Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 815 (2005) 251–260
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Review

Proteomics and physiology of erythritol-producing strains

Yong-Cheol Park[a,1], Do-Yup Lee[a], Dae-Hee Lee[a], Hyo-Jin Kim[a], Yeon-Woo Ryu[b], Jin-Ho Seo[a,*]

[a] *School of Agricultural Biotechnology, Center for Agricultural Biomaterials, Seoul National University, Seoul 151-742, Korea*
[b] *Department of Molecular Science and Technology, Ajou University, Suwon 442-749, Korea*

Received 18 August 2004; accepted 22 October 2004
Available online 24 November 2004

Abstract

In-depth knowledge bases on physiological properties of microbes are required to design a better microbial system at a gene level and to develop an industrially viable process in an optimized scheme. Proteomic analyses of industrially useful microorganisms are particularly important for achieving such objectives. In this review, industrial application of erythritol in food and pharmaceutical areas and proteomic techniques for erythritol-producing microbes were presented. Proteomic technologies for erythritol-producing strains such as *Candida magnoliae* contained protein or peptide sample preparation for two-dimensional electrophoresis and mass spectrometry, analysis of proteome with matrix assisted laser desorption-ionization/time-of-flight mass spectrometry, liquid chromatography/electrospray ionization/tandem mass spectrometry and similarity searching algorithms. The proteomic information was applied to predict the carbon metabolism of erythritol-synthesizing microorganisms.

Keywords: Review; Proteomics; *Candida magnoliae*; Erythritol

Contents

* Corresponding author. Tel.: +82 2 880 4862; fax: +82 2 873 5095.
E-mail address: jhseo94@snu.ac.kr (J.-H. Seo).
[1] Present address: Department of Biochemistry and Cell Biology, Rice University, Houston, TX 77251-1892, USA.

1570-0232/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2004.10.065

1. Introduction

Past several decades, sugars and its derivatives have been developed for matching the consumer's demands such as low caloric and anti-cariogenic property and lowering the blood-glucose level. Sugar alcohols containing a hydroxyl group at the C_1-position of carbohydrate can be used as functional sugar substitute bearing the above characteristics. To date, various sugar alcohols such as mannitol, maltitol, xylitol, arabitol and erythritol have been used as food and medicine ingredients. Erythritol, a four-carbon sugar alcohol, is produced by a fermentation process because of expensive erythrose for chemical or enzymatic conversion to erythritol. Osmophilic erythritol-producing microorganisms are usually classified as unconventional yeast in that metabolism of erythritol and their physiological properties as well as genomic and proteomic databases are insufficient for the systematic analysis of erythritol production. Genetic, proteomic and physiological information of erythritol-producing microorganisms as biocatalyst for a commercial process would be expected to improve the biological process of erythritol production in an industrial environment.

After development of two-dimensional electrophoresis (2-DE) by O'Farrell [1] and definition of "proteome" terminology [2], Human Genome Project triggered the proteomic research of organisms. Several devices such as dry strips with immobilized pH gradients, matrix-assistant desorption and ionization time-of-flight (MALDI-TOF) mass spectrometry, electrospray ionization tandem mass spectrometry (ESI-MS/MS) and web-based data searching services exploded the accumulation of proteomic databases. Recent progress has been made in the development of alternative methods of protein separation for proteomics such as spotted-array-based methods, microfluidic device, the use of affinity tags and large-scale yeast two-hybrid screening [3,4]. Other approaches have been performed to characterize proteins rapidly and quantitatively by the use of mass spectrometry with isotope labeling and development of automated machine combined with protease digestion and electroblotting to a membrane in a single step [5].

This review covers physicochemical properties of erythritol, commercial production of erythritol, applications in foods and pharmaceuticals and proteomic techniques for erythritol-producing microorganisms. Especially, proteomic analysis of erythritol-producing microorganisms focused on protein preparation for 2-DE and mass spectrometric analysis, and techniques of mass spectrometry such as MALDI-TOF and ESI-MS/MS for the case of *Candida magnoliae*. In addition, the proteomic database of erythritol production was applied to elucidate the metabolic network and physiological properties of erythritol-producing microorganisms. This article would be able to provide a guideline for choosing an efficient way to analyze newly identified proteins by using currently available databases without its own genomic information.

1.1. Erythritol

Erythritol, a four-carbon sugar alcohol (polyol), is a natural and moderately sweet bulk sweetener with taste and mouthfeel enhancing properties. It occurs naturally in various fruits and fermented foods including watermelon, pear, grape, wine, sake, beer and soy sauce [6]. Erythritol also exists endogenously in human plasma at about 1.2 mg/L and normally in human urine [7–9]. Erythritol has a 60–80% sweetness relative to sucrose and is a good low-calorie sweetener (1.26 J/g) [10,11]. Its structure is shown in Fig. 1. Erythritol has a positive enthalpy of solubilization (23.3 kJ/mol), thus providing a strong cooling effect in ingestion [12]. Since erythritol does not have an aftertaste, it can be used in combination with intense sweeteners that have a bitter aftertaste like aspartame. Erythritol has been used safely as a noncariogenic sweetener in many countries, owing to inability of erythritol metabolism in cariogenic organisms [13]. More than 90% of ingested erythritol is not metabolized by the human body and is excreted unchanged in the urine without changing blood glucose and insulin levels. Therefore, it could be used advantageously in special foods for people with diabetes and fatness [14]. A little amount of erythritol can be metabolized in some reversible metabolic reactions such as dehydrogenation to D- or L-erythrulose by NAD-dependent cytoplasmtic polyol dehydrogenase or phosphorylation by glycerol kinase to erythritol-1-phosphate followed by dehydrogenation to D-erythrulose-1-phosphate via α-glycerophosphate dehydrogenase in humans [15,16]. In Japan, erythritol has been used since 1990 as sugar substitutes for candies, chocolates, soft

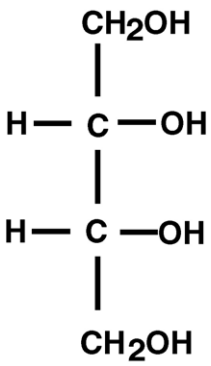

Fig. 1. Structured of erythritol.

drinks, chewing gum, jellies, jams and yogurt [17]. It has been approved in U.S.A. in 2001 and used as a flavor enhancer, formulation aid, humectants, nutritive sweetener, stabilizer, thickener, sequestrant and texturizer at maximum levels of 100% in sugar substitutes [18].

