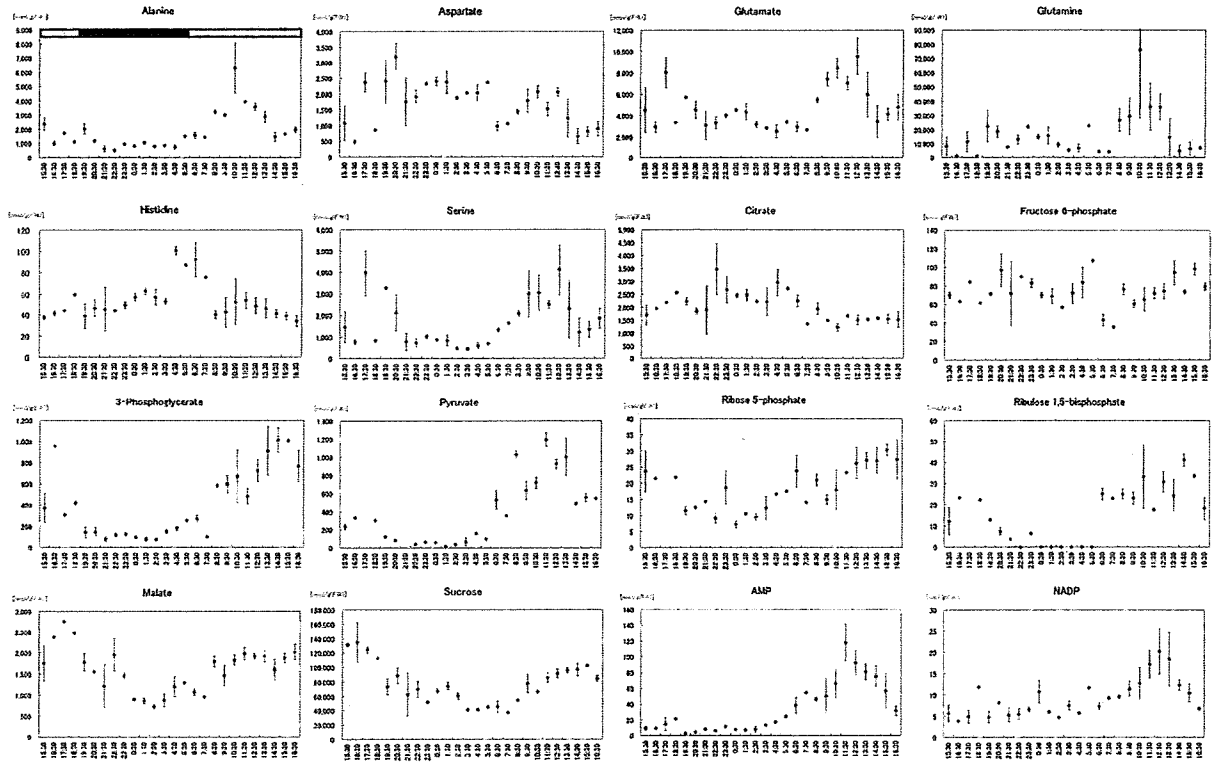

Figure 2

Metabolic time-courses in rice foliage at the third-leaf stage. Plantlets were grown under a 13-hr light – 11-hr dark photocycle. We applied 3 CE-MS methods and a CE-DAD method to analyze 69 major metabolites. Dynamic changes in the metabolite levels were assessed at hourly intervals over a 24 h period. Averages of 2 samples (± SEM) are shown. The top bar (shown in only Ala) indicates light and dark conditions.

colytic pathway (M4), the latter half of the TCA cycle (M5), sugars (M7), and major amino acids (M1). Also included in this group were NADPH and NADH (M6), glutathione and spermidine (M8). Subsets M9 – M12 included the first half of the glycolytic pathway (M9), the first half of the TCA cycle (M10), and minor amino acids (M11); also included were the nucleoside tri- and diphosphates (M12). Thus, our SOM analysis correctly reflected the phenotypic metabolic variations that indicate functioning biochemical pathways, and therefore represents a phenotypic linkage map (PLM).

Table 3: Status of adenine nucleosides and nicotinamide coenzymes in the light and dark period

	ATP AdN*1	ADP AdN	AMP AdN	NAD NiC*2	NADH NiC	NADP NiC	NADPH NiC
Light*3	0.21	0.40	0.40	0.36	0.10	0.09	0.44
Dark*4	0.45	0.43	0.11	0.55	0.09	0.05	0.31

*1 AdN = ATP + ADP + AMP
*2 NiC = NAD + NADH + NADP + NADPH
*3 The average of all data throughout the light period
*4 The average of all data throughout the dark period

The advantages of this analysis became even more apparent upon time-resolved analysis of metabolite levels (Figure 3C), which allowed visualization of the dynamic activity of these metabolic modules (see Discussion).

Discussion

Estimation of unidentified metabolites with SOM analysis

Although S17P could not be directly identified, we hypothesized that its peak could be identified in CE-MS data by combining SOM analysis with knowledge of the chemical structure. We identified a candidate peak among several peaks on selected ion electropherograms using a simple estimation method. As electrophoretic mobility is proportional to the ionic charge of the solute and inversely proportional to the size of the ionic molecule related to the hydrated ionic radius of a spherical mole-

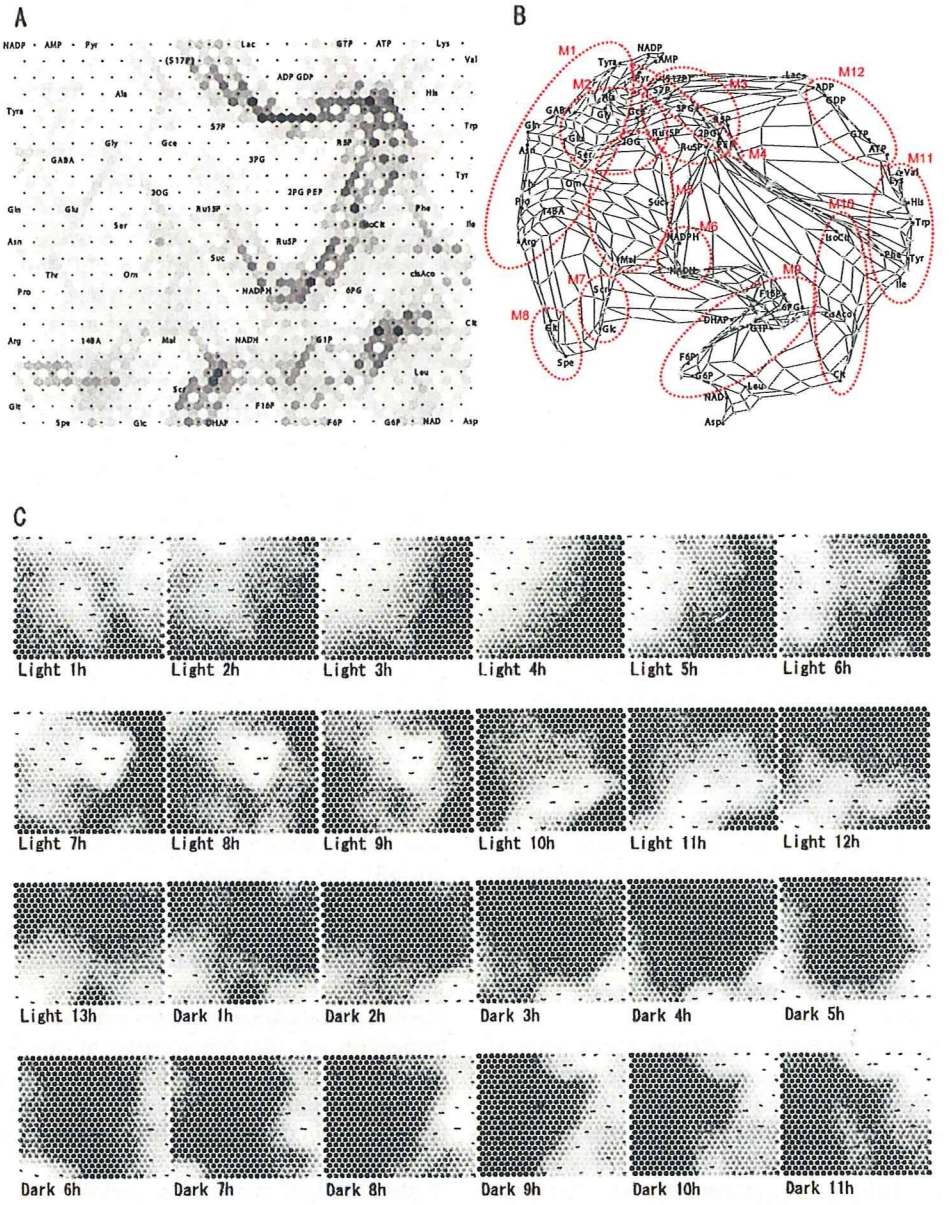

Figure 3

Self-organizing map (SOM) Analysis. A. U-matrix. Measured metabolites (n = 56) were arranged in a 20 × 20 lattice on the basis of diurnal change similarities. Light- and dark shading indicate high and low similarity, respectively. **B.** Phenotypic linkage map (PLM). The linkage among metabolites based on dynamic similarity is expressed as the distance on the quadratic plane. The metabolites were assigned to 14 metabolic modules that fluctuated synchronously; most contained traditional metabolic pathway networks or similar compounds. M1, major amino acid; M2, related to photorespiratory pathway intermediates; M3, pentose phosphate pathway; M4, latter half of the glycolytic pathway; M5, latter half of the TCA cycle; M6, environmental stress response; M7, sugars; M8, NADH and NADPH; M9, first half of the glycolytic pathway; M10; first half of the TCA cycle; M11, minor amino acids; M12, nucleoside tri- and diphosphates. **C.** Time-resolved layout. The relative levels of metabolites are shown for every time point from the start of the light period to the end of the dark period. Light and dark shading indicate high and low levels.

cule [26], we used the cubic root of the molecular weight as a substitute parameter for the radius. Indeed, the cubic root of molecular weights of 3 metabolites of similar chemical structure, Ru5P, F6P and S7P, were linearly correlated with migration time ratios ($r > 0.999$), when PIPES was used as an internal standard (Table 4).

The estimate for S17P was performed using linear approximation with Ru15P and F16P. The estimated migration time ratio (MT/MT_{IS}) of S17P was 0.941 (Table 4). Several peaks were observed at a mass-to-charge ratio (m/z) of 369. A peak of $MT/MT_{IS} = 0.909$ (m/z = 369) was identified within ± 5.0% of the predicted values.

Next, the absence of other metabolites with similar chemical structures was verified with the KEGG ligand database [27]. Note that except for S17P, metabolites were cyclic or non-anionic compounds.

Finally, we obtained the normalized time-course of the putative S17P by calculating the ratio of the peak area of putative S17P to PIPES. Integration of these data into the SOM analysis showed that this putative S17P marker was near metabolites in the reductive pentose phosphate pathway (Figure 3A) or the metabolic module M3 in PLM.

Unfortunately, the above result includes some speculation; most peaks of putative S17P were below the detection limits ($S/N < 3$) and the peak was not detected in the dark period. In the SOM analysis, the peak area of such undetected metabolite was calculated as zero. Nevertheless, the proposed estimation method seems to be effective in identifying unknown metabolites.

Detection of metabolic bottlenecks by pair-wise correlation analysis

In previous studies, Peason's correlation coefficients of metabolite pairs (pair-wise correlation) were applied to construct a metabolic correlation network [5,10,28]. A correlation coefficient is an index of co-linearity between two variables. If two metabolites, A and B, are always equilibrated, i.e., $[A]/[B] = K_{eq}$ (constant), then their relationship is linear and shows a high correlation. Although real metabolic pathways are dynamic and constantly regulated by their influx and/or efflux, the pathway components that are blocked by rate-limiting enzymes should exhibit approximate linearity. For example, 3PG, 2PG, and PEP in the glycolytic pathway are positioned between two rate-limiting enzymes, phosphoglycerate kinase (EC 2.7.2.3) and pyruvate kinase (PK; EC 2.7.1.40), both of which are regulated by the ATP/ADP ratio (Figure 1). The correlation coefficients among these three metabolites throughout a 24-hr period were over 0.90, whereas the correlation coefficient between PEP and Pyr, limited by PK, was under 0.50. Thus, pair-wise correlation analysis is effective for the identification of metabolic modules that are regulated by rate-limiting enzymes.

We used a hierarchical clustering algorithm, Ward's method [29], to classify metabolites in the glycolytic pathway (Figure 1) on the basis of their correlation matrix that was computed using all data throughout the 24-hr period. Indeed, a dendrogram identified the steps regulated by the ATP/ADP ratio (Figure 4A). On the other hand, it did not identify phosphofructokinase I (PFK-1; EC 2.7.1.11) as a rate-limiting enzyme. Although it is regulated by the ATP/ADP ratio in animal cells, another enzyme, pyrophosphate fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90), seems to be active in plant cells and may be independent of the ATP/ADP ratio [30].

