derivative could be estimated with an accuracy of ~30%. As the amount of (underivatized) metabolite used for sample workup was known, the percentage transferred to its derivatized

(19) Hartmans, S.; Smits, J. P.; van der Werf, M. J.; Volkering, F.; de Bont, J. A. M. *Appl. Environ. Microbiol.* **1989**, *55*, 2850−2855.
(20) Tribe, D. E. Novel microorganism and method. Austgen-Biojet International Pty, Ltd. Patent 459302[4,681,852]. 21−7−1987. United States. 18−1−1983.
(21) Jore, J. P. M.; van Luijk, N.; van der Werf, M. J.; Pouwels, P. H. *Appl. Environ. Microbiol.* **2001**, *67*, 499−503.
(22) Pieterse, B.; Jellema, R. H.; van der Werf, M. J. J. *Microbiol. Methods*, published online ahead of publication.
(23) Ruijter, G. J. G.; Visser, J. *J. Microbiol. Methods* **1996**, *25*, 295−302.
(24) Hankemeier, T. Quantitative determination with GC-MS: the principle of molar response. Master of Science thesis, University of Ulm, Germany, 1992.

form could be calculated. The calculation was done using a set of *n*-alkanes as reference compounds.

Method Optimization. The derivatization and GC/MS analysis were optimized using a representative set of test compounds with varying physical and chemical properties. For this purpose, metabolites from different chemical classes, i.e., amino acids, organic acids, sugars, and sugar phosphates, were chosen. Several parameters were optimized, e.g., derivatization solvent (i.e., acetonitrile, dimethylformamide, dimethyl sulfoxide, pyridine, tetrahydrofuran), oximation reagents (hydroxylamine, ethoxyamine), silylation reagents (*N*,*O*-bis(trimethylsilyl)acetamide (BSA), MSTFA, *N*-methyl-*N*-(*tert*-butyldimethylsilyl)trifluoroacetamide, and a mixture of trimethylsilylimidazole/BSA/trimethylchlorosilane 3:3:2 v/v), derivatization times (15−90 min), and temperatures (30−70 °C). The final method parameters were chosen on the basis of the derivatization efficiencies of the test compounds. Also, the volatility of the reagents (and byproducts of the reagents) and solvents was taken into account, to maintain the application range as broad as possible. The derivatization temperatures were kept as low as possible to prevent breakdown of unstable metabolites. The combination of pyridine as solvent, ethoxyamine as oximation reagent, and MSTFA as silylation reagent resulted in the most satisfactory results with respect to derivatization efficiencies and application range (data not shown).

RESULTS

Repeatability and Efficiency of Derivatization. The derivatization efficiency and repeatability of derivatization of 32 standards covering different chemical classes were determined (Table 1). The relative standard deviations (RSDs) of the response and the derivatization efficiencies for most test compounds were satisfactory, i.e., with RSDs below 10% and derivatization efficiencies higher than 50%, respectively. All compounds with high derivatization efficiency (>70%) could be derivatized very reproducibly (RSD <5%). Compounds with phosphate and amide (i.e., glutamine, cytosine) functional groups had lower derivatization efficiencies, but most of these compounds could still be derivatized reproducibly. The calculated derivatization efficiency is an estimate (cf. Experimental Section) and can, therefore, deviate from the actual derivatization efficiency, if, for example, the MS response factor of a compound is lower or higher compared to the reference compounds used for quantification. For example, the ratio of response of glucose 6-phosphate and glucose was higher when analyzed with GC-FID compared to GC/MS. The derivatization efficiency for glucose 6-phosphate calculated from the GC-FID data (using effective carbon number concept[26]) was 80% (data not shown) and was comparable with the derivatization efficiency for glucose.

Linearity of Response. Different volumes of standard solutions of fluorinated or deuterated test compounds were spiked to *E. coli* cell extracts prior to lyophilization to determine the linear range. The concentrations of most test compounds in the sample solution after derivatization ranged from about 0.2 up to 50 ng/μL. As the N^{15},H^2-labeled amino acids were obtained as one reference mix, the concentration ranges of the various amino acids were different. The samples were measured with the GC/MS method, and the calibration curves for the test compounds were calculated (Table 2). The calibration curves for most test compounds were satisfactory with regression coefficients better than 0.996. Some of the compounds, such as glutamine and cholic acid, had a nonlinear response at lower concentrations: the linear dynamic range of these compounds started at higher concentrations compared to the other metabolites tested. This is likely caused by adsorption to the analytical system or breakdown of a small amount of the derivatized compound. These phenomena have a larger influence on the response at lower concentrations.