1.2. Production of erythritol

Erythritol can be produced by a chemical process where dialdehyde starch is converted into erythritol by a high-temperature chemical reaction in the presence of a nickel catalyst [19]. Due to low yields, the chemical process did not reach to industrialization. Erythritol is commercially produced by Bolak Corporation (Whasung, Kyungki-do, Korea), Cargill Food & Pharm Specialties (Blair, Nebraska, USA) and Mitsubishi Chemical Corporation (Tokyo, Japan). Dextrose (glucose) from chemically and enzymatically hydrolyzed wheat and corn starches is used as a major carbon source to produce erythritol by the fermentation of yeast-like fungi such as *Torula* sp., *Moniliella pollinis* and *Trichosporonoides megachiliensis*. Erythritol is purified by ion exchange resin, activated charcoal, ultrafiltration and crystallization. The final crystalline erythritol contains more than 99% purity [20]. It has been reported that sugar- and salt-tolerant yeasts produced and intracellularly accumulated polyols such as glycerol and arabitol in the presence of high concentrations of sugars [21,22] and that the polyols played important roles as compatible solutes [23,24]. Many research efforts of erythritol biosynthesis have been focused on the selection of microbial strains able to produce erythritol with high yield and the optimization of operation strategies in batch and fed-batch fermentation processes. Erythritol-producing strains can be characterized as osmophilic yeasts including *Moniliella tomentosa* var. *pollinis*, *Aureobacidium* sp., *Trichosporonoides* sp., *Trigonopsis variabilis*, *Trichosporon* sp., *Torula* sp., and *Candida magnoliae*. Ishizuka et al. [25] selected a high erythritol-producing mutant of *Aureobasidium* sp. SN124A with 47.6% yield by UV irradiation and nitrosoguanidine (NTG) treatment. Park et al. [26] developed the repeated fed-batch culture of *Trichosporon* sp. to produce erythritol at 1.86 g/L h productivity and 45% of total erythritol conversion yield. A mutant of *Torula corallina* produced erythritol at 48.9% yield and did not produce glycerol and ribitol as by-products [27]. Recently, a high erythritol-producing yeast strain was isolated from honeycombs and identified as *C. magnoliae* KFCC 11023 [28]. To improve the erythritol-producing ability, the parental *C. magnoliae* strain was mutated by UV irradiation and NTG treatment to give a *C. magnoliae* mutant strain. The mutated *C. magnoliae* showed higher erythritol conversion yield and productivity than the wild strain [29]. Various biological processes using the mutant strain of *C. magnoliae* have been developed for maximizing erythritol production. The optimized fed-batch fermentation resulted in 200 g/L erythritol concentration, 1.2 g/L h productivity and 0.43 g/g yield [30,31].

1.2.1. Physiological properties of *C. magnoliae*

Erythritol-producing *C. magnoliae* possesses unique physiological characteristics compared with typical yeasts and bacteria. It is able to grow in a wide range of pH values in the presence of high concentrations of glucose. *C. magnoliae* can grow rapidly in less than 500 g/L glucose and in the pH range of 3–5 [29]. Unlike *Saccharomyces cerevisiae*, *C. magnoliae* does not produce ethanol, but produces erythritol and glycerol during fermentation. There might be several mechanisms in the adaptation to osmotic stress by *C. magnoliae*, which experiences both osmotic stress and ion toxicity when exposed to NaCl, KCl, KH_2PO_4 or $(NH_4)_2SO_4$ [29]. In response to a high external osmotic environment, *C. magnoliae* accumulates erythritol and glycerol, which compensate for differences between the extracellular and intracellular water potential. Microorganisms generally utilize glucose preferable to other sugars as a carbon sources. *C. magnoliae*, however, prefers fructose to glucose when growing in a mixture of fructose and glucose [32]. Since *C. magnoliae* is isolated from honeycombs, the genetic abnormality seems to be related to the evolutionary environment.

2. Proteomics of *C. magnoliae*

Proteomic approach is a powerful tool to explore the biological aspect of microorganisms in response to environmental perturbations. Several technologies involved in preparation, separation, visualization, collection, identification and analysis of proteome have been developed exceedingly. In this section, proteomic analysis of *C. magnoliae* introduces the aspect of sample preparation and peptide analysis tools specific for *C. magnoliae* including MALDI-TOF mass spectrometry, ESI-MS/MS and database searching.

2.1. Sample preparation

Protein preparation is of importance to obtain high quality separation of proteome. Improper techniques of sample preparation can occur horizontal or vertical streaking on gels, leading to insufficient and incorrect information of proteome. The wall of yeast cells contains a thick structure composed of 80–90% polysaccharides such as glucans, mannans and chitin. Both β-1,6- and β-1,3-glucans assembled with α-1,6-linked mannans construct a microfibrillar network with high rigidity [33]. *C. magnoliae* has the strong cell wall structure as other yeasts like *S. cerevisiae* and *Candida albicans*. Methods for disruption of cell walls, solubilization of proteins and in-gel protein digestion are introduced to prepare the protein samples efficiently for two-dimensional electrophoresis and mass spectrometry.

2.1.1. Cell disruption and protein solubilization

To prepare the cell extract, erythritol-producing *Aureobasidium* sp. cells were washed with cooled acetone and diethylether [34]. The cells were dried and stored at −20 °C for

stabilizing cellular components. Inside the cooling chamber with carbon dioxide, an MSK Cell Homogenizer (B. Braun Japan Co. Ltd.) disrupted the acetone-treated cells repeatedly for a short time. To obtain key enzymes in erythritol biosynthesis, *T. corallina* and *C. magnoliae* cells were ruptured by grinding with glass beads of 0.5 mm in diameter [35,36]. The cells were resuspended in disruption buffer containing 10 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 20 mM Tris–HCl (pH 7.8) for *T. coralline* or 50 mM potassium phosphate buffer (pH 7.0) for *C. magnoliae*. For proteomic analysis of cellular proteins in *C. magnoliae* by the isoelectric focusing and two-dimensional gel electrophoresis, proteins were extracted by the following procedure [37,38]. Resuspension of *C. magnoliae* cells with hot SDS sample buffer (1% sodium dodecyl sulfate (SDS) and 100 mM Tris–HCl at pH 7.0 and 95 °C) prevented protein modification and degradation and loss of high molecular mass proteins. A small amount of the cell suspension was disrupted by French® Press (Thermo Spectronic, Rochester, U.S.A.), followed by boiling for 5 min. After the addition of thiourea/urea lysis buffer (2 M thiourea, 7 M urea, 4% CHAPS, 1% dithiothreitol (DTT), and 2% carrier ampholytes; pH 3–10), the suspension was shaken moderately for 1 h and centrifuged at 16,000 rpm for 20 min. The clear supernatant was collected and used for isoelectric focusing on an immobilized pH dry strip [37].