The same cluster analysis was also applied to the TCA cycle intermediates (Figure 1), and the dendrogram revealed the rate-limiting enzymes in the cycle again (Figure 4B): citrate synthase (CS; EC 2.3.3.1), and NADP-dependent isocitrate dehydrogenase (ICDH; EC 1.1.1.42). This suggests that the classification of metabolites along enzymatic steps can help to reveal bottleneck enzymes.

Time-resolved carbon/nitrogen metabolomics

Inspection of the time-course of metabolic modules allowed us to better understand the carbon and nitrogen (C/N) assimilation/dissimilation process and their underlying function during a 24-hr period (Figure 3C).

In the first half of the light period, some accumulation emerged for carbon-fixed products: Pyr, 2OG, and photorespiratory pathway intermediates (metabolic module

Table 4: Estimated migration-time of unidentifiable metabolites based on the molecular weight of similar metabolites

Compound	Formula	M.W.	$M.W.^{1/3}$	MT/MT_{IS}
Ru5P	$CH_2(OH)CO\ [CH(OH)]_2CH_2OPO_3H_2$	230.0192	6.127	1.029
F6P	$CH_2(OH)CO\ [CH(OH)]_3CH_2OPO_3H_2$	260.0298	6.383	1.080
S7P	$CH_2(OH)CO\ [CH(OH)]_4CH_2OPO_3H_2$	290.0403	6.619	1.125
Ru15P	$CH_2(OPO_3H_2)CO\ [CH(OH)]_2CH_2OPO_3H_2$	309.9854	6.768	0.847
F16P	$CH_2(OPO_3H_2)CO\ [CH(OH)]_3CH_2OPO_3H_2$	339.9960	6.980	0.895
S17P	$CH_2(OPO_3H_2)CO\ [CH(OH)]_4CH_2OPO_3H_2$	370.0065	7.179	0.941*

*Estimated value. MT/MT_{IS} was calculated by linear approximation

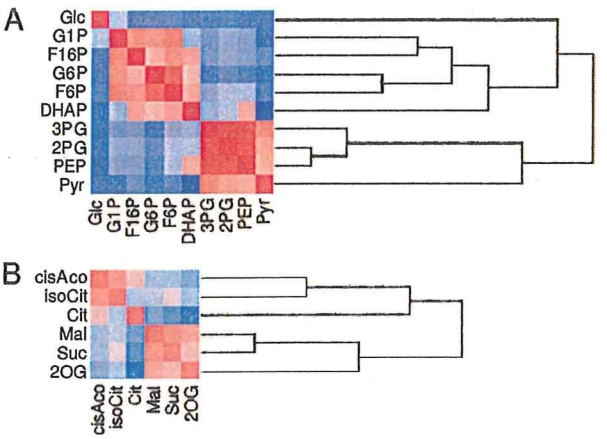

Figure 4
Hierarchical cluster analysis. A. Cluster analysis (Ward's method [26]) was applied to the correlation matrix composed of metabolic intermediates in the glycolytic pathway. The generated dendrogram was clustered into regulatory units by the ATP/ADP ratio; hexokinase (EC 2.7.1.1), phosphoglycerate kinase (EC 2.7.2.3), and pyruvate kinase (EC 2.7.1.40). **B.** As well as in the TCA cycle, the dendrogram was divided into two major groups at the rate-limiting steps; citrate synthase (CS; EC 2.3.3.1), and NADP-dependent isocitrate dehydrogenase (ICDH; EC 1.1.1.42).

M2). This coincides with carbon fixation by activation of several light-dependent enzymes including rubisco (EC 4.1.1.39) at the start of light exposure [31], as shown by the accumulation of Ru15P, Gce and triose derivatives at the beginning of the light period (light 1 – 3 hr). The slow accumulation was partly attributable to the very slow metabolic turnover of rubisco [32]. Likewise, major amino acids and amines including Glu and Gln, the source compounds of nitrogen assimilation as amino-group acceptor/donor [33,34], also accumulated in the first half of the light period (M1). This coincides with the diurnal metabolic dynamics and the activities of key enzymes in tobacco plant [35]. For example, NR activity is known to remarkably increase immediately after the start of light exposure and decrease at midday.

On the other hand, the glycolytic pathway and the reductive pentose phosphate pathway intermediates reached their highest levels (M3, M4) at midday, and sugars peaked at the end of the light period (M7).

We can hypothesize that carbon fixed in the first half of the light period moves down the glycolytic pathway and the TCA cycle, and amino acid biosynthesis progresses using generated Glu, Pyr, and 2OG. In the latter half of the light period, the flow of fixed carbon leads to the accumulation of the intermediates in the pentose phosphate pathway and to sucrose synthesis by inhibiting the production of ammonia, Pyr, and 2OG.

From the end of the light period through the first half of the dark period, we noted an increase in sugar phosphates from the first half of the glycolytic pathway (metabolic module M9). Around midnight, the accumulation of a few organic acids in the first half of the TCA cycle (metabolic module M10) was observed, suggesting the activation of the TCA cycle.

In the latter half of the dark period, the level of minor amino acids was increased (metabolic module M11), although they are synthesized from diverse biochemical pathways. The good correlation among these minor amino acids, also reported in potato and wheat [36], is attributable to the fact that the ratio between Gln and 2OG regulate minor amino acids in bacteria and fungi through the reaction Glu + 2-oxo acid ↔ amino acid + 2OG [37]. Under our experimental conditions, the Glu/2OG ratio was much higher in the dark- than in the light period (22.9 vs. 7.2) and the amino group can easily transferred to 2-oxo acids to produce amino acids.

Adenine nucleoside and nicotinamide coenzyme status

ATP and ADP were placed in the dark-activated group in PLM (metabolic module M12); they were accumulated at the end of the dark period, and decreased by illumination (Figure 3C). On the other hand, AMP was placed in the light-activated group peaking at midday. The reason for fluctuations of adenylate is unknown. Previous observations also do not coincide in the adenylate levels during the light- and dark period. In sugar beet leaves, all adenylate levels increased in the light period [38]. In spinach leaves and wheat leaf protoplast, ATP increased but ADP and AMP decreased under light [39,40]. In Crassulacean-acid metabolism (CAM) species, on the contrary, ATP decreased but ADP and AMP increased [41]. Such differences may result from different dynamics in cytosol, chloroplasts, and mitochondria [40].

We extrapolate that the lower ATP ratio during the light period was caused by an excess demand of ATP by intra- and extra cellular processes for carbon fixation and nitrogen assimilation against ATP supply from photosynthesis. In theory, the amount of ATP consumption in the reductive pentose phosphate pathway and the photorespiratory pathway is more than ATP production in the photophosphorylation [42]. Beside this, nitrogen assimilation process, intracellular transport of the assimilation products, and sucrose synthesis and its translocation are also accompanied by ATP. Therefore the dark respiration makes a considerable contribution to produce ATP even in the light. However, granted that ATP supply is insufficient in the light, high metabolic turnover of adenylate

kinase (EC 2.7.4.3) would immediately work to reproduce ATP from ADP that leads to increase of AMP. Further investigation is necessary to clarify the adenylate dynamics among cell compartments.

In our analysis, NADPH and NADH behaved similarly (metabolic module M6), whereas NADP and NAD did not. As NADPH and NADH were respectively generated by their unique reaction of reducing NADP and NAD, dependence on the intracellular oxidation-reduction state shifted the formation of oxidation and reduction. In PLM, however, NADP was placed in the light-activated- and NAD in the dark-activated group. This suggests that highly concentrated NAD in the dark is converted to NADPH via NADP in the light period. It was reported that the NADPH/NAD ratio is the inverse of the ATP/ADP ratio in guard cell protoplast, which indicates that ATP phosphorylates NAD in the light period by NAD kinase (EC 2.7.1.23) and the generated NADP is reduced to NADPH in the course of photosynthesis [43].

The ratios of NADH to NAD and NADPH to NADP were 0.16–0.29 and 6.2–6.6. The observed difference in the tendency of oxidized- or reduced form indicates their different cellular roles. NADH is used for oxidative phosphorylation, and a low NADH/NAD ratio constrains this process. On the other hand, NADPH is used for the reductive biosynthesis of metabolites, and the high ratio of NADPH/NADP favors the reduction of metabolites.

Environmental stress response

It is remarkable that Glt (GSH; gamma-glutamylcysteinyl glycine) and Spe exhibited similar fluctuation patterns (metabolic module M8). Both peaked at the end of the light period and again just after midnight, suggesting the existence of common regulatory factors. GSH plays a central role in the antioxidant defense by eliminating harmful peroxide during photosynthesis and oxidative phosphorylation [44]. Polyamines, including spermidine, are also effective antioxidants under various environmental stress conditions [45]. During photosynthesis, GSH is converted to oxidized dithiol (GSSH) to eliminate oxidative stress, and upon the reduction of NADPH, GSSH can be converted back to GSH by glutathione reductase (GR; EC 1.8.1.7, annotated in rice plant). Our finding that NADPH reached its highest level at a few hours before the end of the light period is consistent with the above observation (Figure 3C), although the connection remains speculative. The relative contribution of NADPH and NADH to the generation of GSH and spermidine requires further investigation.

Conclusion

We intended to analyze the rice plant metabolism and to reconstruct its phenotypic networks in an effort to explain underlying biological functions. Our CE-MS technology provided a comprehensive high-throughput system with easy sample preparation and facilitated the generation of high-resolution metabolic time-courses. Data mining with statistical techniques and SOM analysis revealed synchronous dynamics in metabolic modules downstream of C and N assimilation and dissimilation processes and stress responses. Our system was able to discriminate unidentified metabolites and identify bottleneck enzymatic steps. In a comprehensive approach such heuristics become increasingly important because with current technology, the determination of all network components is virtually impossible. For a more precise investigation of biochemical networks, expansion of target metabolites and determination of metabolite levels in each cellular compartment may be suggested. There are technical hurdles, however, in separating organelles without disturbing a wide range of metabolites inside them. Without much technical advancement, therefore, it seems difficult to repeat our time-course measurement for any single cellular compartment although there are reports for such a challenge [46]. Finally, for the analysis part, it is necessary to couple biological information with computer simulations based on large-scale time-resolved measurements of metabolites, proteins, and mRNAs.

Methods

Plant materials

Young seedlings of rice plants, *Oryza sativa* L. ssp. *japonica* Haenuki, at the third-leaf stage were cultured as follows. Rice seeds were germinated on filter paper soaked with Milli-Q water and kept at 30 °C in a dark room for 2 days. After germination, the plantlets were placed on rock fiber (35 × 35 × 40 mm; Nittobo, Tokyo, Japan), and grown in a growth chamber (FLI-301N, Tokyo Rika Kikai, Tokyo, Japan) for 18 days. The temperature and light conditions were 25 °C and 365 $\mu E \cdot m^{-2}s^{-1}$ for 9 hr (light), 20 °C and 0 $\mu E \cdot m^{-2}s^{-1}$ for 11 hr (dark), and 150 $\mu E \cdot m^{-2}s^{-1}$ for 2 hr between light and dark. The plants were watered with Kasugai water culture solution (18.9 mg/L $(NH_4)_2SO_4$, 10.1 mg/L $Na_2HPO_4 \cdot 12H_2O$, 4.7 mg/L KCl, 0.79 mg/L $CaCl_2$, 3.0 mg/L $MgCl_2$, 0.17 mg/L $\cdot FeCl_3 \cdot 6H_2O$, and HCl to adjust the pH to 5.0 – 5.5) [47].

Reagents

Piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES) was purchased from Dojindo (Kumamoto, Japan), methionine sulphone from Avocado Research (Heysham, Lancashire, UK). All other reagents were obtained from conventional commercial sources. Individual stock solutions, at a concentration of 10 or 100 mM, were prepared in Milli-Q water, 0.1 N HCl, or 0.1 N NaOH. The working standard mixture was prepared by diluting these stock solutions with Milli-Q water just before injection. All chemicals used were of analytical or reagent grade. Water

was purified with a Milli-Q purification system (Millipore, Bedford, MA, USA).

Sample preparation

Leaves were harvested (fresh weight approximately 100 mg (6 seedlings)) and frozen in liquid nitrogen to stop enzymatic activity. They were mashed in a Multi-Beads Shocker (Yasuikikai, Osaka, Japan) at 2000 rpm for 10 sec and 0.5 mL of ice-cooled methanol, including 400 μM PIPES and methionine sulphone as an internal standard, was added to dissolve phospholipid membranes and inactive enzymes. Then 0.5 mL ice-cold Milli-Q water was added and the sample was ultrafiltered through a 5-kDa cut-off filter at 9058 g for 10 min to remove proteins, phospholipids, chlorophyll, and other high-molecular-weight impurities. The filtrate was analyzed by CE-MS and CE-DAD methods. To obtain sufficient sensitivity for the analysis of nucleotides, coenzymes, and sugars, the filtrate was concentrated 5-fold by lyophilization [17].

Instruments

All CE-MS experiments were performed by Agilent CE capillary electrophoresis. We used a 1100 series MSD mass spectrometer, a 1100 series isocratic HPLC pump, a G1603A CE-MS adapter kit, and a G1607A CE-ESI-MS sprayer kit (Agilent Technologies). CE-DAD experiments were performed by Agilent CE capillary electrophoresis with a built-in diode-array detector. G2201AA Agilent ChemStation software for CE was used for system control, data acquisition and analysis, and MSD data evaluation.