Table 2. Linearity and Quantification Limit of Standards Spiked to *E. coli* Cell Extracts

compound	linear from (ng/μL)	*n*	$r^{2\,d}$	quantification limit: pg on column[b] (S/N ratio)	quantification limit: Mmol/g DGT[c]
alanine[a]	≥0.2	8	0.9956	180 (70)	0.5
valine[a]	≥0.1	8	0.9979	110 (25)	0.2
leucine[a]	≥0.2	8	0.9981	200 (70)	0.4
isoleucine[a]	≥0.05	8	0.9979	70 (20)	0.1
threonine[a]	≥0.1	8	0.9977	280 (9)	0.6
proline[a]	≥0.15	8	0.9973	130 (30)	0.3
glycine[a]	≥0.15	8	0.9963	150 (100)	0.5
serine[a]	≥0.25	7	0.9976	260 (20)	0.6
histidine[a]	≥0.7	5	0.9958	700 (35)	1.1
tyrosine[a]	≥0.5	5	0.9975	200 (16)	0.3
methionine[a]	≥1	4	0.9995	40 (10)	0.1
aspartic acid[a]	≥5	4	0.9999	230 (10)	0.4
phenylalanine[a]	≥0.2	8	0.9981	200 (50)	0.3
asparagine[a]	≥5	3	0.9840	25000 (500)	4.3
glutamine[a]	≥10	3	0.9600	1200 (135)	2.0
glutamic acid[a]	≥0.25	8	0.9955	240 (30)	0.4
lysine[a]	≥5	4	0.9978	3600 (35)	2.3
chlorolactic acid	≥0.2	8	0.9980	200 (15)	0.4
6-fluoro-6-deoxyglucose	≥0.2	8	0.9999	500 (35)	0.7
5-fluorocytosin	≥0.5	7	0.9930	500 (35)	0.4
ribose 5-phosphate	≥ 5	5	0.9760	250 (20)	0.3
fructose 6-phosphate	≥ 5	5	0.9900	250 (15)	0.2
cholic acid-d_4	≥10	4	0.9910	2500 (40)	1.5

[a] Present in ^2H,^{15}N-labeled amino acid mixture, all hydrogens (except for NH, OH, and SH) and all nitrogen atoms labeled. Three of the labeled amino acids from the mix are not included in the table; arginine could not be detected with the GC/MS method; the tryptophan content of the mix was too low for detection (0.2% w/w) and due to low solubility in the spike solvent cysteine was detected only in two highest calibration standards cysteine concentration of >10 ng/μL) [b] Quantification limit (pg on-column) is the lowest calibration standard injected with a S/Nratio of ≥9. [c] Quantification limit (mmol/g of dry weight) is based on the fact that ∼4 mg of dry weight per sample was used for the sample workup for GC/MS analysis. The detection limit can therefore be improved by lyophilizing a larger portion of cell extract. [d] Regression coefficient is calculated for calibration line starting at the concentration given in the second column of this table, (mostly) up to 50 ng/μL.

Quantification Limit. Quantification limits of several compounds were determined by the analysis of *E. coli* extracts spiked with different amounts of labeled metabolite standards. The quantification limit was defined as the concentration of a compound resulting in a peak with a signal-to-noise ratio (S/N ratio) of nine. In most cases, the lowest spiked concentration had a higher S/N ratio, and the actual quantification limit was lower than the lowest spiked concentration. In such instances, the lowest spiked concentration was reported as the quantification limit, together with the corresponding S/N ratios (Table 2).

(25) Pierce, A. E. *Silylation of organic compounds: A technique for gas phase analysis*; Pierce Chemical Co.: Rockford, IL, 1968.
(26) Poole, C. F. *The essence of chromatography*, 1st ed.; Elsevier Science: Amsterdam, 2002.

Table 3. Recovery of Standards Spiked to *E. coli* Cell Extract

compounds	recovery (%)	compounds	recovery (%)
alanine[a]	75	glutamic acid[a]	120
valine[a]	100	asparagine[a]	100
chlorolactic acid	105	glutamine[a]	135
leucine[a]	105	lysine[a]	100
isoleucine[a]	100	tyrosine[a]	105
proline[a]	95	histidine[a]	70
glycine[a]	85	cholic acid-d_4	105
serine[a]	95	5-fluorocytosine	140
threonine[a]	95	2-fluorophenylalanine	115
methionine[a]	95	6-fluoro-6-deoxyglucose	105
aspartic acid[a]	95	5-fluorotryptophan	90
cysteine[a]	105	ribose 5-phosphate	110
phenylalanine[a]	115	fructose 6-phosphate	120

[a] From ^2H,^{15}N-amino acid mixture, all hydrogens (except for NH, OH, and SH) and all nitrogen atoms labeled. Two of the labeled amino acids from the mix are not included; arginine could not be measured with the GC/MS method, and tryptophan was present in a very low concentration (0.2% w/w) in the mix.