Prefractionation of cellular organelles was proposed to reduce protein sample complexity, increase protein solubility and enhance spot detection of low level proteins in two-dimensional gel electrophoresis [39]. *C. albicans*, which is similar to *C. magnoliae*, was used to separate membrane proteins from cytoplasmic proteins [40]. The *C. albicans* cells were resuspended in a Tris buffer (pH 7.0) containing a concentrated protease inhibitor mixture (DNase and RNase A). They were disrupted in a bead beater using 425–600 μm glass beads. After separation of the supernatant from pellets, the supernatant was used for analyzing the cytoplasmic proteins. Pellets containing the membrane components were solubilzed with 6 M urea, 2 M thiourea, 4 mM tributylphosphine, 2% CHAPS, 2% *N*-decyl-*N*,*N*-dimethyl-3-ammonio-1-propane sulfate (SB 3–10), 0.5% Pharmalyte (pH 3–10) and 0.25% Pharmalyte (pH 4–6.5).

2.1.2. In gel digestion for mass spectrometric analysis

Solubilized and/or fractionated proteins are separated by two-dimensional electrophoresis techniques according to their isoelectric points and molecular weights. Several staining agents such as Commassie blue, colloidal Coomassie blue, silver nitrate or ammonia-silver complex, SyproRuby and other fluorescence dyes, radiolabeling with ^{3}H, ^{14}C, ^{35}S, ^{32}P, ^{33}P and ^{125}I, and reverse stain with zinc, copper and cobalt ions, visualize protein spots on SDS-PAGE gel. Their properties and applications to proteomics were well summarized in a recent review [41]. After staining and destaining of proteins on 2-DE, in-gel digestion in which protease, especially trypsin, cuts the peptide bonds of a protein inside SDS-polyacrylamide gel facilitates sample preparation for further mass spectrometric analysis. As an example for *C. magnoliae*, procedures of in-gel digestion and peptide extraction from 2-DE are described below [37,38]. A protein spot in silver-stained gel without glutaraldehyde treatment was sliced and destained by removing the silver bound with proteins [42]. After destaining, the gel slab was dehydrated in acetonitrile and incubated for 15 min, followed by drying in a vacuum centrifuge for 20 min. Reduction of proteins was obtained by mixing the gel pieces with 10 mM DTT in 100 mM ammonium bicarbonate at 56 °C for 1 h. After cooling at room temperature, the same volume of 55 mM iodoacetamide in 100 mM ammonium bicarbonate replaced the DTT solution. The gel pieces were washed with 50 μL of 100 mM ammonium bicarbonate for 10 min by occasional vortex mixing at ambient temperature in dark, dehydrated by addition of acetonitrile, swelled by rehydration in 100 mM ammonium bicarbonate, shrunk again by the addition of the same volume of acetonitrile, and completely dried in a vacuum centrifuge for 25 min. In-gel protein digestion was carried out by swelling the gel pieces in a digestion buffer containing 50 mM ammonium bicarbonate, 5 mM $CaCl_2$, and 12.5 ng/μL of trypsin (Boehringer Mannheim, sequencing grade) in an ice-cold bath for 45 min. After changing the supernatant with 5–10 μL of the same buffer without trypsin, the gel pieces were kept wet during enzymatic cleavage at 37 °C for 12 h. Peptides were extracted by one time change of 20 mM ammonium bicarbonate and three times changes of 5% trifluoroacetic acid in 50% acetonitrile, followed by air drying. Each extraction step was undertaken at room temperature for 20 min.

2.2. Mass spectrometric tools for C. magnoliae

In the early 1990s, DNA sequencing of cDNAs derived from pools of mRNA generated large numbers of expressed sequence tags (ESTs) [43]. Gene sequences (ESTs and others) provided a resource that could greatly accelerate protein identification by correlating the experimentally derived sequence segments with publicly available sequences in databases. Even with sequencing of genomes, the rapid identification of proteins was limited only by experimental capacity to extract the sequence information from proteins or peptides, and to correlate such information with the sequence databases. Mass spectrometry and database search algorithms rapidly filled this gap [44,45]. The high precision of mass spectrometric measurement can distinguish the peptide fragment among closely related species, and tandem mass spectrometry can provide structural information of proteins from molecular ions.

In addition to measuring peptide mass, ESI mass spectrometry can isolate specific ions from a peptide mixture on the basis of their mass-to-charge ratio (m/z) and record tandem mass spectra. Several algorithms comparing MS/MS spectra with sequence databases have greatly facilitated the mass spectrometry-based protein identification by this ap-

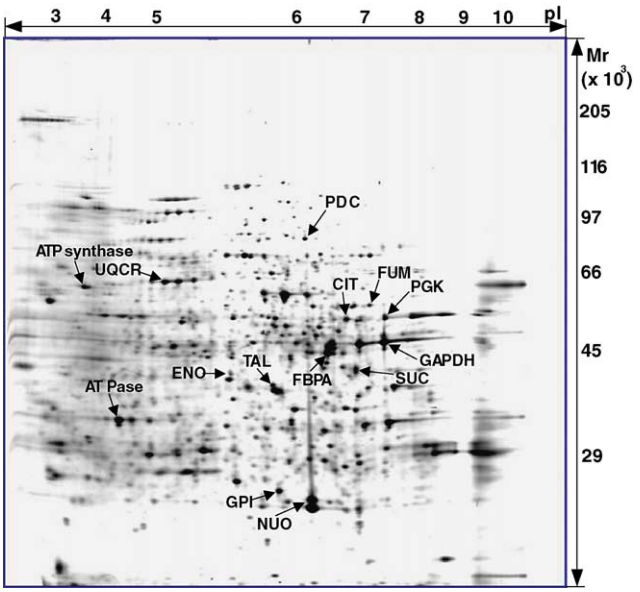

Fig. 2. Typical picture of two-dimensional electrophoresis of *C. magnoliae* stained with silver nitrate. Arrows indicate metabolic enzymes involved in glucose metabolisms such as glycolysis, citric acid cycle, pentose phosphate pathway and oxidative phosphorylation. GPI, glucose-6-phosphate isomerase; FBPA, fructose bisphosphate aldolase; GAPDH, glyceraldehydes-3-phosphate dehydrogenase; PGK, phosphoglycerate kinase; ENO, enolase; PDC, pyruvate decarboxylase; CIT, citrate synthase; SUC, succinyl-CoA ligase; FUM, fumarase; TAL, transaldolase; NUO, NADH:ubiquinone oxidoreductase; UQCR, ubiquinol:cytochrome *c* reductase.