Analytical conditions

The compounds were analyzed in four groups using three CE-MS methods and one CE-DAD method.

a) Cationic metabolites (amino acids and amines) were analyzed with a fused-silica capillary (50 μm i.d. × 100 cm total length), with 1 M formic acid as the electrolyte. The sample was injected at an injection pressure of 5.0 kPa for 3 sec (approximately 3 nL). The applied voltage was set at 30 kV. The capillary temperature was set to 20°C, and the sample tray was cooled to below 5°C. The sheath liquid (5 mM ammonium acetate in 50% [v/v] methanol-water) was delivered at 10 μL/min. ESI-MS was conducted in positive ion mode; the capillary voltage was set at 4000 V. A flow rate of heated dry nitrogen gas (heater temperature 300°C) was maintained at 10 L/min [12].

b) Anionic metabolites (organic acids and sugar phosphates) were analyzed with a cationic polymer-coated SMILE(+) capillary (Nakalai Tesque, Kyoto, Japan). The electrolyte for CE separation was a 50 mM ammonium acetate solution (pH 8.5). The sample was injected at an injection pressure of 5.0 kPa for 30 sec (approximately 30 nL). The applied voltage was set at -30 kV, and the capillary temperature was set to 30°C. ESI-MS was conducted in negative ion mode; the capillary voltage was set at 3500 V. Other conditions were as in the cationic metabolite analysis [13].

c) Nucleotides and coenzymes were analyzed with an uncharged polymer-coated gas chromatograph capillary, polydimethylsiloxane (DB-1) (Agilent Technologies). The electrolyte for CE separation was 50 mM ammonium acetate solution (pH 7.5). The applied voltage was set at -30 kV and a pressure of 5.0 kPa was added to the inlet capillary during the run. Other conditions were as in the anion analysis [14].

d) Sugars were analyzed with a fused-silica capillary (50 μm i.d. × 112.5 cm total length, 104 cm effective length). Basic anion buffer for CE (Agilent Technologies) was the electrolyte. The sample was injected at a pressure of 5.0 kPa for 10 sec (approximately 10 nL). The applied voltage was set at -25 kV; the capillary temperature, regulated with a thermostat, was 25°C. Sugars were detected by indirect UV detection using a diode-array detector. The signal wavelength was set at 350 nm with a reference at 230 nm [48].

Self-organizing map (SOM) analysis

A free software package, SOM -PAK [49], was used to compute both the SOM and the Sammon map. Before SOM analysis, the observed time-course data for 58 metabolites (including an estimate of S17P) were smoothed by averaging the adjacent data points using a sliding window of width 3, to reduce high-frequency noise presumably originating from individual differences in plant seedlings, rapid oscillations in metabolism, or measurement errors. The missing data points were extrapolated by linear approximation between prior and subsequent data values. Among the 57 metabolites evaluated at 26 time points, only 30 data points could be extrapolated due to the detection limit or contamination of other unidentifiable peaks. The SOM is a map from the input n-dimensional data space (input layer) to a two-dimensional array of nodes (output layer). The vectors in the output layer are the parametric reference vector $\boldsymbol{m}_i$, which has n elements. An input data vector, $\boldsymbol{x}$, is compared with $\boldsymbol{m}_i$, and the best-match vector, which is the smallest Euclidean distance $|\boldsymbol{x} - \boldsymbol{m}_i|$, is mapped onto this location. During learning, nodes that are topographically close in the array up to a certain distance activate each other to learn from the same input vector, and the reference vectors are corrected so that they become close to the input vector. Thus,

$$\boldsymbol{m}_i(t + 1) = \boldsymbol{m}_i(t) + h_{ci}(t)\,[\boldsymbol{x}(t) - \boldsymbol{m}_i(t)],$$

where t is an integer, the discrete-time coordinate, and $h_{ci}(t)$ is the neighborhood kernel, a function defined over

the lattice points. The neighborhood size, N_c, around node c is a function of time, and h_{ci} is defined as

$$h_{ci} = \alpha(t) \quad (i \in N_c)$$
$$h_{ci} = 0 \quad (i \notin N_c),$$

where $\alpha(t)$ is a monotonic decreasing function of time ($0 < \alpha(t) < 1$) called the "learning rate". The learning rate function was defined as

$$\alpha(t) = \alpha(0)(1.0 - t/T),$$

where $\alpha(0)$ is the initial learning rate and T the running length (number of steps) in training. In this study, 58 metabolic time-courses were formatted and classified in a 24 × 24 hexagonal lattice. The applied SOM parameters were: initial radius of the training area = 12, initial learning rate = 0.025, running length = 65 000.

Metabolic pair-wise correlation

Significance levels for Pearson correlation coefficient r were computed depending on the number of metabolite pairs n found throughout the light and dark period, respectively, by calculating t-scores given by $t = r\,(n - 2)^{0.5}/(1 - r)^{0.5}$. The critical t-score was set to correspond to the commonly used p-value of 0.05 in two-sided tests.

Hierarchical clustering

Among several algorithms for clustering analysis, we chose Ward's method [29] in JMP software (ver. 6.0.0; SAS Institute Inc. Cary, NC). Starting from trivial clusters each containing one object only, Ward's method iteratively merges two clusters that will result in the smallest increase in the sum of the square of their differences (i.e., variance). At each step, all possible mergers of two clusters are tried and their variance is computed. The difference between clusters is calculated by the equation:

$$d(a,b) = \frac{n_a n_b}{n_a + n_b}(x_a - x_b)^2$$

Authors' contributions

SS conceived this study, performed the biochemical- and the computational experiments, and wrote the manuscript. MA provided intellectual help for the computational analysis and together wrote the manuscript. TN advised the experimental design. TS and MT supervised the research. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by a grant for the Development of Rice Genome Simulators from MAFF, Japan, and by the Ministry of Education, Culture, Sports, Science and Technology, and a Grant-in-Aid for the 21st Century Center of Excellence (COE) Program entitled "Understanding and Control of Life's Function via Systems Biology (Keio University)". This work was also supported, in part, by Grant-in-Aid for Scientific Research on Priority Areas "Systems Genomics" from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. Fridman E, Pichersky E: **Metabolomics, genomics, proteomics, and the identification of enzymes and their substrates and products.** *Curr Opin Plant Biol* 2005, **8(3)**:242-248.
2. Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, Dam K, Oliver SG: **A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations.** *Nat Biotechnol* 2001, **19(1)**:45-50.
3. Allen J, Davey HM, Broadhurst D, Heald JK, Rowland JJ, Oliver SG, Kell DB: **High-throughput classification of yeast mutants for functional genomics using metabolic footprinting.** *Nat Biotechnol* 2003, **21(6)**:692-696.
4. Morgenthal K, Wienkoop S, Scholz M, Selbig J, Weckwerth W: **Correlative GC-TOF-MS-based metabolite profiling and LC-MS-based protein profiling reveal time-related systematic regulation of metabolite-protein networks and improve pattern recognition for multiple biomarker selection.** *Metabolomics* 2005, **1(2)**:109-121.
5. Weckwerth W, Loureiro ME, Wenzel K, Fiehn O: **Differential metabolic networks unravel the effects of silent plant phenotypes.** *Proc Natl Acad Sci USA* 2004, **101(20)**:7809-7814.
6. Hirai MY, Klein M, Fujikawa Y, Yano Y, Goodenowe DB, Yamazaki Y, Kanaya S, Nakamura Y, Kitayama M, Suzuki H, Sakurai N, Shibata D, Tokuhisa J, Reichelt M, Gershenzon J, Papenbrock J, Saito K: **Elucidation of gene-to-gene and metabolite-to-gene networks in arabidopsis by integration of metabolomics and transcriptomics.** *J Biol Chem* 2005, **280(27)**:25590-25595.
7. Kohonen T: *Self-Organizing Maps* Springer-Verlag, Heidelberg, Germany; 1995.
8. Mounet F, Lemaire-Chamley M, Maucourt M, Cabasson C, Giraudel JL, Deborde C, Lessire R, Gallusci P, Bertrand A, Gaudillère M, Rothan C, Rolin D, Moing A: **Quantitative metabolic profiles of tomato flesh and seeds during fruit development: Complementary analysis with ANN and PCA.** *Metabolomics* 2007, **3(3)**:273-288.
9. Panagiotou G, Kouskoumvekaki I, Jónsdóttir , Olsson L: **Monitoring novel metabolic pathways using metabolomics and machine learning: Induction of the phosphoketolase pathway in *Aspergillus nidulans* cultivations.** *Metabolomics* 2007, **3(4)**:503-516.
10. Fiehn O: **Metabolic networks of *Cucurbita maxima* phloem.** *Phytochemistry* 2003, **62(6)**:875-886.
11. Yeung KY, Medvedovic M, Bumgarner RE: **From co-expression to co-regulation: How many microarray experiments do we need?** *Genome Biol* 2004, **5(7)**:R48.
12. Soga T, Heiger DN: **Amino acid analysis by capillary electrophoresis electrospray ionization mass spectrometry.** *Anal Chem* 2000, **72**:1236-1241.
13. Soga T, Ueno Y, Naraoka H, Ohashi Y, Tomita M, Nishioka T: **Simultaneous determination of anionic intermediates for *Bacillus subtilis* metabolic pathway by capillary electrophoresis electrospray ionization mass spectrometry.** *Anal Chem* 2002, **74**:2233-2239.
14. Soga T, Ueno Y, Naraoka H, Matsuda K, Tomita M, Nishioka T: **Pressure-assisted capillary electrophoresis electrospray ionization mass spectrometry for analysis of multivalent anions.** *Anal Chem* 2002, **74**:6224-6229.
15. Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T: **Quantitative metabolome analysis using capillary electrophoresis mass spectrometry.** *J Proteome Res* 2003, **2**:488-494.
16. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Ho PY, Kakazu Y, Sugawara K, Igarashi S, Harada S, Masuda T, Sugiyama N, Togashi T, Hasegawa M, Takai Y, Yugi K, Arakawa K, Iwata N, Toya Y, Nakayama Y, Nishioka T, Shimizu K, Mori H, Tomita M: **Multiple high-throughput analyses monitor the response of E. coli to perturbations.** *Science* 2007, **316(5824)**:593-7.
17. Sato S, Soga T, Nishioka T, Tomita M: **Simultaneous determination of the main metabolites in rice leaves using capillary**