Recovery of Metabolites from Cell Extract. To study the recovery of metabolites from cell extracts, standard solutions of several deuterated or fluorinated compounds were spiked to cell extracts of *E. coli* prior to lyophilization at a concentration of ~15 ng/μL in the cell extract. The response of the compounds in the cell extract was compared with that of a standard solution of these compounds. The recoveries of all metabolites from cell extracts were satisfactory, i.e., 70−120% (Table 3). For glutamine and 5-fluorocytosine, a higher recovery of 135−140% was obtained. In the presence of matrix, the influence of adsorption to the analytical system or breakdown of the derivative on liner or column was less than in standard solutions, resulting in higher recoveries from cell extract compared to standard solutions for some compounds.

Stability of GC/MS System. The stability of the performance of the GC/MS system was investigated for the repetitive analysis of 30 cell extracts of *P. freudenreichii* and 18 standards injected between the microbial cell extract. The injection liner was not exchanged during the whole series (Table 4).

The RSD were good (i.e., better than 10%) for most spiked compounds and metabolites detected in the sample. Only the RSDs for phosphoenolpyruvic acid and 2-phosphoglyceric acid were high, i.e., 32 and 21% for the sample, but these two metabolites are suspected to be unstable. In general, the RSDs for the analysis of standards and the RSDs for the analysis of cell extracts were comparable (Table 4). However, for some compounds, i.e., phosphoenolpyruvic acid and cholic acid, slightly higher RSDs were obtained for standards compared to cell extracts. In addition, a decrease in the response of the cholic acid standard was observed after ~20 analyses. As a result, the RSD of the cholic acid standard after 15 microbial samples was 6%, after 20 microbial samples 12%, and after 30 microbial samples 26%.

The somewhat higher RSDs for phosphoenolpyruvic acid and cholic acid in the standard solutions compared to their RSDs in cell extract can probably be attributed to the presence or increase of active places in the analytical system when samples are injected. These active places are deactivated by compounds present in the sample matrix of the cell extract.[27] In standard solutions, these "protective" compounds from the matrix are not present, causing the RSDs to be higher than in cell extracts.

In general, 20 samples could be analyzed using the same injection liner. The performance of a few quality standards added to the microbial samples was checked for each measurement; if the performance of the quality standards deteriorated, the injection liner was changed. In some cases, it was necessary to remove a small piece from the front of the analytical column in addition to the exchange of the injection liner to restore the performance of the system to the initial level.

Precision of quantification. The intrabatch precision and interbatch precision of quantification in standards and cell extracts was tested by analyzing derivatized standard solutions and derivatized *B. subtilis* cell extracts 0, 1, 2. and 6 weeks after

Table 4. Stability of the Analysis of Metabolites in Standards and in *P. freudenreichii* Cell Extracts[a]

	amount injected (ng/μL)		RSD (%)	
metabolite	standard	cell extract	standard (n = 18)	cell extract (n = 30)
leucine-d_3	15	15	3	2
malic acid	38	19	3	2
phosphoenolpyruvic acid	31	31	(46)[c]	(32)[c]
phenylalanine-d_5	15	15	3	3
glutamic acid-d_3	15	15	7	5
2-phosphoglyceric acid	23	23	(24)[c]	(21)[c]
citric acid	38	19	3	5
fructose	38	19	1	1
ribose 5-phosphate	38	19	4	1
glucose 6-phosphate	38	19	6	2
lactose	38	19	2	2
cholic acid-d_5	15	15	12[b]	8
alanine	not present[a]	*a*	not present[a]	2
valine	not present[a]	*a*	not present[a]	2
proline	not present[a]	*a*	not present[a]	4
glycine	not present[a]	*a*	not present[a]	6
succinic acid	not present[a]	*a*	not present[a]	2

[a] Metabolite present in sample, concentration not known; not present in standard. [b] RSD of cholic acid in the standards after 15 microbial samples was 6% (n = 8), after 20 samples 12% (n = 10), and after 30 samples 26% (n = 18). [c] Unstable metabolite.

Table 5. Precision of Quantification[a] in Standard Solutions and Cell Extracts of *B. subtilis*

	Standard Solutions				
	metabolite	malic acid	fructose	glucose 6-phosphate	cholic acid
	act. conc (μg/L)	0.600	1.410	0.670	0.110
repeatability[b]	SD (μg/L)	0.018	0.041	0.038	0.007
(i.e., interbatch precision)	CV (%)	3	3	6	6
	df	26	26	26	26
reproducibility	SD (μg/L)	0.051	0.16	0.069	0.0200
(i.e., interbatch precision)	CV (%)	8	11	10	18
	df	3	3	3	3