proach [46,47]. Because the sequences of a peptide fragment can be used to identify a protein, the specificity of MS/MS-based protein identifications is often much higher than that of peptide mass mapping. MS/MS spectra are also ideally suited to search translated EST and other sequence databases containing incomplete sequences. The complexity of the peptide mixtures by tryptic digestion of proteins needs additional separations. ESI-MS/MS combined with liquid chromatography such as ion exchange, gradient reversed-phase and smaller diameter capillary types increases separation efficiencies (peak capacities of ~1200) [48,49]. The proteome for erythritol-producing *C. magnoliae* was separated in 2-DE gel according to p*I* and molecular weight (Fig. 2). After tryptic digestion in gels, the peptide mixtures were analyzed by MALDI-TOF, ESI-MS/MS and de novo sequencing. The methods of mass spectrometric analysis of *C. magnoliae* proteome and of identification of protein spots were presented briefly [37,38].

2.2.1. MALDI-TOF

The in-gel tryptic digest of *C. magnoliae* proteome was analyzed on the Voyager-DE STR MALDI-TOF mass spectrometer (Applied Biosystems, Framingham, MA, USA) using the thin-layer method in which the matrix solution contained α-cyano-4-hydroxycinnamic acid dissolved in a nitrocellulose in acetone/isopropanol solvent [38]. After washing the peptide dot on the dried matrix material with 0.1% trifluoroacetic acid, the delayed extraction modes at 20 kV of accelerating voltage, 150 ns delay, and desorption and ionization modes using 337 nm nitrogen laser gave the MS spectra of the tryptic peptides.

MALDI-TOF mass spectrometry and peptide mass fingerprint search using fungi databases were applied to analyze 24 protein spots of *C. magnoliae* on 2-DE [38]. Four protein spots contained the same peptide sequence of WWSLVPLGR, which was a tryptic peptide of carbonyl reductase with the peptide sequence coverage higher than 30%. One of the MS spectra was shown in Fig. 3. Two spots analyzed by MASCOT were identified to be malate dehydrogenase of *Talaromyces emersonii* and the hypothetical protein 8D4.190 [imported] of *Neurospora crassa*. Most of the protein samples could not be identified by MALDI-TOF analysis alone due to low sequence coverage. When protein similarity between species is less than 70% identity, peptide mass fingerprinting (PMF) alone was not suitable for divergent species jumps [50]. In case of *C. magnoliae* with unsequenced genome, peptide sequencing of tandem mass spectrometry is more specific for protein identification than simple peptide mass mapping.

2.2.2. ESI-MS/MS

A liquid chromatography/mass spectrometry (LC/MS) system to analyze protein fragments of *C. magnoliae* consists of a nano-LC system (LC Packings, Netherlands) and a quadrupole time-of-flight mass spectrometer (Q-TOF2, Micromass, U.K.) complemented with an electrospray ionization (ESI) source [37]. Desalting and concentration of peptides using a C_{18} precolumn (i.d. 300 μm, length 1 mm, particle size 5 μm; LC Packings) proceeded their separation by a C_{18} nano-column (i.d. 75 μm, length 150 mm, particle size 5 μm; LC Packings). For coupling the nanospray with the on-line nano-LC, the end of the capillary tube from the nano-LC column was connected with the emitter with a pico-tip silica tube (i.d. 5 μm, New Objectives, U.S.A.). Electrospray condition of the union was 1.5–2 kV and cone voltage was 30 V. Argon as a collision gas was introduced at a pressure of 68.9 kPa. At collision energy increasing stepwise to 25, 30 and 35 eV, MS/MS spectra were obtained in a data dependent MS/MS mode. Mascot (Matrix science, U.K.) or manually sequenced by Masslynx software (Micromass, U.K.) was used for protein identification from MS/MS spectra. For example, an MS/MS spectrum shown in Fig. 4 led to identification of an amino acid fragment, VALTGLTVAEYFR, which was present in F1-ATPase δ and ε subunits. Among 52 proteins of *C. magnoliae* identified by ESI-MS/MS and de novo sequencing, the information of enzymes involved in carbon and energy metabolic pathways is summarized in Table 1 and their protein spots are shown in Fig. 2.

2.3. Similarity search

The goal of database searching is to be able to quickly and accurately identify large numbers of proteins. The success of database searching depends on the quality of the data obtained in mass spectrometry, the quality of the database