electrophoresis mass spectrometry and capillary electrophoresis diode array detection. *The Plant J* 2004, **40**:151-163.
18. Sammon JW Jr: **A nonlinear mapping for data structure analysis.** *IEEE Transactions Computers* 1969, **C-18(5)**:401-409.
19. **KEGG pathway database** [http://www.genome.ad.jp/kegg/pathway.html]
20. **Swiss-Prot database** [http://au.expasy.org/sprot/]
21. **Rice Annotation Project Data Base** [http://rapdb.dna.affrc.go.jp/]
22. Stafford HA, Magaldi A, Vennesland B: **The enzymatic reduction of hydroxypyruvic acid to $_D$-glyceric acid in higher plants.** *J Biol Chem* 1954, **207**:621-629.
23. Rippert P, Matringe M: **Purification and kinetic analysis of the two recombinant arogenate dehydrogenase isoforms of *Arabidopsis thaliana*.** *Eur J Biochem* 2002, **269**:4753-4761.
24. Boldt R, Edner C, Kolukisaoglu Ü, Hagemann M, Weckwerth W, Wienkoop S, Morgenthal K, Bauwe H: **$_D$-Glycerate 3-kinase, the last unknown enzyme in the photorespiratory cycle in Arabidopsis, belongs to a novel kinase family.** *The Plant Cell* 2005, **17**:2413-2420.
25. Duncan K, Edwards RM, Coggins JR: **The pentafunctional *arom* enzyme of *Saccharomyces cerevisiae* is a monofunctional domains.** *Biochem J* 1987, **246**:375-386.
26. Kuhr WG: **Separation of small organic molecules.** In *Capillary Electrophoresis: Theory and Practice* Edited by: Camilleri P. CRC Press LLC, Boca Raton, FL, USA; 1997:91-133.
27. **KEGG ligand database** [http://www.genome.ad.jp/kegg/ligand.html]
28. Steuer R, Kurths J, Fiehn O, Weckwerth W: **Observing and interpreting correlations in metabolic networks.** *Bioinformatics* 2003, **19(8)**:1019-1026.
29. Ward JH: **Hierarchical grouping to optimize an objective function.** *J Am Stat Assoc* 1963, **58**:236-245.
30. Ashihara H, Sato F: **Pyrophosphate: Fructose-6-phosphate 1-phosphotransferase and biosynthetic capacity during differentiation of hypocotyls of Vigna seedlings.** *Biochim Biophys Acta* 1993, **1156**:123-127.
31. Farr TJ, Huppe HC, Turpin DH: **Coordination of chloroplastic metabolism in N-limited chlamydomonas reinhardtii by redox modulation (I. The activation of phosphoribulokinase and glucose-6-phosphate dehydrogenase is relative to the photosynthetic supply of electrons).** *Plant Physiol* 1994, **105**:1037-1042.
32. Woodrow IE, Berry J: **Enzymatic regulation of photosynthetic CO_2 fixation in C_3 plants.** *Ann Rev Plant Physiol Plant Molec Biol* 1988, **39**:533-594.
33. Lea PJ, Ireland RJ: **Nitrogen metabolism in higher plants.** In *Plant Amino Acids, Biochemistry and Biotechnology* Edited by: Singh BK. New York: Marcel Dekker, Inc; 1999:1-47.
34. Ireland RJ, Lea PJ: **The enzymes of glutamine, glutamate, asparagines, and aspartate metabolism.** In *Plant Amino Acids, Biochemistry and Biotechnology* Edited by: Singh BK. New York: Marcel Dekker, Inc; 1999:49-109.
35. Scheible WR, Krapp A, Stitt M: **Reciprocal diurnal changes of phosphoenol pyruvate carboxylase expression and cytosolic pyruvate kinase, citrate synthase and NADP-isocitrate dehydrogenase expression regulate organic acid metabolism during nitrate assimilation in tobacco leaves.** *Plant Cell and Environ* 2000, **23**:1155-1167.
36. Noctor G, Novitskaya L, Lea PJ, Foyer CH: **Co-ordination of leaf minor amino acid contents in crop species: Significance and interpretation.** *J Exp Bot* 2002, **53(370)**:939-945.
37. Ferrario-Méry S, Suzuki A, Kunz C, Valadier MH, Roux Y, Hirel B, Foyer CH: **Modulation of amino acid metabolism in transformed tobacco plants deficient in Fd-GOGAT.** *Plant and Soil* 2000, **221**:67-79.
38. Rao IM, Arulanantham AR, Terry N: **Diurnal changes in adenylates and nicotinamide nucleotides in sugar beet leaves.** *Photosynthesis Res* 1990, **23**:205-212.
39. Bonzon M, Hug M, Wagner E, Greppin H: **Adenine nucleotides and energy charge evolution during the induction of flowering in spinach leaves.** *Planta* 1981, **152**:189-194.
40. Stitt M, Lilley RM, Heldt HW: **Adenine nucleotide levels in the cytosol, chloroplasts, and mitochondria of wheat leaf protplasts.** *Plant Physiol* 1982, **70**:971-977.
41. Chen LS, Nose A: **Day-Night changes of energy-rich compounds in crassulacean acid metabolism (CAM) species utilizing hexose and starch.** *Ann Bot* 2004, **94**:449-455.
42. Leegood RC: **Photosynthesis in C_3 plants: The Benson-Calvin cycle and photorespiration.** In *Plant Biochemistry and Molecular Biology* 2nd edition. Edited by: Lea PJ, Leegood RC. John Wiley & Sons Ltd; 1999:29-50.
43. Hampp R, Schnabl H: **Adenine and pyridine nucleotide status of isolated Vicia guard cell protoplasts during K+-induced swelling.** *Plant and Cell Physiol* 1984, **25(7)**:1233-1239.
44. May MJ, Vernoux T, Leaver C, Montagu MV, Inze D: **Glutathione homeostasis in plants: Implications for environmental sensing and plant development.** *J Exp Bot* 1998, **49(321)**:649-667.
45. Lovaas E: **Antioxidant and metal-chelating effects of polyamines.** In *Advances in Pharmacology. Antioxidants in Disease Mechanisms and Therapy 38* Edited by: Sies H. New York: Academic Press; 1996:119-149.
46. Farré EM, Tiessen A, Roessner U, Geigenberger P, Trethewey RN, Willmitzer L: **Analysis of the compartmentation of glycolytic intermediates, nucleotides, sugars, organic acids, amino acids, and sugar alcohols in potato tubers using a nonaqueous fractionation method.** *Plant Physiol* 2001, **127**:685-700.
47. Kasugai S: **Studies of water culture.** *Jpn J Soil Sci Plant Nutr* 1939, **13**:669-822. (in Japanese)
48. Soga T, Ross GA: **Simultaneous determination of inorganic anions, organic acids, amino acids and carbohydrate by capillary electrophoresis.** *J Chromatogr A* 1999, **837**:231-239.
49. **SOM-PAK** [http://www.cis.hut.fi/research/som_lvq_pak]

Depiction of metabolome changes in histidine-starved *Escherichia coli* by CE-TOFMS†

Yoshiaki Ohashi,[ab] Akiyoshi Hirayama,[b] Takamasa Ishikawa,[a] Seira Nakamura,[a] Kaori Shimizu,[b] Yuki Ueno,[a] Masaru Tomita[ab] and Tomoyoshi Soga*[ab]

Received 13th September 2007, Accepted 25th October 2007
First published as an Advance Article on the web 15th November 2007
DOI: 10.1039/b714176a

Metabolic changes in response to histidine starvation were observed in histidine-auxotrophic *Escherichia coli* using a capillary electrophoresis time-of-flight mass spectrometry (CE-TOFMS)-based metabolomics technique. Prior to the analysis, we prepared an *E. coli* metabolome list of 727 metabolites reported in the literature. An improved method for metabolite extraction was developed, which resulted in higher extraction efficiency in phosphate-rich metabolites, *e.g.*, ATP and GTP. Based on the results, 375 charged, hydrophilic intermediates in primary metabolisms were analysed simultaneously, providing quantitative data of 198 metabolites. We confirmed that the intracellular levels of intermediates in histidine biosynthesis are rapidly accumulated in response to a drop in histidine level under histidine-starved conditions. Simultaneously, disciplined responses were observed in the glycolysis, tricarboxylic acid cycle, and amino acid and nucleotide biosynthesis pathways as regulated by amino acid starvation.

Introduction

Metabolomics, the unbiased determination of metabolite levels, is expected to be a valuable approach for the characterisation of bioprocesses in combination with genomics, transcriptomics, and proteomics. Different from the other "omics", metabolomics includes methodological problems derived from heterogeneity in chemical properties, that is, it involves the development of extraction methods of metabolites that can be applied to a broad range of chemical species, and analytical methods that are achieved by as few processes as possible. However, it is impossible for all metabolites to be detected by single analytical method at present.

The chemical properties of metabolites are widely diverse, but the majority of them have a few common properties in terms of physico-chemical traits: molecular weight less than 500, high hydrophilicity,[1] and charge in aqueous solution.[2] On the basis of these properties, we have developed a metabolome analysis method using capillary electrophoresis electrospray ionization mass spectrometry (CE-MS) to determine the majority of metabolic intermediates.[3–7] We found 150 metabolites in sporulating *Bacillus subtilis*, and analysed the changes in profiles in major energy metabolisms.[6] However, it is unclear how many metabolites in bacterial cells can be determined by our analytical system. Extraction of metabolites from cells is also one of the most dominant factors in metabolomics. Extraction must be adapted to the target metabolites, organisms, and analytical methods employed. To date, many extraction methods have been comparatively studied, with cold methanol extraction,[8,9] and acidic acetonitrile extraction[10] as the recommended techniques. We have employed a quenching method by pure methanol followed by methanol-water-chloroform extraction to remove hydrophobic metabolites.[6] While this method satisfies particular constraints in CE-MS analysis, the efficiency and data reproducibility is not sufficient especially in phosphate-rich metabolites, *e.g.*, ATP. In this study, we modified the method to improve these problems, and applied it to determine the intermediates of histidine biosynthesis in *Escherichia coli* histidine-auxotroph.

The stringent response is known to be a starvation-stimulated adaptation mechanism which includes functional arrests of chromosome replication, cell division, transcription, translation, and metabolisms.[11] Translation arrest induced by amino acid starvation on ribosomes activates RelA protein, resulting in temporal accumulation of a bacterial alarmone, guanosine 3′-diphosphate 5′-diphosphate (ppGpp). Recently, transcriptome and proteome analyses have been demonstrated in bacterial stringent responses, indicating the possibility that metabolisms are dramatically influenced by amino acid starvation.[12–14] However, no metabolomics approach has been used, raising a question of how to change the metabolism in the starved, auxotrophic situation in bacteria.

In this study, we prepared a complete metabolite list for *E. coli* by searching databases and the literature. The method of metabolome extraction was also improved, realising good metabolome analyses using capillary electrophoresis electrospray ionization time-of-flight mass spectrometry (CE-TOFMS). We applied the method to the analysis of the *E. coli* histidine-auxotroph and demonstrated through the analysis of 375 metabolites including primary metabolism intermediates that the metabolome profile of *E. coli* is

[a]*Human Metabolome Technologies, Inc., Tsuruoka, Yamagata 997-0052, Japan*
[b]*Institute for Advanced Biosciences, Keio University, 246-2 Mizukami, Kakuganji, Tsuruoka, Yamagata 997-0052, Japan.*
E-mail: soga@sfc.keio.ac.jp; Fax: +81-235-29-0534; Tel: +81-235-29-0528

† Electronic supplementary information (ESI) available: Supplementary tables and high quality figures. See DOI: 10.1039/b714176a

dramatically changed during amino acid starvation. The quantitative data obtained for 198 hydrophilic, charged metabolites provides new insights in bacterial starvation.

Results and discussion

Target metabolites of CE-MS based metabolomics

The definition of the metabolome that is widely accepted in various fields is the entire set of low-molecular weight intermediates in metabolisms including amino acids, amines, nucleotides, sugars, lipids, and other substances. Generally, DNA, RNA, and protein are excluded from the definition, but their digests are occasionally recognised to be the targets of metabolomics. These metabolites are widely diverse in their chemical properties, *i.e.*, polarity that is a determinant of solubility, electronic charge of ions, volatility, and molecular weight. Metabolites found in *E. coli* cells are summarised in the EcoCyc database,[15,16] and their physico-chemical properties have been previously observed.[1] Prior to the metabolome analysis, we prepared a list of *E. coli* metabolites to understand the target metabolites in CE-TOFMS based metabolomics. Our survey of the known *E. coli* metabolites, carried out through a database and literature search (see Experimental), resulted in a list of metabolome including 727 metabolites (supplementary Table 1†). Of these, 453 (62%) belong to primary metabolites and their degradation intermediates (classes I and II), which are expected to be found in cells grown in minimal medium (Fig. 1A). Since most of the primary metabolites, for which the number is calculated as 92% (hydrophilic in Fig. 1B) × 96% (anion, cation, and nucleotide in Fig. 1C) = 88% (402 metabolites), are hydrophilic and charged, it is expected that CE-TOFMS, which principally targets water-soluble, charged metabolites is an extremely promising analysis method for metabolomics in primary pathways. Since some of the metabolites (*e.g.*, acetic acid) cannot be detected rationally, the number of the target metabolites in CE-TOFMS analysis was 375, which corresponds to 83% of the primary metabolites in *E. coli*.

Optimisation of metabolome extraction procedure

A problem was inherent in our previous studies regarding bacterial metabolome analyses.[5,6] The extraction efficiencies of phosphate-rich metabolites such as ATP were often relatively low, affecting reproducibility of the analyses. We considered that the problem may be due to the interaction between phosphate groups of metabolites and phospholipids in the cell membrane. To improve this problem, we introduced sonication treatment during methanol extraction of metabolites from the cells (Fig. 2 and Experimental section). By this additional process, cells are peeled from the filter and completely suspended in methanol. Here, an ultrasonic syringe is suitable rather than a cell disruptor, since the methanol is transpired by excessively strong ultrasonic of cell disruptor, resulting in fluctuation and enhanced error of the metabolomics data (data not shown). To compare the extraction efficiencies, we used four methods for the methanol extraction process; (i) with sonication (improved method), (ii) without sonication (conventional method), (iii) with sonication, but cells are removed by brief centrifugation before chloroform addition (sonicated cell removal method), and (iv) with incubation at −80 °C for 12 h, but cells are removed (cold method).

To examine the reproducibility of the improved method, we repeatedly collected the data obtained in the improved method (the data are distinguished as "1st" and "2nd"). Fig. 3A shows the reproducibility and improvement of extraction efficiency by sonication in nucleotide metabolites containing phosphate groups. The metabolome source was *E. coli* W3110 grown in rich medium at 37 °C until the middle-logarithmic growth phase. Strikingly, extraction efficiency of GTP was improved 130-fold relative to the conventional method reported previously.[6] For adenosine-phosphates, ATP, ADP, and AMP, the signal intensities were enhanced 39-, 17-, and 1.6-fold, respectively, by the improved method. Enhancement of detected signal intensities depended on the number of phosphate groups in the metabolites, indicating that efficient extraction of phosphate-rich metabolites is difficult by methanol extraction alone. It has been assumed that the phenomenon is due primarily to interaction of the metabolites with cell membranes by the intense negative charge of phosphate, but further investigations are needed. The improvement of extraction efficiency was not achieved by the sonicated cell removal method or cold method (Fig. 3A). This result suggests that the cells (and enzymes in the cells) fixed in methanol are damaged or inverted by contact with chloroform, and release the metabolites including phosphate groups into the solvent. The extraction efficiency of most amino acids was independent of the extraction methods except for lysine and arginine (Fig. 3B). This suggests that extraction of basic metabolites is also conditional in methanol extraction. Another basic amino acid, histidine, was efficiently extracted by all methods, possibly due to its more feeble imidazole group charge. As the isoelectric points of arginine, lysine, and histidine are 10.76, 9.74, and 7.59, respectively, the extremely intense positive charges of arginine and lysine most likely interact with the phosphate group in the phospholipids of the cell membrane. On the other hand, a problem was raised in the improved method. In some metabolites, NADH, NADPH, and coenzyme A (CoA), the extraction efficiency was worsened relative to that of the conventional method (Fig. 3A). Levels of oxidised forms of these metabolites, NAD^+, $NADP^+$, and CoA dimer (CoA–CoA) were increased in the improved method. This was not found in the other methods. Total amounts of these oxidised and reduced forms of metabolites were enhanced in the improved method (data not shown), suggesting that the extraction of them is basically improved. However, decreased levels of the oxidised forms most likely suggest that the reduced forms of the metabolites are oxidised during preparation procedures. Although it is not clear why oxidation of the metabolites is accelerated in the improved method, enhanced levels of some oxidants, *e.g.*, flavin nucleotides may stimulate the oxidation.