	Cell Extracts			
metabolite	dihydroxy-benzoic acid	disaccharide	citric acid	glucose 6-phosphate
average area	61348	98004	142394	130069
repeatability SD (units)	3533	6757	8383	7490
(i.e., intrabatch precision) CV (%)	6	7	6	6
df	10	10	10	10
reproducibility SD (units)	7994	7230	16762	11796
(i.e., interbatch precision) CV (%)	13	7	12	9
df	3	3	3	3

[a] The repeatability and reproducibility were calculated according to analysis of variance calculation (one-way ANOVA): Miller, C.; Miller, J. N. *Statistics for Analytical Chemistry*, 2nd ed.; Ellis Horwood: Chichester, 1988. Abbreviations: act. conc, actual concentration; CV, coefficient of variance (=RSD); df, degree of freedom; SD, standard deviation. [b] The samples analyzed within one week were considered to belong to the same batch. The samples analyzed within various weeks were considered to belong to different batches.

storage. The peak areas of four metabolite standards, i.e., malic acid, fructose, glucose 6-phosphate and cholic acid, and four metabolites in cell extracts, i.e., dihydroxybenzoic acid, citric acid, glucose 6-phosphate, and an unknown disaccharide, were determined by integration of a peak of a characteristic mass from the mass spectrum for each metabolite, in appropriate reconstructed ion chromatograms. The areas were corrected for variations in injection volume and MS response with an internal standard, dicyclohexyl phthalate. For the metabolites in the standard solutions, the precision of quantification was estimated by calculating the concentrations via relative response factors in a database, with dicyclohexylphthalate as the reference. For the metabolites in the cell extracts, the peak areas were used to determine the precision (Table 5).

The intrabatch precision expressed as relative standard deviation was 3−6% in standards and 6−7% in cell extracts. The intrabatch precision is a measure for the repeatability when metabolites are measured within one series. This value is in agreement with the results in the previous paragraph; i.e., the RSDs for stable metabolites were generally 10% or better. The interbatch precision is a measure for the comparability of concentrations found in different sequences or when determining the concentration using the response factors stored in a database. For stable metabolites, the reproducibility (or interbatch precision) was about 8−11% in standard solutions and 8−14% in cell extracts, which allows a good comparison of samples analyzed in different series or quantification of samples using response factors stored in a database.

Application Range. *B. subtilis* was used as a model organism to establish the application range of the GC/MS method. Based on its annotated genome sequence, it was estimated that this bacterium could contain 580 different metabolites.[28] Over 80% of all commercially available standards of these metabolites (~300 compounds) were derivatized and analyzed. The GC/MS method allowed the detection of 70% (~200 metabolites) of all commercially available metabolites (unpublished results). For 160 of the metabolites, the expected derivative was formed and the recoveries for these compounds were satisfactory, i.e., larger than ~50%. For ~40 of the compounds, multiple peaks or degradation was observed, i.e., adenosine 5′-phosphosulfate, or the recoveries of the standards were low (<10%), i.e., uridine 5′-monophosphate. Compounds that could not be detected or quantified with the GC/MS method were mostly compounds more suitable for LC or CE analysis because of their high molecular weight or thermal instability, e.g., nucleotides and CoA esters. In some cases, the derivatized metabolites were too volatile to be measured with the described GC/MS method, e.g., acetic acid and glyoxilic acid.

DISCUSSION

In metabolomics, the aim is to determine the differences in the complete metabolomes of cells, body fluids, or tissue. The nontargeted approach can result in new insights in the functioning of biological systems, but the development of sensitive, quantitative, and reproducible comprehensive analytical methods needed to achieve this goal is very challenging. An extensive method optimization and validation is needed in order to reliably determine the differences between samples, sometimes measured over a larger period of time. The method performance has to be optimized and assessed for a broad range of different compound classes. For the development and optimization of the GC/MS

(27) Anastassiades, M.; Mastovska, K.; Lehotay, S. J. *J. Chromatogr., A* **2003**, *1015*, 163−184.

(28) Hall, R.; Beale, M.; Fiehn, O.; Hardy, N.; Sumner, L.; Bino, R. J. *Plant Cell* **2002**, *14*, 1437−1440

method described in this paper, an extensive set of test compounds was used. From our experience, microbial samples are more difficult to analyze than plant or mammalian samples, due to the high concentrations of compounds from growth media or buffers used during sample workup. Therefore, after the first optimization of the derivatization and analysis using standards of metabolites, the method performance was assessed by spiking deuterated and fluorated standards to "real-life" microbial samples. Three different and phylogentically different microorganisms, e.g., *B. subtilis*, *E. coli*, and *P. freudenreichii*, were used to asses the method performance in different sample matrixes.

In view of their analytical performance, metabolites could be classified into three groups.