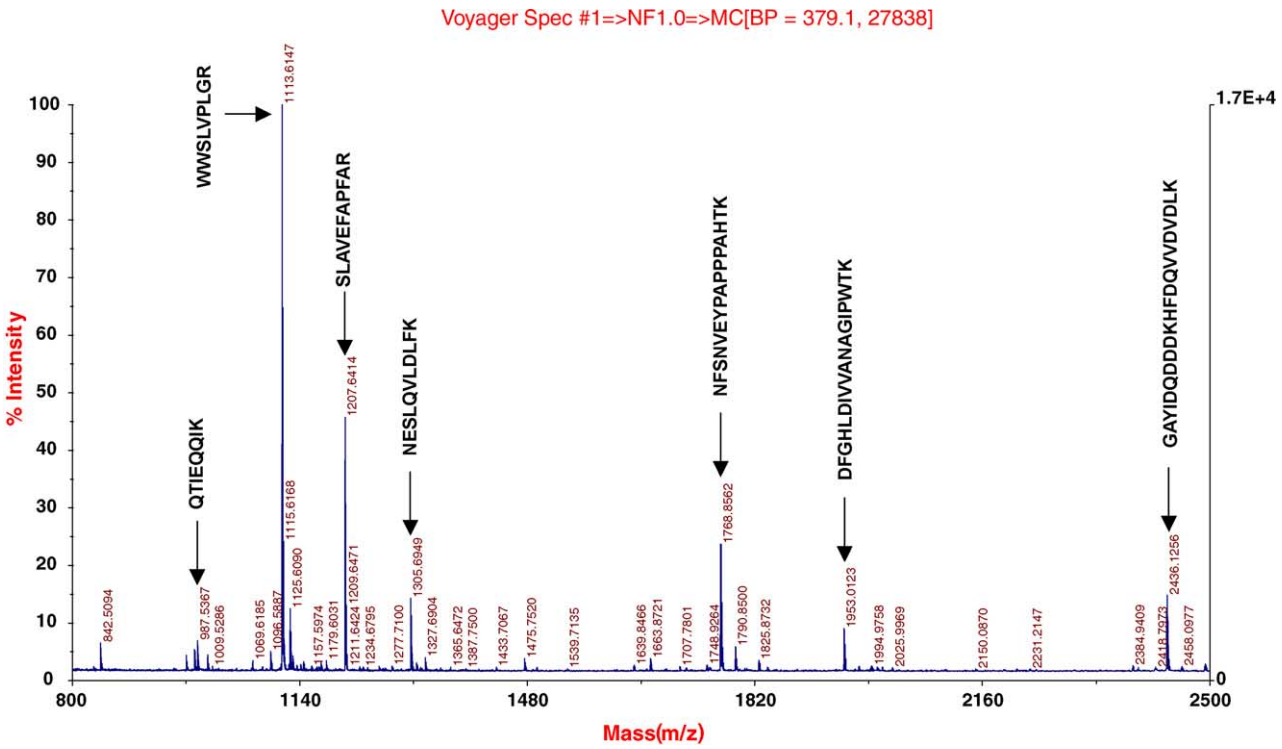

Fig. 3. An MS spectrum obtained with MALDI-TOF mass spectrometry. A trypsin-digested amino acid fragment with 1113.6147 *m*/*z* was identified as WWSLVPLGR present in carbonyl reductase of *C. magnoliae*.

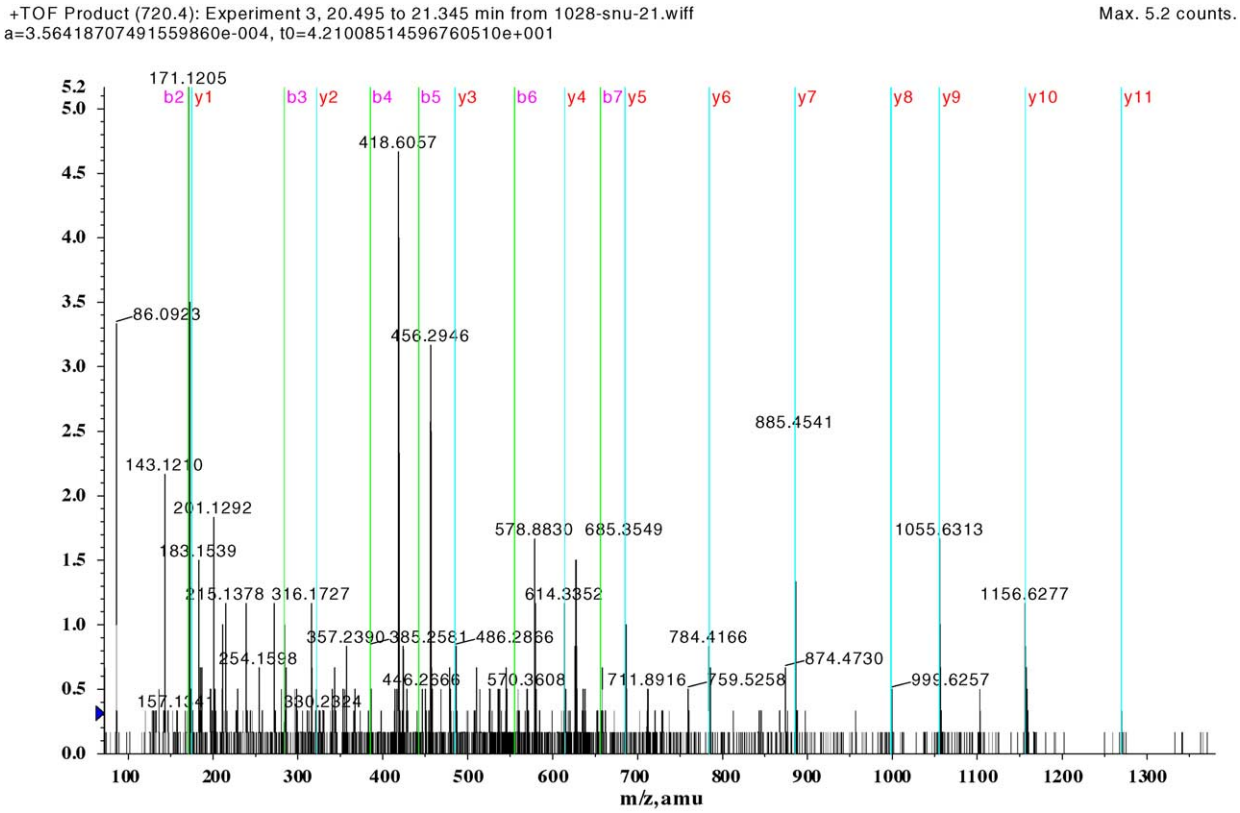

Fig. 4. Analysis of a peptide spectrum by ESI-MS/MS resulting in VALTGLTVAEYFR. F1-ATPase δ and ε subunits of *S. cerevisiae* contain this peptide sequence.

Table 1
C. magnoliae metabolic enzymes identified using proteomic technologies [38]