Determination of histidine-biosynthesis intermediates using improved methods

Histidine is synthesised *via* ten steps which are evolutionarily conserved in all organisms that synthesise histidine.[17]

Table 1 List of metabolites clustered by HCA

Cluster No.	Pathway	Class[a]	Metabolite	Abbreviation	KEGG ID	Analytical mode
1	Arginine biosynthesis	I	*N*-Acetyl-L-glutamic acid	N-AcGlu	C00624	Cation
1	Cysteine biosynthesis	I	*O*-Acetyl-L-serine	AcSer	C00979	Cation
1	Leucine biosynthesis	I	3-Isopropylmalic acid	3IPMA	C04411	Anion
1	Methionine biosynthesis	I	Homocysteine	HomoCys	C05330	Cation
1	Pyruvic acid oxidation	I	Acetyl-CoA	AcCoA	C00024	Nucleotide
1	Salvage of Ade, HypXan	I	Adenine	—	C00147	Cation
1	Serine biosynthesis	I	3-Phosphoserine	3PSer	C01005	Anion
2	Arginine biosynthesis	I	L-Citrulline	Citrulline	C00327	Cation
2	Arginine biosynthesis	I	L-Arginosuccinic acid	ArgSuccinate	C03406	Cation
2	Biotin biosynthesis	I	7,8-Diaminononanoic acid	7,8 DANA	C01037	Cation
2	Cholanic acid biosynthesis	I	α-D-Galactose-1-phosphate	Gal1P	C00446	Anion
2	Cholanic acid biosynthesis	I	UDP-galactose	UDP-Gal	C00052	Nucleotide
2	Chorismic acid biosynthesis	I	3-Deoxy-D-arabinoheptulosonate-7-phosphate	3DAH7P	C04691	Anion
2	Chorismic acid biosynthesis	I	Shikimate-3-phosphate	S3P	C03175	Anion
2	dTDP-rhamnose biosynthesis	I	dTDP-D-glucose	dTDP-Glc	C00842	Nucleotide
2	Fravin biosynthesis	I	5-Amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione	5A6RbA2,4PD	C04732	Cation
2	Glutamic acid biosynthesis	I	L-Glutamic acid	Glu	C00025	Cation
2	Glutathione biosynthesis	I	L- γ-Glutamylcysteine	GluCys	C00669	Cation
2	Glycolysis(Embden-Meyerhof-Parnas)	I	α-D-Glucose-6-phosphate	G6P	C00092	Anion
2	Glycolysis (Embden-Meyerhof-Parnas)	I	D-Fructose-1,6-diphosphate	F1,6P	C00354	Anion
2	Glycolysis (Embden-Meyerhof-Parnas)	I	D-Glyceraldehyde-3-phosphate	G3P	C00661	Anion
2	Glycolysis (Embden-Meyerhof-Parnas)	I	Dihydroxyacetonephosphate	DHAP	C00111	Anion
2	Inorganic metabolite	I	Pyrophosphate	Ppi	C00013	Anion
2	Isopentenyldiphosphate biosynthesis	I	1-Hydroxy-2-methyl-2(*E*)-butenyl-4-diphosphate	HMeBuPP	C11811	Anion
2	Leucine biosynthesis	I	2-Ketoisocaproic acid	2KICA	C00233	Anion
2	Leucine biosynthesis	I	L-Leucine	Leu	C00123	Cation
2	Lysine biosynthesis	I	L,L-Diaminopimelic acid	DPA	C00666	Anion
2	Lysine biosynthesis	I	L-Lysine	Lys	C00047	Cation
2	Methionine biosynthesis	I	*O*-Succinyl-L-homoserine	SucHomoSer	C01118	Cation
2	Methionine biosynthesis	I	L-Methionine	Met	C00073	Cation
2	Methionine oxidation	I	L-Methionine sulfoxide	MetSFX	C02989	Cation
2	NAD biosynthesis	I	Deamido-NAD	—	C00857	Nucleotide
2	NAD biosynthesis	I	Nicotinamideadenine dinucleotidephosphate (reduced form)	NADP+	C00006	Nucleotide
2	NAD biosynthesis	I	Nicotinamideadenine dinucleotidephosphate (oxidized form)	NADPH	C00005	Nucleotide
2	Pentose phosphate pathway (non-oxidative)	I	D-Sedoheptulose-7-phosphate	S7P	C05382	Anion
2	Pentose phosphate pathway (non-oxidative)	I	D-Erythrose-4-phosphate	E4P	C00279	Anion
2	Pentose phosphate pathway (non-oxidative)	I	D-Ribose-5-phosphate	R5P	C00117	Anion
2	Pentose phosphate pathway (oxidative)	I	6-Phospho-D-gluconic acid	6-PG	C00345	Anion
2	Phenylalanine biosynthesis	I	Phenylpyruvic acid	—	C00166	Anion
2	Polyamine biosynthesis	I	Agmatine	—	C00179	Cation
2	Polyamine biosynthesis	I	*S*-Adenosylmethioninamine	dAdoMet	C01137	Cation
2	ppGpp biosynthesis	I	Guanosine-5′-diphosphate 3′-diphosphate	ppGpp	C01228	Nucleotide
2	Proline biosynthesis	I	L-Proline	Pro	C00148	Cation
2	PRPP biosynthesis	I	5-Phosphoribosyldiphosphate	PRPP	C00119	Nucleotide
2	PRPP biosynthesis	I	D-Ribose-1,5-diphosphate	R1.5P	C01151	Nucleotide
2	Purine biosynthesis	I	5′-Phosphoribosyl-*N*-formylglycineamide	FGAR	C04376	Nucleotide
2	Purine biosynthesis	I	Adenosine-5′-diphosphate	ADP	C00008	Nucleotide
2	Purine biosynthesis	I	Adenosine-5′-triphosphate	ATP	C00002	Nucleotide

Table 1 List of metabolites clustered by HCA (*Continued*)

Cluster No.	Pathway	Class[a]	Metabolite	Abbreviation	KEGG ID	Analytical mode
2	Purine biosynthesis	I	Xanthosine-5-phosphate	XMP	C00655	Anion
2	Purine biosynthesis	I	Deoxyadenine-5′-triphosphate	dATP	C00131	Nucleotide
2	Purine biosynthesis	I	Guanosine-5′-diphosphate	GDP	C00035	Nucleotide
2	Purine biosynthesis	I	Guanosine-5′-triphosphate	GTP	C00044	Nucleotide
2	Pyridine nucleotide cycling	I	Nicotinamide	—	C00153	Cation
2	Pyridoxal-10′-phosphate biosynthesis	I	Pyridoxal-5′-phosphate	Pyr5′P	C00018	Anion
2	Pyrimidine biosynthesis	I	Carbamoyl-L-aspartic acid	Carbamoyl-Asp	C00438	Cation
2	Pyrimidine biosynthesis	I	Dihydroorotic acid	—	C00337	Anion
2	Pyrimidine biosynthesis	I	Deoxycytidine-5′-triphosphate	dCTP	C00458	Nucleotide
2	Salvage of Ade, HypXan	I	Adenosine	—	C00212	Cation
2	Salvage of Gua, Xan	I	Deoxyguanosine	dGuanosine	C00330	Cation
2	Salvage of pyridoxal-6′-phosphate	I	Pyridoxamine	—	C00534	Cation
2	Serine biosynthesis	I	L-Serine	Ser	C00065	Cation
2	Sulfuric acid assimilation	I	Adenosine-5′-phosphosulfate	APS	C00224	Nucleotide
2	Sulfuric acid assimilation	I	Adenosine-3′,5′-diphosphate	pAp	C00054	Nucleotide
2	TCA cycle	I	Citric acid	—	C00417	Anion
2	TCA cycle	I	*cis*-Aconitic acid	—	C00417	Anion
2	TCA cycle	I	2-Oxoglutaric acid	2-OG	C00026	Anion
2	TCA cycle	I	Succinyl-CoA	SucCoA	C00091	Nucleotide
2	TCA cycle	I	Succinic acid	—	C00042	Anion
2	TCA cycle	I	Fumaric acid	—	C00122	Anion
2	TCA cycle	I	Malic acid	—	C00497	Anion
2	Thiamine biosynthesis	I	Thiaminephosphate	ThiamineP	C010181	Cation
2	Threonine biosynthesis	I	*O*-Phospho-L-homoserine	P-HomoSer	C01102	Anion
2	Tyronine biosynthesis	I	*p*-Hydroxyphenylpyruvic acid	HPP	C01179	Anion
2	UDP-*N*-acetylglucosamine	I	UDP-*N*-acetyl-D-glucosamine	UDP-GlcNAc	C00043	Nucleotide
2	Valine biosynthesis	I	2-Acetolactic acid	2-Ac	C06010	Anion
2	Valine biosynthesis	I	2-Ketoisovaleric Acid	2-KIV	C00141	Anion
2	Valine biosynthesis	I	2,3-Dihydroxyisovaleric acid	2,3-DIV	C04039	Anion
2	Valine biosythesis	I	L-Valine	Val	C00183	Cation
2	Acetylglucosamine dissimilation	II	*N*-Acetylneuraminic acid	NeuNAc	C00270	Anion
2	Arginine degradation	II	N^2-Succinylglutamic acid	N2SucGLu	C05931	Cation
2	Lysine degradation	II	Cadaverine	—	C01672	Cation
2	Methylmalonyl pathway	II	D-Methylmalonyl-CoA	—	C02557	Nucleotide
3	Alanine biosynthesis	I	L-Alanine	Ala	C00041	Cation
3	Arginine biosynthesis	I	L-Arginine	Arg	C00062	Cation
3	Asparagine biosynthesis	I	L-Asparagine	Asn	C00152	Cation
3	Aspartic acid biosynthesis	I	L-Aspartic acid	Asp	C00049	Cation
3	Chorismic acid biosynthesis	I	5-Enolpyruvylshikimate-3-phosphate	5EPS3P	C01269	Anion
3	Common antigen biosynthesis	I	dTDP-4-acetamido-4,6-dideoxy-D-galactose	dTDP-AADDGal	-	Nucleotide
3	Glutamine biosynthesis	I	L-Glutamine	Gln	C00064	Cation
3	Glutathione biosynthesis	I	Glutathione (oxidized form)	GSSG	C00127	Cation
3	Glycine biosynthesis	I	Glycine	Gly	C00037	Cation
3	Histidine biosynthesis	I	Phosphoribosyl-AMP	PR-AMP	C02741	Nucleotide
3	Histidine biosynthesis	I	Phosphoribosylformimino-AICAR-phosphate	PRFAP	C04896	Nucleotide
3	Histidine biosynthesis	I	D-Erythroimidazoleglycerolphosphate	EIGP	C04666	Nucleotide
3	Histidine biosynthesis	I	Imidazoleacetolphosphate	IAP	C01267	Nucleotide
3	Histidine biosynthesis	I	L-Histidinolphosphophosphate	Histidinol-P	C01100	Nucleotide
3	Histidine biosynthesis	I	L-Histidinol	Histidinol	C00860	Cation

Table 1 List of metabolites clustered by HCA (*Continued*)