The analytical performance of metabolites containing hydroxylic and carboxylic functional groups (so-called group I compounds), e.g., sugars and organic acids, were in general very satisfactory. These metabolites could be derivatized with high efficiency (60−115%) and could be measured very repeatable (RSDs < 5% in cell extracts and standard solutions) and reproducible (RSDs 8−14%). These metabolites had a large linear dynamic range (0.2−50 ng on-column), and determination limits were low (<500 pg on-column in cell extracts in full scan mode).

Metabolites containing amine or phosphoric functional groups (so-called group II compounds) could also be measured with satisfactory derivatization efficiencies (30−110%), repeatability (1−7%), and reproducibility (10%). For some of these metabolites, however, detection limits were somewhat higher (i.e., lysine) or the linear dynamic range started at somewhat higher concentrations (i.e., ribose 5-phosphate and fructose 6-phosphate).

Metabolites with amide, thiol, or sulfonic functional groups (so-called group III compounds) were more difficult to analyze. For example, asparagine and glutamine, both metabolites with an amide functional group, had higher detection limits and linear ranges started at higher concentrations than for most metabolites tested. Still, if the analytical system was in good shape, reasonable analytical performance was obtained at concentration levels within the linear range.

The difference in system performance of the different groups is related to the stability of the silylated reaction product. In general, the reactivity of the parent compound and the stability of the derivatized compound decreases in the following order: hydroxyl (alcohol) > hydroxyl (phenol) > carboxylic acid > amine > amide.[17,25] All group III compounds have relatively weak bonds with silicium and are, in fact, very good leaving groups, even when compared to *N*-methyltrifluoroacetamide, the leaving group of MSTFA. Despite their relatively low reactivity toward silylation, derivatives of group III compounds are formed, due to the large excess of MSTFA in the solution, but these derivatives are the first to react with active places in the analytical system or to break down on the injection liner or column.

As the performance of all metabolites and especially group III metabolites was depending on the overall state of the analytical system, quality control using internal and external standards was essential. Therefore, when the metabolomes of (different) microorganism were measured, a set of deuterated internal standards spiked at the different steps during sample workup, was used to monitor loss or disturbances during extraction (phenylalanine-d_5), lyophilization (glutamic acid-d_3), derivatization (glucose-d_7 for oximation and phenylalanine-d_5 for silylation), and GC/MS analysis (alanine-d_4, dicyclohexylphthalate). Phenylalanine-d_5 was also used to compensate for differences between samples in the amount of biomass used for sample workup. Degradation of the system performance could be detected in an early stage when monitoring the performance of external standards without the presence of matrix (cf. Experimental Section, stability of the GC/MS method). Therefore, these were used to determine whether the injection liner had to be changed, and eventually a short piece of the analytical column had to be removed.

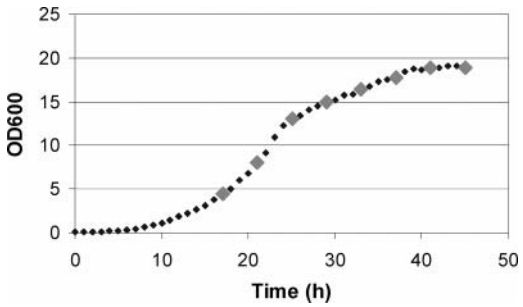

Figure 1. Growth curve of *E. coli* batch fermentation (phosphate limited). The eight sample points are represented by squares; OD600, optical density at 600 nm.

The derivatization efficiency was the only method performance parameter that was determined in standard solutions instead of in "real-life" samples. For the calculation of the derivatization efficiency of a metabolite, the full scan response of the derivatized compound is needed. This is only possible when no other metabolites coelute with the derivative of interest. As the chromatograms of cell extracts contain hundreds of different components, there are almost no compounds completely resolved.

However, by combining the results of the recovery from cell extracts (80−140%) and the linearity in cell extracts (linear dynamic range from 100 to 250 pg up to 50 ng for most of the metabolites) with the derivatization efficiencies in standards, it can be concluded that the derivatization efficiencies in real-life samples at high (50 ng/μL) as well as low concentrations (100−250 pg/μL) were comparable with the derivatization efficiencies in standard solutions.

We demonstrated that the method was quantitative and precise and the method performance was stable; the method was applied for a large number of studies and the repeatability and reproducibility of quality standards added to samples and metabolites present in the samples were generally better than 10 and 15%, respectively. Also, in the presence of high (varying) concentrations of matrix compounds from different growth media and extraction buffers, metabolites could be analyzed reliably. The recoveries of the internal quality standards in extracts of different organisms grown on clean mineral media, but also complex industrial media, were satisfactory (80−120%, data not shown). The derivatization was robust; the results for quality standards were comparable when the derivatization reaction was carried out in tubes and vials with different volumes or when extracts of different microorganisms were analyzed. In addition, the internal quality standards were able to detect variations introduced during sample workup or analyses that influenced method performance to a large extent. All together, the method allowed the comparison of large numbers of samples, measured over a larger period of time.