pI/Mr	Enzyme	Role in metabolism	Peptides	Type of analysis
Glycolysis				
5.8/28,000	Glucose-6-phosphate isomerase	Isomerization of glucose-6-phosphate to fructose-6-phosphate	EFSDAVR	MS/MS
6.2/43,000	Fructose bisphosphate aldolase	Splitting of fructose-1,6-bisphosphate into glyceraldehyde-3-phosphate and dihydroxyacetone phosphate	NSPGAAGGGFASLAGAVAAATAGAPR, LSNNQMASFLDK, VNLDTDTQYAYLGGGGAR	de novo
7.0/45,000	Glyceraldehyde-3-phosphate dehydrogenase	Conversion of glyceraldehyde-3-phosphate into 1,3-bisphosphoglycerate, NADH formation	VVDLVLELAAK, VLSWYDNEYGYSAR, VMGFVPGPYVVSTD-FVFAGGAMSSGGGAAK	de novo
7.0/53,000	Phosphoglycerate kinase	Substrate-level phosphorylation (ATP formation) by dephosphorylation of 1,3-bisphosphoglycerate	VDFNVPLDGKT, LSHVSTGGGASLELLEGK, ALENPERPFLAILGGAK	MS/MS
5.4/42,000	Enolase	Formation of phosphoenolpyruvate	ALLGVAVAAAR, HSTGSWALELR, LNQLLR, IEEELGDK, IGLDCASSEFFK, LGANAILGVSMAAAR	de novo, MS/MS
6.3/75,000	Pyruvate decarboxylase	Decarboxylation of pyruvate in alcohol metabolism	AALNDAFATR, EEPTVPLGTYLLYK, MSAPVNAAYIDR, QNVVVTAVGGGAYK	de novo
Citric acid cycle				
6.7/55,000	Citrate synthase	Condensation of acetyl-CoA and oxaloacetate	YLWDTLNSGR, AIGVLPQLIIDR	MS/MS
6.7/43,000	Succinyl-CoA ligase	Nucleotide-dependent conversion of succinyl-CoA to succinate	VLFQGFTGK, LVGPNCPGIIAPGQCK	MS/MS
6.8/58,000	Fumarase	Transhydration of fumarate to L-malate	QGGTAVGT	de novo
Pentose phosphate pathway				
5.8/39,000	Transaldolase	Transferring of a keto group from ketose to aldose	ILDWYK, TIVMGASFR, ATGTTVVADTGDFESIAK	MS/MS
Oxidative phosphorylation				
6.0/28,000	NADH:ubiquinone oxidoreductase	Electron transfer from NADH to coenzyme Q	YDGFLLRPTR, AFSVDATGVAWAK, LSVFQAMLTEPAAGGALSK	de novo
4.8/62,000	Ubiquinol:cytochrome *c* reductase	Electron transfer from coenzyme Q to cytochrome *c*	ENTVYFAK, ASILLSLDGTTAVAEDIGR, EAVLELVK, DAGTFDMREEIYGK	MS/MS
3.1/60,000	ATP synthase β chain	Direct catalysis of ATP formation	TVFIQELINNIAK, LVLEVAQHLGENTVR, SLQDIIAILGMDELSEQDKLTVER, GSVTSVQAVYVPADDLTDPAPAT	MS/MS
4.0/35,000	F1-ATPase δ and ε subunits		IGLFGGAGVGK, VVDLLAPYAK, FTQAGSEVSALLGR, VALTGLTVAEYFR, VALVYGQMNEPPGAR	MS/MS

searched, and the method used to search the database. The ExPASy server provides convenient tools for proteomics and programs for protein identification [51]. Various algorithms for protein identification from mass spectrometric analysis are publicly available on the World Wide Web, for example Mascot, MS-Fit, ProFound and SEQUEST [52]. The critical disadvantage of performing tandem mass spectrometry is a lack of flexibility in the search programs. If a single mistake is made in the assignment of a y- or b-ion, the amino acid sequence would be incorrect and the database search bring up irrelevant proteins. Often it is necessary to confirm that the sequence obtained in the mass spectrometer matches the peptide sequence obtained from the database. This can be done by performing a theoretical fragmentation of the peptide from the database and comparing the two mass spectra. Another approach for protein identification is to obtain de novo sequence data from peptides by MS/MS and then to use all the peptide sequences to search appropriate databases. The key advantage of this method is the capability of searching peptide sequence information across both DNA and protein databases. This search method is particularly useful for organisms with insufficient genomic and proteomic databases. Since this method requires peptide information of three or four amino acids, it may not be the first choice for peptide identification. Rather, the much faster methods of peptide mass fingerprinting or peptide ion searching can be used first. If these search methods are not sufficient, de novo sequence information can be used to identify the protein of interest.

A large number of programs are now available for the identification of proteins by using uninterpreted MS/MS

data, for example, Mascot and SEQUEST [53]. However, searches against unannotated or untranslated DNA databases with uninterpreted MS/MS data have problems with correct identification of a protein, due to polymorphisms, sequencing errors and conservative substitutions. The development of uninterpreted MS/MS search algorithms that are error tolerant may overcome some of these handicaps. For *C. magnoliae*, information on the peptide mass fingerprinting, MS/MS spectra, isoelectric point and molecular weight estimated from 2-DE, MALDI-TOF and ESI-MS/MS were used to search for the protein homology. Peptide fragments were identified by correlation of uninterpreted tandem mass spectra to entries in the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/BLAST/), *Saccharomyces* Genome Database (SGD, http://genome-www.stanford.edu/Saccharomyces/) and sequence data of *C. albicans* at the Stanford Genome Technology Center (http://www-sequence.stanford.edu/group/candida) using MS/MS Ion Search on Mascot (http://www.matrixscience.com). To increase the alignment score, NCBI protein blast search for *short nearly exact matches* was undertaken with peptide-mapping data.

2.4. Proteomic application to erythritol metabolism in *C. magnoliae*

High throughput analysis of proteome using two dimensional electrophoresis and mass spectrometry-based database searching makes an advance to predict and elucidate the metabolic network of carbon metabolism and stress response under various environmental conditions. Applications of proteomics to microbial metabolisms were achieved in the response of the perturbed metabolic network of yeast galactose metabolism [54], analysis of metabolic enzymes in the recombinant xylose-fermenting *S. cerevisiae* [55,56], optimization of high cell density cultivation of *Escherichia coli* [57], glucose starvation of *Bacillus subtilis* [58], characterization of immunogenic enzymes engaged in carbon metabolism of *C. albicans* [59] and prediction of the metabolic network for erythritol production by *C. magnoliae* [37]. To date, several *C. magnoliae* enzymes involved in carbon metabolisms such as glycolysis, citric acid cycle, pentose phosphate pathway and oxidative phosphorylation were identified using proteomic technologies (Table 1). An effort to compare the wild type of *C. magnoliae* with its mutant overproducing erythritol was made to explain the metabolic changes by chemical mutation [37]. Down-regulation of enolase in the mutant probably directed the flux of erythrose-4-phosphate to erythritol production rather than shikimate biosynthesis. Up-regulation of fumarase in the mutant coincided with the result in the inhibitory effect of fumarate on erythritol production [60].