Cluster No.	Pathway	Class[a]	Metabolite	Abbreviation	KEGG ID	Analytical mode
3	Isopentenyldiphosphate biosynthesis	I	2-*C*-Methyl-D-erythritol-2,4-cyclodiphosphate	MeEry2, 4CP	C11453	Anion
3	Methylglyoxal pathway	I	*S*-Lactoylglutathione	SL-GSH	C03451	Cation
3	Peptideglycan biosynthesis	I	D-Alanyl-D-alanine	AlaAla	C00993	Cation
3	Peptideglycan biosynthesis	I	L-Alanyl-D-glutamyl-*meso*-A2pm	AlaGluA2pm	—	Cation
3	Phenylalanine biosynthesis	I	L-Phenylalanine	Phe	C00079	Cation
3	Polyamine biosynthesis	I	5′-Methylthioadeosine	MTA	C00170	Cation
3	Polyamine biosynthesis	I	Spermidine	—	C00315	Cation
3	Purine biosynthesis	I	Aminoimidazole carboxamide ribonucleotide	AICAR	C04677	Nucleotide
3	Pyrimidine biosynthesis	I	Uridine-5′-phosphate	UMP	C00105	Nucleotide
3	Pyrimidine biosynthesis	I	Cytidine-5′-triphosphate	CTP	C00063	Nucleotide
3	Salvage of pyridoxal-7′-phosphate	I	Pyridoxamine-5′-phosphate	PX5′P	C00647	Anion
3	Salvage of pyrimidine	I	Cytidine-5′-phosphate	CMP	C05822	Nucleotide
3	SAM biosynthesis	I	*S*-Adenosyl-L-methionine	SAM	C00019	Cation
3	Tryptophan biosynthesis	I	L-Tryptophan	Trp	C00078	Cation
3	Tryrosine biosynthesis	I	L-Tyrosine	Tyr	C00082	Cation
3	Ubiquinone biosynthesis	I	*p*-Hydroxybenzoic acid	PHBA	C00156	Antion
3	Arginine degradation	II	N^2-Succinyl-L-arginine	N2SucArg	C03296	Cation
4	AppppA biosynthesis	I	*P*(1), *P*(4)-Bis(5′-adenosyl)tetraphosphate	AppppA	C01260	Nucleotide
4	Arginine biosynthesis	I	*N*-α-Acetylornithine	N-AcOrn	C00437	Cation
4	Arginine biosynthesis	I	L-Ornithine	Ornthine	C00077	Cation
4	Biotin biosynthesis	I	Dethiobiotin	—	C01909	Cation
4	Fravin biosynthesis	I	Flavinmononucleotide (oxidized form)	FMN	C00061	Nucleotide
4	Fravin biosynthesis	I	Flavinadeninedinucleotide (oxidized form)	FAD	C00016	Nucleotide
4	Glycolysis (Embden-Meyerhof-Parnas)	I	3-Phosphoglyceric acid	3-PG	C00597	Anion
4	Glycolysis (Embden-Meyerhof-Parnas)	I	Phosphoenolpyruvic acid	PEP	C00074	Anion
4	Homoserine biosynthesis	I	L-Homoserine	HomoSer	C00263	Cation
4	Lactic acid fermentation	I	D-Lactic acid	Lactic acid	C01432	Anion
4	Lysine biosynthesis	I	*N*-Succinyl-2-amino-6-ketopimelic acid	SAKPA	C04462	Anion
4	Lysine biosynthesis	I	*N*-Succinyl-L,L-2,6-diaminopimelic acid	SDPA	C04421	Anion
4	Methionine biosynthesis	I	Cystathionine	—	C00542	Cation
4	NAD biosynthesis	I	Nicotinamideadeninedinucleotide (oxidized form)	NADH	C00004	Nucleotide
4	Peptideglycan biosynthesis	I	UDP-*N*-acetylmuramic acid	UDP-MurNaAcA	C01050	Nucleotide
4	Purine biosynthesis	I	5-Phosphoribosyl-*N*-formylglycineamidine	FGAM	C04640	Nucleotide
4	Pyrimidine biosynthesis	I	Uridine-5′-triphosphate	UTP	C00075	Nucleotide
4	Tetrahydrofolic acid biosynthesis	I	2-Amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridinediphosphate	AHHMeDHPDP	C04807	Nucleotide
4	Arginine degradation	II	N^2-Succinylornithine	N2SucOrn	C03415	Cation
4	*S*-Adenosylhomocysteine catabolism	II	*S*-D-Ribosyl-L-homocysteine	RibHomoCys	C03539	Cation
4	Threonine degradation	II	Propionyl-CoA	—	C00100	Nucleotide
5	CoA biosynthesis	I	Dephospho-CoA	dePCoA	C00882	Nucleotide
5	CoA biosynthesis	I	Coenzyme A	CoA	C00010	Nucleotide
5	Glutathione biosynthesis	I	Glutathione (reduced form)	GSH	C00051	Cation
5	Glutathionylspermidine metabolism	I	Glutathionylspermidine	GSH-spermidine	C05730	Cation
5	Glycolysis (Embden-Meyerhof-Parnas)	I	Pyruvic acid	—	C00022	Anion
5	Histidine biosynthesis	I	L-Histidine	His	C00135	Cation
5	Isoleucine biosynthesis	I	L-Isoleucine	Ile	C00407	Cation
5	KDO biosynthesis	I	D-Arabinose-5-phosphate	Ara5P	C01112	Anion
5	Leucine biosynthesis	I	2-Isopropylmalic acid	2IPMA	C02504	Anion
5	Pantothenic acid biosynthesis	I	2-Dehydropantoic acid	2DHPA	C00966	Anion
5	Pantothenic acid biosynthesis	I	Pantothenic acid	—	C00864	Anion

Table 1 List of metabolites clustered by HCA (*Continued*)

Cluster No.	Pathway	Class[a]	Metabolite	Abbreviation	KEGG ID	Analytical mode
5	Pentose phosphate pathway (oxidative)	I	D-Ribulose-5-phosphate	Ru5P	C00199	Anion
5	Polyamine acetylation	I	N^1-Acetylspermidine	N1AcSpermidine	C00612	Cation
5	Proto-and siroheme biosynthesis	I	Protoheme IX	—	C00032	Anion
5	Pyrimidine biosynthesis	I	Uridine-5′-diphosphate	UDP	C00015	Nucleotide
5	Pyrimidine biosynthesis	I	Cytidine-5′-diphosphate	CDP	C00112	Nucleotide
5	Pyrimidine biosynthesis	I	Deoxythymidine-5′-diphosphate	dTDP	C00363	Nucleotide
5	Salvage of Ade, HypXan	I	Hypoxanthine	—	C00262	Cation
5	Salvage of Ade, HypXan	I	Inosine	—	C00294	Cation
5	Salvage of Ade, HypXan	I	Xanthine	—	C00385	Cation
5	Salvage of Gua, Xan	I	Guanine	—	C00242	Cation
5	Salvage of Gua, Xan	I	Guanosine	—	C00387	Cation
5	Salvage of pyrimidine	I	Cytosine	—	C00380	Cation
5	Salvage of pyrimidine	I	Thymine	—	C00178	Cation
5	Threonine biosynthesis	I	L-Threonine	Thr	C00188	Cation
5	UDP-*N*-acetylglucosamine biosynthesis	I	*N*-Acetylglucosamine-1-phosphate	GlcNAc-P	C04256	Anion
5	Histidine degradation	II	Imidazole lactic acid	Im-lactate	C05568	Cation
5	Threonine degradation	II	Aminoacetone	—	C01888	Cation

[a] Classes: I, primary matabolism; II, degradation of primary metabolites; III, Degradation of environmental compounds; IV, secondary or unconventional metabolism; V, pathway unknown; VI, intermediates in putative *in vitro* reaction.

Fig. 1 Classification of *E. coli* metabolites. A. Distribution of metabolites by metabolic pathways. The classes are described in the Experimental section. B. Hydrophobicity of metabolites in classes I and II. H_2O was excluded from the calculation. C. Distribution of analytical modes in CE-TOFMS analysis for hydrophilic metabolites in classes I and II. The neutral metabolites cannot be analysed by CE-MS.

However, standard chemicals of most of the intermediates are not commercially available, making pathway analysis difficult (Fig. 4A). Providing the migration times of the intermediates in CE-TOFMS analysis enables the analysis of histidine biosynthesis as a part of metabolome analysis. Since, most intermediates in histidine biosynthesis are phosphate-including complex compounds which are detected by CE-TOFMS in nucleotide mode, we performed CE-TOFMS analyses of the metabolites extracted by the improved preparation method presented here. To accumulate the histidine intermediates, we employed an *E. coli* mutant (JW2002) carrying a null mutation in the *hisD* allele that codes for histidine dehydrogenase catalysing the conversion of L-histidinol to L-histidine *via* L-histidinal.[18]

The histidine biosynthesis pathway is repressed, as a posttranslational regulation, by histidine at the first step of the pathway,[19] and is down-regulated by attenuation of the gene expression under conditions of excess histidine.[20] The expression of *his* operon is also induced by amino acid starvation in a ppGpp-mediated manner.[21] Hence, it is expected that amino acid starvation stimulates the accumulation of histidine biosynthesis intermediates. *E. coli* JW2002 was grown in a minimal medium supplemented with histidine and then resuspended in a minimal medium without histidine at the middle-to-late logarithmic growth phase. Accumulation of histidine biosynthesis intermediates are shown in Fig. 4B. The intermediates for which standard chemicals are unavailable were identified based on their calculated *m/z* values (Fig. 4C). A downshift of histidine in the medium activates histidine biosynthesis provoking accumulation of its

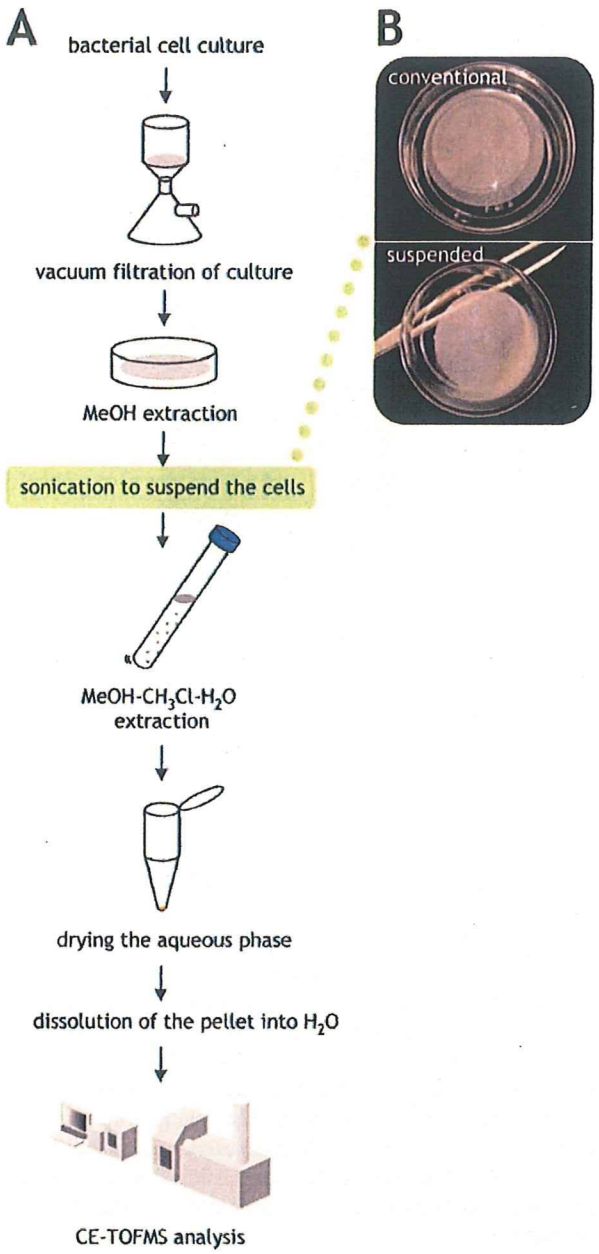

Fig. 2 Metabolome extraction procedure modified in this study. A. Schematic of the metabolome extraction procedure. Modified point is indicated in the green box. B. Cell suspension prepared by sonication of methanol.

intermediates. Actually, the intracellular level of histidine dropped rapidly within 30 min of the downshift, resulting in the activation of the histidine biosynthesis pathway. Though phosphoribosyl-ATP was not detected here, chronological accumulation profiles of the intermediates were observed. The phosphoribosylpyrophosphate (PRPP) level was increased at the onset of starvation (0–15 min), and then slightly decreased in concert with the drop of histidine concentrations (30–60 min). These results are consistent with the previous observations,[19] and clearly indicate the accumulation of intermediates in histidine biosynthesis. Additionally, among the intermediates only PRPP and L-histidinol are commercially available at present, and thus the results yield important information of CE-TOFMS analysis in histidine biosynthesis.

Metabolome profiles under histidine starvation conditions

Amino acid starvation induces the accumulation of ppGpp, stimulating stringent response in cells.[11] Actually, ppGpp levels were significantly increased after the histidine downshift (supplementary Table 2†). This observation led us to view all the metabolite changes during histidine starvation as the stringent response. We determined the metabolome profiles including 375 metabolites belonging to classes I and II (supplementary Table 2). Of the metabolites, a total of 198 were successfully detected in cells during histidine starvation, and are depicted in the *E. coli* metabolic pathway map shown in Fig. 5 (high resolution, printable version is available as supplementary Fig. 1†). Of the metabolites detected in cells, 45 were determined based on their *m/z* values because the standard chemicals are not available. At a glance, dramatic changes were found in regional primary metabolisms, but most of the metabolites were not changed. Since the metabolic network in *E. coli* is generally robust against environmental perturbations,[22] the metabolic changes found here seem to involve homeostasis of metabolic processes.

To identify the metabolic changes, we performed clustering analysis of the metabolome data using hierarchical clustering analysis (HCA), generating five significant clusters. Clustering analysis is applicable and valuable in transcriptomic analysis, since the accumulation of a specific mRNA directly shows the activation of the gene expression.[23] However, it should be noted in metabolomics that the accumulation of a metabolite does not indicate activation of the metabolic pathway, it should be estimated in combination with the change of its down- and upstream metabolites. The clusters generated here are shown in Fig. 6 and listed in Table 1. The metabolites for which levels were unchanged were excluded from HCA. Clusters 1, 2, and 3 resembled each other in the profiles and tend to be accumulated after histidine downshift (Fig. 6B). Cluster 1 metabolites were temporary accumulated during 15 to 30 min after the histidine downshift, and then decreased (Fig. 6B). Cluster 2 metabolites were rapidly accumulated at 15 min and maintained thereafter (Fig. 6B). Accumulations of cluster 3 metabolites were delayed to 60 min (Fig. 6B). Clusters 4 and 5 tended to be decreased gradually and rapidly, respectively (Fig. 6B).