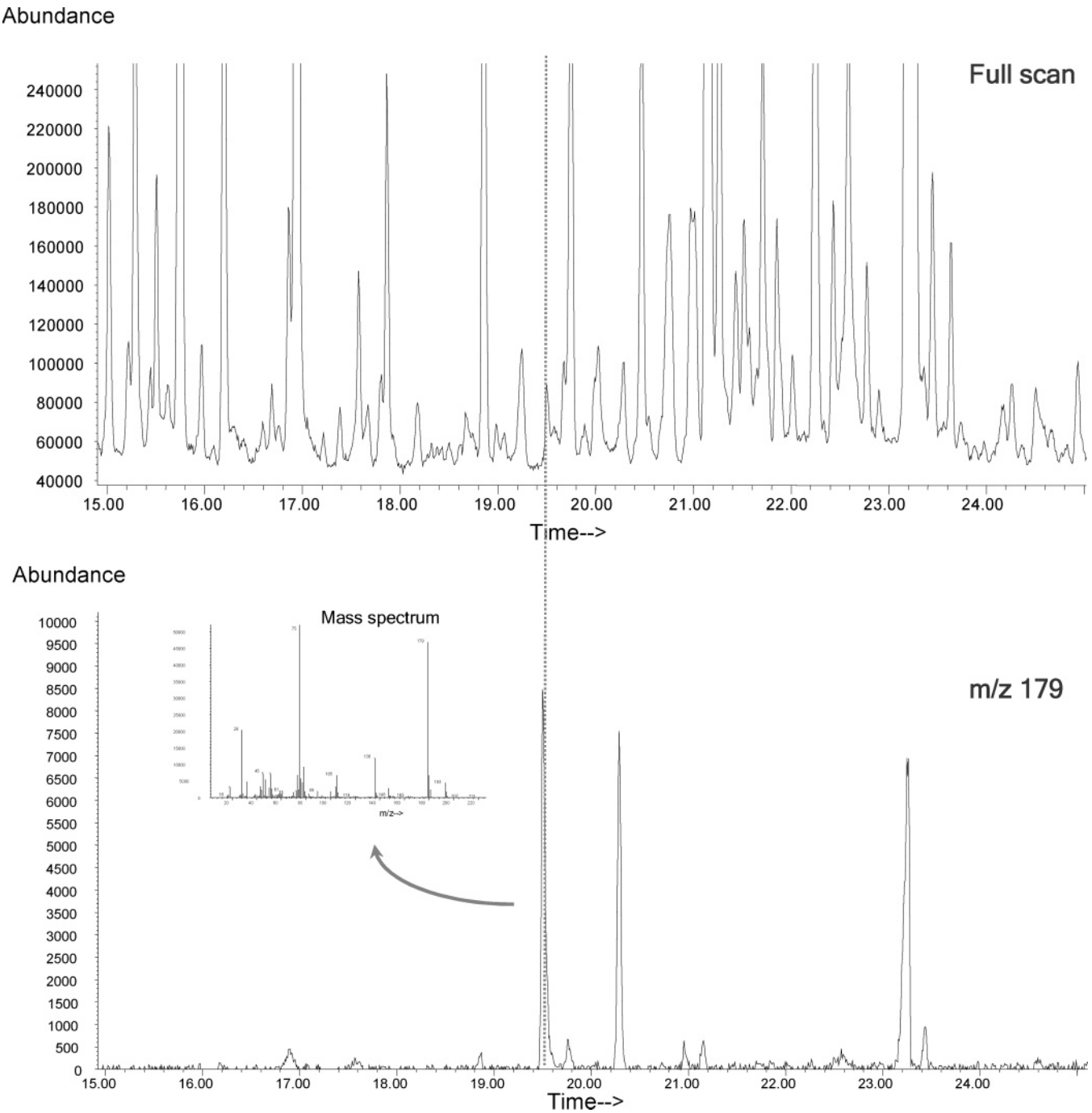

Figure 2. Targeted detection of metabolite nicotinamide in *E. coli* extracts, by reconstructing the ion chromatogram of $m/z = 179$ from the full scan GC/MS chromatogram.

The optimized GC/MS method was suitable for the analysis of a large variety of metabolite classes important for the biological functioning of cells, namely, alcohols, aldehydes, amino acids (also acyl amino acids and succinyl amino acids), amines, fatty acids, organic acids, phosphoorganic acids, sugars, sugar acids, (acetyl) sugar amines, sugar monophosphates, purines, pyrimidines, and aromatic compounds. The method covered a large volatility range; compounds as volatile as 1,2-butanediol up to trisaccharides (e.g., cellotriose) could be analyzed. In addition to the described GC/MS method, a complementary comprehensive LC/MS method was developed (Coulier et al., in preparation). The GC/MS and LC/MS methods together allowed the detection of 93% of the commercially available metabolites of the in silico metabolome of *B. subtilis* (unpublished results).