Information on the metabolic pathway involved in the formation of erythritol in osmophilic yeasts or bacteria is not sufficient to explore the complete network at a protein or gene level. Enzymatic analysis involved in erythritol formation from glucose proposed the erythritol metabolism in a heterolactic bacterium, *Oenococcus oeni* (formerly *Leuconostoc oenos*) [61]. Phosphoketolase [EC 4.1.2.22] expressed in *O. oeni* catalyzes the phospho-cleavage reaction of fructose 6-phosphate into acetyl phosphate and erythrose 4-phosphate, which is an intermediate in the erythritol metabolic pathway. On the basis of these results, the pathway was proposed for the formation of erythritol from glucose 6-phosphate by *O. oeni* (Fig. 5). It is unclear which enzyme or reaction is engaged in the final step of

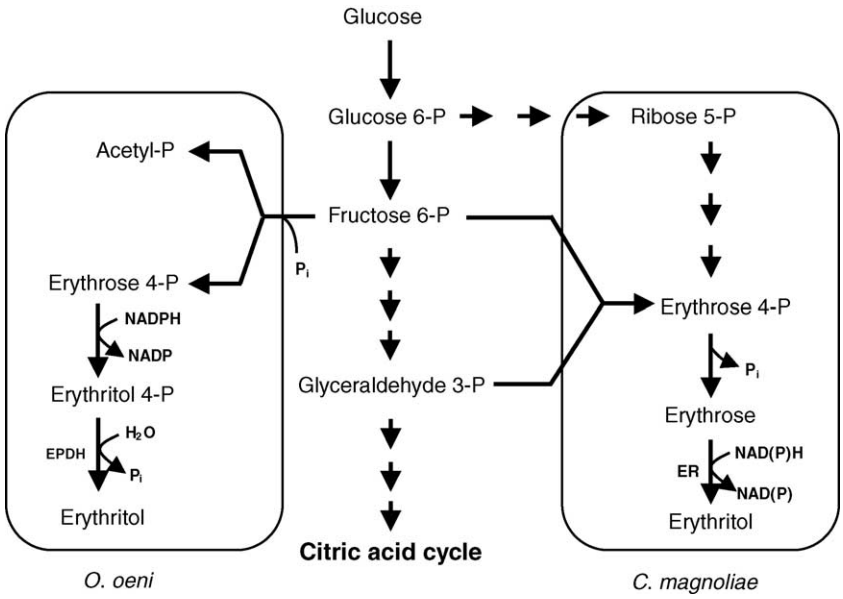

Fig. 5. Erythritol metabolisms of *O. oeni* (left side) and *C. magnoliae* (right side). EPDH, erythritol 4-phosphate dehydrogenase; ER, erythrose reductase; P_i, inorganic phosphate.

erythritol production. In fact, (i) the dephosphorylation rate of erythritol 4-phosphate hydrolysis in extracts was more than twice as high as that of erythrose 4-phosphate hydrolysis, and (ii) the dehydrogenase acting on erythrose 4-phosphate was much more active than the enzyme acting on erythrose. In *O. oeni*, erythrose 4-phosphate is supposed to be dehydrogenated into erythritol 4-phosphate, followed by dephosphorylation of erythritol 4-phosphate to erythritol.

In contrast to the bacterial strains producing erythritol, eukaryotes contain erythrose reductase to catalyze the hydrogenation of erythrose. *T. corallina* isolated from high concentrated sucrose solution expressed erythrose reductase with 7.12 mM of K_m and 26 μmol/min mg protein of V_{max} [60]. Fumarate and 1.8-dihydroxynaphthalene-melanin inhibited the activity of erythrose reductase in an uncompetitive and non-competitive type, respectively [37,60]. Characterization of erythrose reductase in *C. magnoliae* indicated that it was one of the aldose reductase with a high preference for NADH in contrast to the typical aldose reductase acting with NADPH [36]. Recently, our laboratory identified the nucleotide sequence of erythrose reductase in *C. magnoliae* with N-terminal homology to the typical aldose reductase (data not shown). The biosynthetic mechanism of erythritol in *C. magnoliae* is suggested that 1 mol of glucose is converted to 1 mol of erythrose-4-phosphate via the pentose phosphate pathway, and then erythrose-4-phosphate is dephosphorylated into erythrose, which is reduced into erythritol by erythrose reductase (Fig. 5).

3. Conclusion

Erythritol is a biologically manufactured sweetener with good characteristics for food and pharmaceutical application. Production of erythritol has focused on the mutation and selection of high erythritol-producing microbial mutants, characterization of these microorganisms at a protein level and optimization of fermentation technologies at a reactor level. The proteome of erythritol-producing microorganisms was analyzed with two-dimensional electrophoresis, mass spectrometry and similarity searching. Proteomic methods for the erythritol-producing microorganisms were introduced in details and especially, 55 proteins in *C. magnoliae* were identified by MALDI-TOF and ESI-MS/MS techniques. Experimental results obtained by proteomic and enzymatic approaches allowed the plausible explanation for erythritol biosynthesis and will eventually be used to genetically modify the erythritol-producing microorganisms. The proteomic technology developed for *C. magnoliae* will also be applied to other industrial microorganisms whose genomic, proteomic and enzymatic databases are insufficient.

Acknowledgement

This study was supported by Ministry of Education and Human Resources Development through BK21 program and Ministry of Science and Technology through the 21st Century Frontier Microbial Genomics and Application Center Program (Grant MG 02-303-002-1-0-0).