During stringent responses, DNA replication, transcription, translation, and cell division are arrested in a ppGpp-dependent manner.[11] Actually, growth was seen to be arrested by monitoring optical density at 600 nm at the onset of histidine starvation (data not shown). High-energy carrier molecules, ATP and GTP, were accumulated slightly (cluster 2), suggesting that DNA replication, and transcription of ribosomal RNAs are stopped, as described elsewhere.[11] This is considered to be a reservation for further growth when *de novo* histidine biosynthesis is activated (growth cannot occur because of *hisD* disruption in the cells). The building blocks of DNA, dNTP, were maintained after histidine downshift. Amino acid levels were significantly accumulated after

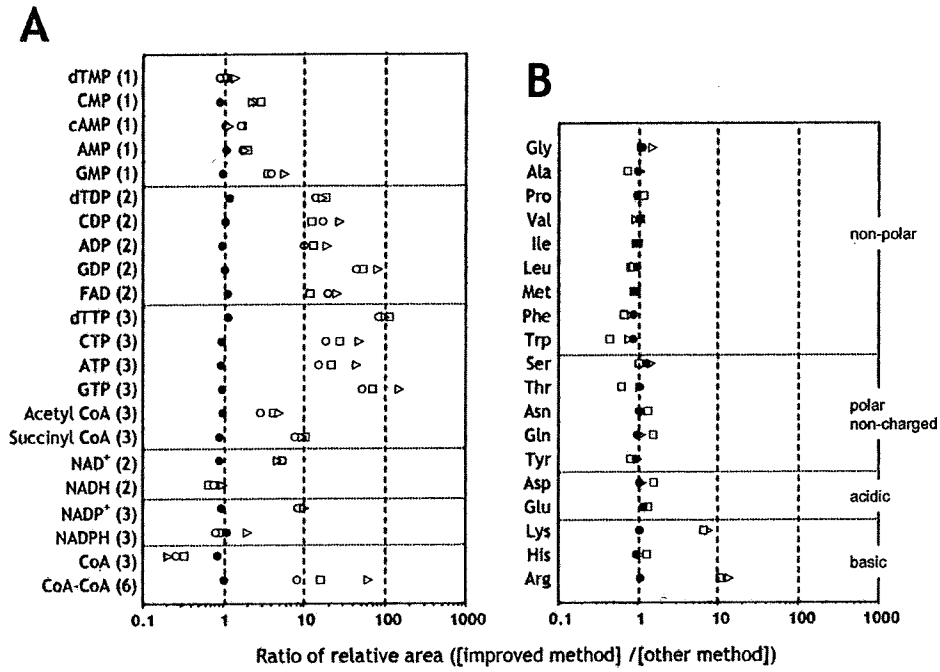

Fig. 3 Improved extraction efficiency of nucleotide metabolites by the improved method developed in this study. A. Comparison of the detected levels of nucleotide metabolites. The number of phosphate groups that are contained in the metabolite is shown in parentheses. B. Comparison of the detected levels of amino acids. ●, area ratio of [improved method (1st)] to [improved method (2nd)], that is, the reproducibility of the improved method; △, area ratio of [improved method (1st)] to [conventional method]; □, area ratio of [improved method (1st)] to [sonicated cell removal method]; ○, area ratio of [improved method (1st)] to [cold method]. See text for details of the extraction procedures.

downshift: most are distributed to clusters 2 and 3. In particular, accumulation of lysine, leucine, arginine, and phenylalanine was prominent. Of amino acids, threonine, except histidine, was assigned to cluster 5, indicating threonine was consumed after histidine-downshift. These results suggest that threonine biosynthesis from aspartic acid was shut down during starvation, resulting in the drop of isoleucine and following stimulation of valine and leucine biosynthesis.[24]

Next, we focused on the behavior of cluster 2 metabolites which were clustered with ppGpp in order to discuss the candidate pathways under stringent control (Table 1). The first half of glycolysis, from glucose-6-phosphate (G6P) to glyceraldehyde-3-phosphate (G3P) and dihydroacetonephosphate (DHAP), was accumulated in concert with the intermediates in the pentose phosphate pathway (Fig. 5, positions B-2 and C-2). This effect seems to involve increased histidine biosynthesis *via* PRPP. However, the middle of glycolysis, from 3- or 2-phosphoglyceric acid (3- or 2-PG) to phosphoenolpyruvic acid (PEP), was unchanged (Fig. 5, position C-3), while late step of glycolysis, pyruvic acid, dropped remarkably (Fig. 5, position C-3). These results suggest that the behavior of glycolysis is complicated during histidine starvation: that is, the first half is activated with the pentose phosphate pathway, but the flux is not extended to pyruvic acid synthesis. Moreover, the entire tricarboxylic acid (TCA) cycle intermediates were distributed into cluster 2, indicating that the TCA cycle was activated to yield precursors of downstream pathways and consumes intracellular pyruvic acid. Although we cannot exclude the possibility that the aforementioned pathways are regulated in a ppGpp-independent manner, the results provide new insights for understanding the stringent response in bacteria. Metabolomics observations in the *relA* mutant are needed to unveil the mechanism of stringent response in metabolism.

Experimental

Preparation of *E. coli* metabolite list

Metabolites in *E. coli* were selected from the EcoCyc[15,16,25] and KEGG[26,27] databases. The metabolites listed in the databases were identified one by one in the literature. Finally, 727 metabolites were selected as *E. coli* metabolites (supplementary Table 1†) and then classified based on the metabolism pathways as follows. Class I metabolites belong to primary metabolisms including biosynthesis of building block metabolites of cells, energy metabolisms such as the glycolysis and TCA cycle, and other biosynthesis pathways. Class II includes the intermediates of degradation pathways of primary metabolites, *e.g.*, amino acid degradation. The metabolites in classes I and II are expected to be found in cells grown in general minimal medium. Class III involves the degradation pathways of environmental compounds which are not included in general minimal medium. Class IV includes secondary metabolism or unconventional metabolisms. Class V includes the metabolites found in *E. coli* cells previously, but the pathways are unknown. Class VI metabolites are intermediates and products of *in vitro* enzymatic reactions using putative gene products.

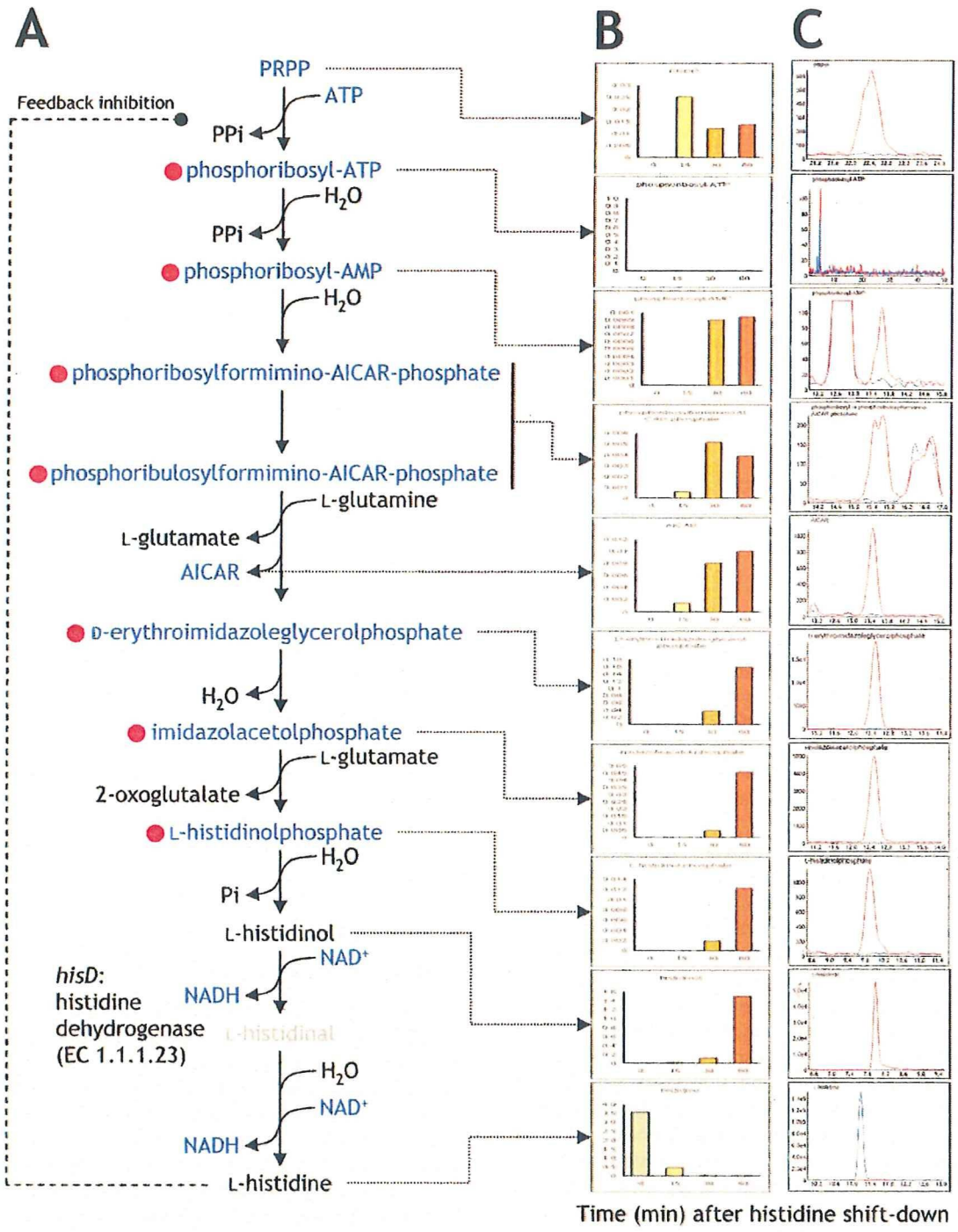

Fig. 4 Signal identification of histidine biosynthesis intermediates. A. Histidine biosynthesis pathway in *E. coli*. Metabolites that are detected in nucleotide mode by CE-TOFMS analysis are indicated in blue. Red circles show the metabolites for which standard chemicals are not commercially available. L-Histidinal coloured gray cannot be detected because it is not released during the overall reaction catalysed by histidine dehydrogenase.[31] B. Accumulation of intermediates of histidine biosynthesis after histidine downshift. The data are indicated as the area values relative to that of internal standard. C. Electropherograms of the intermediates obtained by CE-TOFMS analyses. The metabolites except for L-histidinol and L-histidine were detected with nucleotide mode. The data are indicated as the signal intensities (arbitrary unit).

Bacterial strains and growth conditions

E. coli W3110 (laboratory stock) grown in Luria-Bertani (LB) liquid medium[28] was used for optimisation of metabolome extraction procedure. The *hisD* deletion mutant, JW2002, was distributed from the Keio collection.[29] The cells were pre-grown on LB plates at 37 °C for 12 h. The fresh colonies that appeared were inoculated into M9 liquid medium[28] containing