In conclusion, the presented method is a reliable and generic method that fulfills the requirements for metabolomics studies of microorganisms: the variation due to the overall analytical method (<10% for most metabolites) is typically much smaller than the biological variation we encountered in microbial metabolomics studies (see below).

Application: Differences in the Metabolome of *E. coli* during Growth. To prove the suitability of the GC/MS method for microbial metabolomics, the method was applied to a set of samples from phenylalanine producing *E. coli*. A total of eight

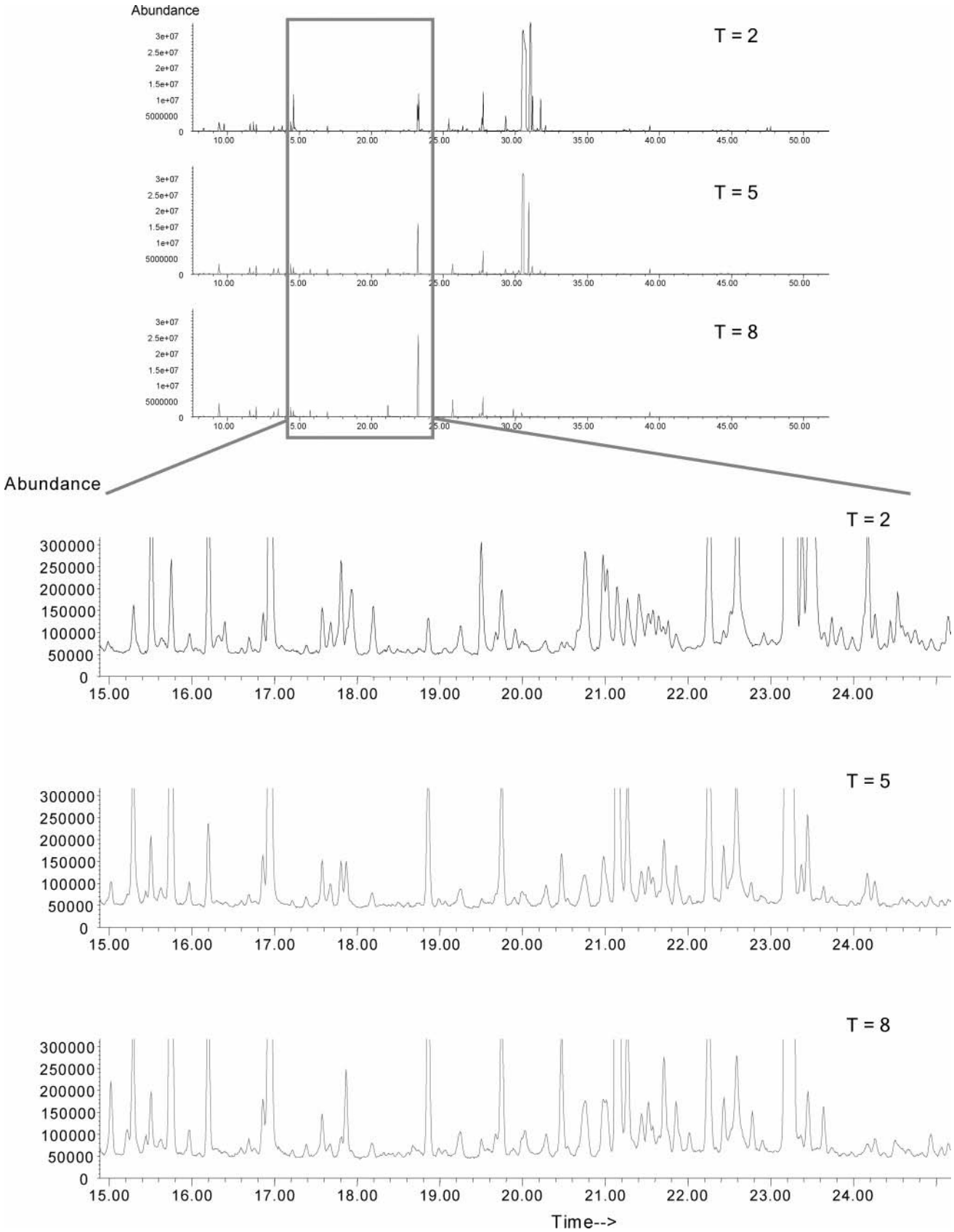

Figure 3. Full scan GC/MS chromatograms of *E. coli* samples taken at different time points during growth. (For the time points, cf. Figure 1.)

samples, harvested at different time points during fermentation (Figure 1), were analyzed and the concentrations for the different metabolites were compared. The first five samples were harvested in the logarithmic growth phase, and the last three samples were

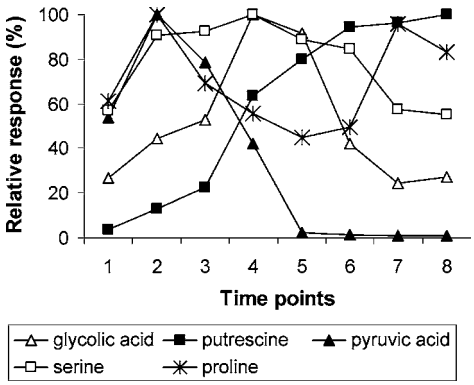

Figure 4. Relative responses of five metabolites with increasing (putrescine), decreasing (pyruvic acid), and stable (serine) concentration and with maximum (glycolic acid) and minimum (proline) concentration in the middle of the growth curve. Time point assignment corresponds to Figure 1. Highest response for each metabolite is set at 100%.

harvested in the stationary growth phase. Glucose was used as the carbon source. Glucose depleted after ~36 h, 1 h before the sixth sample was taken.

AMDIS software (Automated Mass Spectral Deconvolution and Identification System, V2.61,[29]) was used to determine the number of metabolites in the chromatograms. About 200−250 different metabolites could be detected in *E. coli* with the optimized GC/MS method. From these compounds, 60 were of known identity and present in our own database.

The response for each of the identified metabolites was determined in all samples, by manually integrating a peak of a specific mass from its mass spectrum in a reconstructed ion chromatogram (Figure 2). For each individual metabolite, the variation in concentration (RSD of peak areas) at different time points during growth was determined. The peak areas were corrected for differences in biomass by means of an internal standard (alanine-d_4) added to the sample before extraction of the intracellular metabolites from the *E. coli* cells (cf. Experimental Section). In Figure 3, the full scan GC/MS chromatograms of *E. coli* samples at three different time points are shown.

The RSDs of the response of the added internal quality standards were less than 10% and the recoveries were between 90 and 130%, indicating the performance of the GC/MS method was stable and satisfactory during the study. The variation in concentration of the determined metabolites at the different time points was significantly larger than the variation of the analytical method. Differences in metabolite concentrations up to a factor of 240 (RSD 280%) were observed. Some metabolites were only detected during specific time periods during growth (e.g., beginning of the growth curve).

(29) Stein, S. E. *J. Am. Soc. Mass. Spectrom.* **1999**, *10*, 770−781.

(30) U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), and Center for Veterinary Medicine (CVM). Guidance for Industry; Bioanalytical Method Validation, 2001.

Different trends in metabolite concentrations over time were observed (Figure 4), (1) concentrations increasing or (2) decreasing in time, (3) metabolites with a maximum or (4) minimum concentration in the middle of the growth curve, and (5) metabolites with an almost constant concentration during growth. Currently, the results of this application are studied in more detail.

It can be concluded that the GC/MS method was suitable to reliably analyze differences in the metabolome of *E. coli* during batch cultivation.

CONCLUSION

In this paper, an analytical method is presented for the quantitative analysis of microbial metabolites, consisting of a derivatization reaction combined with gas chromatographic analysis coupled to mass spectrometry. This method is part of a metabolomics platform in our laboratory consisting of several GC- and LC-based methods to analyze, ultimately, all metabolites from cells, body fluids, and tissue.

Notwithstanding the fact that the developed GC/MS method is a comprehensive method with a very large application range, still the analytical performance (with respect to stability of performance, precision, recoveries, and linear ranges) for the majority of metabolites, meets the requirements for target analysis in biological matrixes.[30] Even more importantly, the analytical variation of the method was much smaller than the biological variation in the *E. coli* samples, proving the suitability of the method to analyze differences in the metabolome of microorganisms.

The generic method has been applied for the analysis of a large number of microorganisms, e.g. *B. subtilis*, *P. freudenreichii*, *Trichoderma reesei*, and *Pseudomonas putida* (data not shown). In addition, the method has also been successfully applied to mammalian (blood plasma and serum, urine, tissue) and plant metabolomics studies (data not shown). Currently, the research is aimed at improving the performance of the method for critical metabolites (i.e., metabolites with amide, thiol, and sulfonic functional groups) and lowering the detection limits.

ACKNOWLEDGMENT

The authors thank Roelie Bijl, Machtelt Braaksma, Nicole van Luijk, and Karin Overkamp for cultivating the microorganisms and extracting the samples, William van Dongen for reviewing the manuscript, Roche Vitamins (currently DSM Nutritional Products), DSM Food Specialties and Friesland Foods for financial support, and Markus Wyss and Werner Bretzel (currently DSM Nutritional Products) for fruitful discussions.

Received for review September 21, 2005. Accepted December 14, 2005.

AC051683+