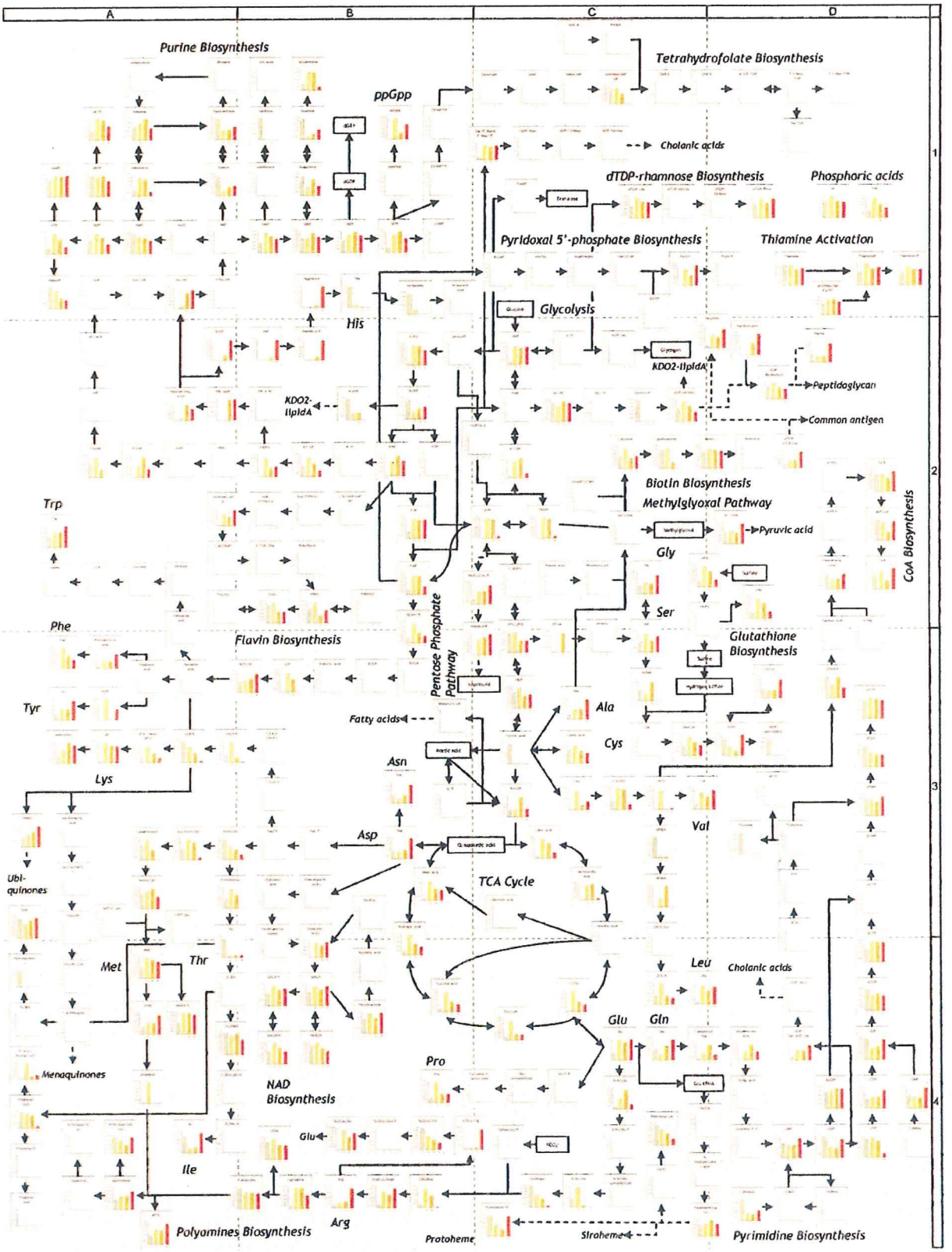

Fig. 5 Metabolomic profile of *E. coli* after histidine downshift. The data of each metabolite level are summarized in supplementary Table 2. Enlarged, printable version is available as supplementary Figure 1.†

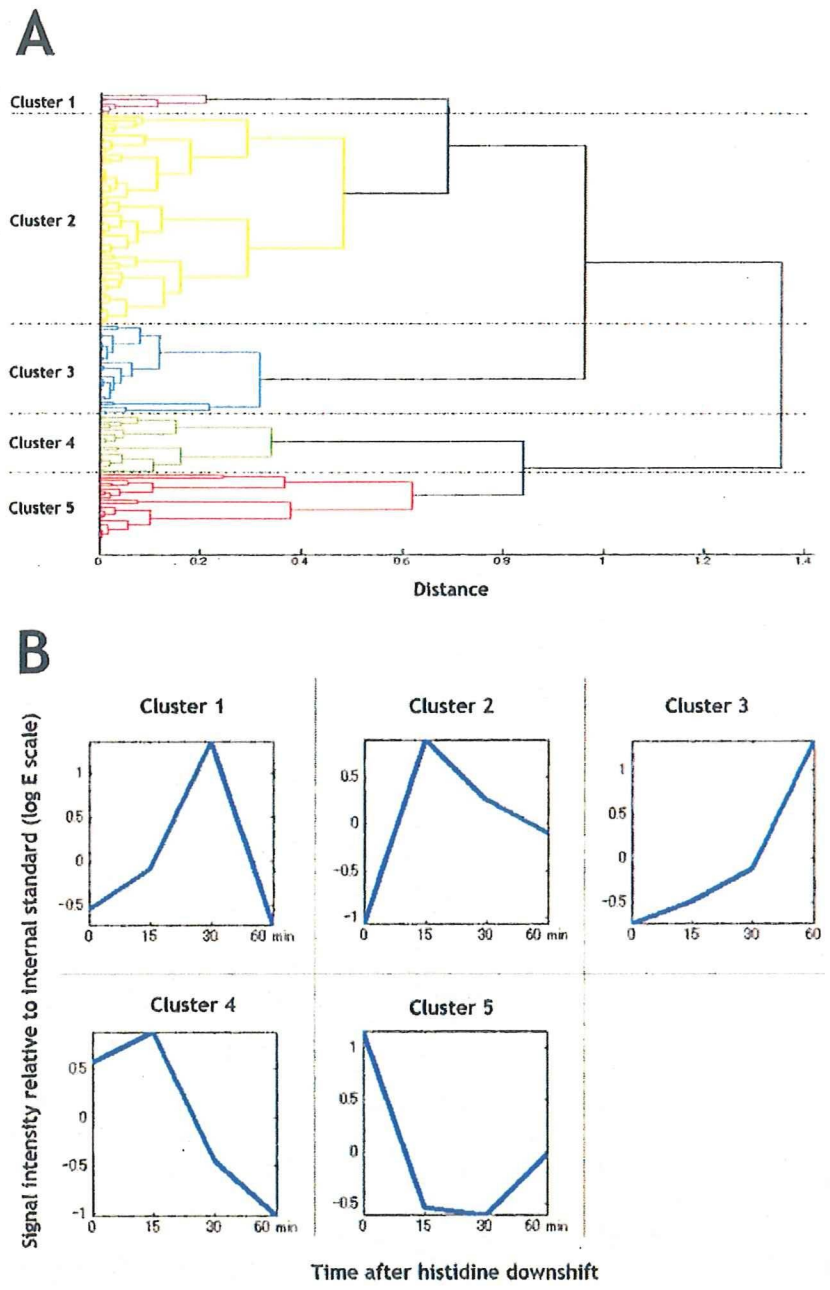

Fig. 6 Clustering analysis of metabolites for which intracellular levels were changed during histidine starvation. A. Tree view of distances between the clusters generated by HCA. B. Average view of the clusters. Signal intensities relative to that of internal standard were standardised.

50 μg ml^{-1} of L-histidine, and this was followed by incubation at 37 °C with shaking. When the cells reached the middle-to-late logarithmic phase (optical density at 600 nm = 0.8), they were harvested by brief centrifugation, washed with M9 medium without histidine, and then suspended in the M9 medium. The culture was continuously shaken at 37 °C until the sampling times.

Metabolite extraction procedure

Intracellular metabolites were extracted as described previously with some modifications.[6] Schematic of the metabolome extraction procedure is shown in Fig. 2. Cells grown as described above were collected for metabolome analysis at indicated times. Culture including approximately 10^9 cells (calculated as optical density at 600 nm × sampling volume of culture (ml) = 20) was filtered by a vacuum filtration system using a 0.4 μm pore size filter. The residual cells on the filter were washed twice with 5 ml of Milli-Q water. The filter was immersed in 2 ml of methanol including 5 μM each of internal standards, methionine sulfone and D-camphor-10-sulfonic acid (CSA). The dish was sonicated for 30 s using an Elma Transsonic T460/H ultrasonic syringe (not an ultrasonic

cell disrupter) (Elma Hans Schmidbauer GmbH & Co., Singen, Germany) to suspend the cells completely. Effects of this process on metabolite extraction efficiency are discussed in the results and discussion section. A 1.6 ml portion of the methanol cell suspension was transferred to a Falcon Blue Max Jr., 352097 centrifugal tube (15 ml) (Becton Dickinson & Co., New Jersey, USA), and mixed with 1.6 ml of chloroform and 640 μl of Milli-Q water. After vortexing well, the mixture was centrifuged at 4600 × *g* and 4 °C for 5 min. The aqueous layer (750 μl) was distributed to three Amicon Ultrafree-MC ultrafilter tips (Millipore Co., Massachusetts, USA) and centrifuged at 9100 × *g* and 4 °C for approximately 2 h. The filtrate was dried and preserved at −80 °C until CE-MS analysis. Prior to analysis, the sample was dissolved in 25 μl of Milli-Q water.

Instrumentation

CE-TOFMS was carried out using an Agilent CE Capillary Electrophoresis System equipped with an Agilent 6210 Time-of-Flight mass spectrometer, Agilent 1100 isocratic HPLC pump, Agilent G1603A CE-MS adapter kit, and Agilent G1607A CE-ESI-MS sprayer kit (Agilent Technologies, Waldbronn, Germany). The system was controlled by Agilent G2201AA ChemStation software version B.03.01 for CE (Agilent Technologies, Waldbronn, Germany). Data acquisition was performed by Analyst QS Build: 7222 software for Agilent TOF (Applied Biosystems, California, USA/MDS Sciex, Ontario, Canada).

CE-TOFMS conditions

Separations and detections of metabolites were basically performed as described previously for cationic metabolites,[3,6] anionic metabolites,[4,6] and nucleotides.[30] For cationic metabolites, capillary electrophoreses were performed using a fused silica capillary. The electrolyte was 1 M formic acid. Methanol-water (50% v/v) containing 0.5 μM reserpine (the lock mass for exact mass measurements) was delivered as the sheath liquid at 10 μl min^{-1}. For anionic metabolites, a polymer coated SMILE(+) capillary (Nacalai tesque, Kyoto, Japan) was used. The electrolyte was 50 mM ammonium acetate (pH 8.5). Ammonium acetate (5 mM) in 50% (v/v) methanol-water containing 1 μM reserpine was delivered as the sheath liquid at 10 μl min^{-1}. For nucleotides, separations were performed using a fused silica capillary. The electrolyte was 50 mM ammonium acetate (pH 7.5). The sheath liquid for anionic metabolites was used as the electrolyte. The capillary was pretreated with preconditioning buffer including 25 mM ammonium acetate and 75 mM sodium phosphate at pH 7.5. Pressure of 50 mbar was applied to inlet capillary during run to reduce the analysis time. For all analytical modes, inner diameter and total length of capillary are 50 μm and 100 cm, respectively. The applied voltage was set at +30 kV and −30 kV for cation and anion modes and nucleotide mode, respectively.

Electrosplay ionisation-TOFMS was operated in the positive ion mode (4 kV), the negative ion mode (3.5 kV), and the negative ion mode (3.5 kV) for cationic metabolites, anionic metabolites, and nucleotides, respectively. A flow rate of heated dry nitrogen gas (heater temperature 300 °C) was maintained at 10 psig. Exact mass data were acquired over a 50–1000 *m/z* range.

Cluster analysis

The metabolome data set was prepared by selection of metabolites based on a coefficient of variation of more than 0.2 to exclude the metabolites for which levels were unchanged. HCA was performed on the data set using MATLAB 2007a (The Math Works, Massachusetts, USA). Distances between the metabolites were calculated by the equation, correlation coefficient + 1.

Conclusion

In this study, we produced a list of previously reported *E. coli* metabolites together with analytical information for CE-TOFMS-based metabolome analysis. This list is also useful for other analytical methods using MS as a detector. The metabolome extraction method was also improved, realising quantitative analyses of phosphate-rich metabolites and other hydrophilic, charged metabolites. However, this method cannot be applied to the analysis of metabolites oxidised in the metabolome pool. Extraction methods inhibiting oxidation of the metabolites are needed to determine them. We successfully determined 198 metabolites in primary metabolisms of *E. coli* histidine-auxotroph. The metabolome profiles were analysed by mapping data to metabolic pathways and HCA, suggesting separated activation of amino acid biosynthesis pathways, glycolysis, and the TCA cycle. Although more detailed studies using the *relA* mutant is required to unveil the entire view of stringent responses mediated by ppGpp, we believe that metabolome profiling should provide new insights into bacterial adaptation to environmental changes.

Acknowledgements

We are grateful to Hirotada Mori, Tomoya Baba, and Kenji Nakahigashi for providing the *E. coli* mutant strain and to Masatomo Hirabayashi and Yuji Sakakibara for providing excellent technical expertise in CE-TOFMS analysis. We also thank Mineo Morohashi for helpful discussions.

References

1 I. Nobeli, H. Ponstingl, E. B. Krissinel and J. M. Thornton, *J. Mol. Biol.*, 2003, **334**, 697–719.
2 M. R. Monton and T. Soga, *J. Chromatogr., A*, 2007, **1168**, 237–246.
3 T. Soga and D. N. Heiger, *Anal. Chem.*, 2000, **72**, 1236–1241.
4 T. Soga, Y. Ueno, H. Naraoka, Y. Ohashi, M. Tomita and T. Nishioka, *Anal. Chem.*, 2002, **74**, 2233–2239.
5 T. Soga, Y. Ueno, H. Naraoka, K. Matsuda, M. Tomita and T. Nishioka, *Anal. Chem.*, 2002, **74**, 6224–6229.
6 T. Soga, Y. Ohashi, Y. Ueno, H. Naraoka, M. Tomita and T. Nishioka, *J. Proteome Res.*, 2003, **2**, 488–494.
7 T. Soga, R. Baran, M. Suematsu, Y. Ueno, S. Ikeda, T. Sakurakawa, Y. Kakazu, T. Ishikawa, M. Robert, T. Nishioka and M. Tomita, *J. Biol. Chem.*, 2006, **281**, 16768–16776.
8 S. G. Villas-Bôas, J. Højer-Pedersen, M. Åkesson, J. Smedsgaard and J. Nielsen, *Yeast*, 2005, **22**, 1155–1169.
9 R. P. Maharjan and T. Ferenci, *Anal. Biochem.*, 2003, **313**, 145–154.