References

[1] P. O'Farrell, J. Biol. Chem. 250 (1975) 4007.
[2] V.C. Wasinger, S.J. Cordwell, A. Cerpa-Polijak, J.X. Yan, A.A. Gooley, M.R. Wilkins, M.W. Duncan, R. Harris, K.L. Williams, I. Humphery-Smith, Electrophoresis 16 (1995) 1090.
[3] K.H. Lee, Trends Biotechnol. 19 (2001) 217.
[4] D.J. Cahill, E. Nordhoff, J. O'Brien, J. Klose, H. Eickhoff, H. Lehrach, in: S.R. Pennington, M.J. Dunn (Eds.), Proteomics: From Protein Sequence to Function, Bios Scientific Publishers Ltd., Oxford, UK, 2001, p. 1, Chapter 1.
[5] M.J. Dutt, K.H. Lee, Curr. Opin. Biotechnol. 11 (2000) 176.
[6] T. Shindou, H. Ishizuka, Food Sci. Technol. 2 (1996) 82.
[7] T. Niwa, K. Tohyama, Y. Kato, J. Chromatogr. 613 (1993) 914.
[8] G.P. Roberts, A. McDiarmid, P. Gleed, Res. Vet. Sci. 20 (1976) 254.
[9] J. Goossens, H. Röper, Int. Food Ingred. 1–2 (1994) 27.
[10] P. Cock, World Rev. Nutr. Diab. 85 (1999) 110.
[11] W.O. Bernt, J.F. Borzelleca, G. Flamn, I.C. Munro, Regul. Toxicol. Pharmacol. 24 (1996) S191.
[12] J. Goossens, H. Röper, Confectionery Product. 24 (1994) 182.
[13] K.K. Mäkinen, K.P. Isotupa, T. Kivilompolo, P.L. Mäkinen, J. Toivanen, E. Söderling, Caries Res. 35 (2001) 129.
[14] M. Hiele, Y. Ghoos, P. Rutgeerts, G. Vantrappen, Br. J. Nutr. 69 (1993) 169.
[15] N.J. Chü, C.E. Ballow, J. Am. Chem. Soc. 83 (1961) 1711.
[16] W. Maret, D.S. Ault, Biochem. J. 27 (1988) 1622.
[17] Mitsubishi and Nikken, Private communication from Mitsubishi Chemical Corp. and Nikken Chemical Co., 1996, in press.
[18] U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, Office of Food Additives Safety: Agency Response Letter: GRAS notice No. GRN 000076, September 2001.
[19] V.F. Pfeifer, V.E. Sohn, H.E. Conway, E.B. Lancaster, S. Dabic Jr., E.L. Griffin, Ind. Eng. Chem. 52 (1960) 201.
[20] Cerestar Holding BV, Mitsubishi Chemical Corporation and Nikken Chemicals Co. Ltd, Application for assessment of erythritol prior to its authorization, 1999.
[21] C. Larsson, C. Morales, L. Gustafsson, L. Adler, J. Bacteriol. 172 (1990) 1769.
[22] K. Tokuoka, T. Ishitani, W.C. Chung, J. Gen. Appl. Microbiol. 38 (1992) 35.
[23] A.D. Brown, Adv. Microb. Physiol. 17 (1978) 181.
[24] H. Onishi, Y. Shiromaru, FEMS Microbiol. Lett. 25 (1984) 175.
[25] H. Ishizuka, K. Wako, T. Kasumi, T. Sasaki, J. Ferment. Bioeng. 68 (1989) 310.
[26] J.B. Park, B.C. Seo, J.R. Kim, Y.K. Park, J. Ferment. Bioeng. 86 (1998) 577.
[27] S.Y. Kim, S.R. Jung, D.K. Oh, J.K. Lee, Korea Patent 0364105 (2002).
[28] J.H. Seo, Y.W. Ryu, S.R. Jung, S.Y. Kim, U.S. patent 6,287,830 B1 (2001).
[29] K.H. Lee, J.H. Seo, Y.W. Ryu, Kor. J. Biotechnol. Bioeng. 17 (2002) 509.
[30] S.W. Yang, J.B. Park, N.S. Han, Y.W. Ryu, J.H. Seo, Biotechnol. Lett. 21 (1999) 887.
[31] E.S. Koh, T.H. Lee, D.Y. Lee, H.J. Kim, Y.W. Ryu, J.H. Seo, Biotechnol. Lett. 25 (2003) 2103.
[32] J.H. Choi, M.D. Kim, J.H. Seo, J.W. Ahn, Korean J. Food Sci. Technol. 35 (2003) 708.
[33] G.M. Walker (Ed.), Yeast Physiology and Biotechnology, John Wiley & Sons Ltd., Chichester, 1998, p. 17, Chapter 2.

[34] H. Ishizuka, K. Tokuoka, T. Sasaki, H. Taniguchi, Biosci. Biotech. Biochem. 56 (1992) 941.
[35] J.K. Lee, H.M. Jung, S.Y. Kim, Appl. Environ. Microbiol. 69 (2003) 3427.
[36] J.K. Lee, S.Y. Kim, Y.W. Ryu, J.H. Seo, J.H. Kim, Appl. Environ. Microbiol. 69 (2003) 3710.
[37] D.Y. Lee, Y.C. Park, H.J. Kim, Y.W. Ryu, J.H. Seo, Proteomics 3 (2003) 2330.
[38] H.J. Kim, D.Y. Lee, D.H. Lee, Y.C. Park, D.H. Kweon, Y.W. Ryu, J.H. Seo, Proteomics 4 (2004) 3588.
[39] A. Pitarch, M. Sánchez, C. Nombela, C. Gil, J. Chromatogr. B 787 (2003) 101.
[40] J.M. Bruneau, I. Maillet, E. Tagat, R. Legrand, F. Supatto, C. Fudali, J.P. Le Caer, V. Labas, D. Lecaque, J. Hodgson, Proteomics 3 (2003) 325.
[41] W.F. Patton, J. Chromatogr. B 771 (2002) 3.
[42] F. Gharahdaghi, C.R. Weinberg, D.A. Meagher, B.S. Imai, S.M. Mische, Electrophoresis 20 (1999) 601.
[43] M.D. Adams, Science 252 (1991) 1651.
[44] A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S.G. Oliver, Science 274 (1996) 563.
[45] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, K. Baldwin, K. Devon, et al., Nature 409 (2001) 860.
[46] J.K. Eng, A.L. McCormack, J.R. Yates III, J. Am. Soc. Mass Spectrom. 5 (1994) 976.
[47] R. Gras, M. Muller, Curr. Opin. Mol. Ther. 3 (2001) 526.
[48] M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, M. Mann, Nature 379 (1996) 466.
[49] R.D. Smith, Trends Biotechnol. 20 (2002) 3.
[50] P.J. Lester, S.J. Hubbard, Proteomics 2 (2002) 1392.
[51] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Methods Mol. Biol. 112 (1999) 531.
[52] D. Lin, D.L. Tabb, J.R. Yates III, Biochim. Biophys. Acta 1646 (2003) 1.
[53] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (1999) 3551.
[54] T. Ideker, V. Thorsson, J.A. Ranish, R. Christmas, J. Buhler, J.K. Eng, R. Bumgarner, D.R. Goodlett, R. Aebersold, L. Hood, Science 292 (2001) 929.
[55] Y.S. Chung, M.D. Kim, W.J. Lee, Y.W. Ryu, J.H. Kim, J.H. Seo, Enzyme Microb. Technol. 30 (2002) 809.
[56] L. Salusjärvi, M. Poutanen, J. Pitkänen, H. Koivistoinen, A. Aristidou, N. Kalkkinen, L. Ruohonen, M. Penttilä, Yeast 20 (2003) 295.
[57] S.H. Yoon, M.J. Han, S.Y. Lee, K.J. Jeong, J.S. Yoo, Biotechnol. Bioeng. 81 (2003) 753.
[58] R. Schmid, J. Bernhardt, H. Antelmann, A. Völker, H. Mach, U. Völker, M. Hecker, Microbiology 143 (1997) 991.
[59] M. Pardo, M. Ward, A. Pitarch, M. Sánchez, C. Nombela, W. Blackstock, C. Gil, Electrophoresis 21 (2000) 2651.
[60] J.K. Lee, B.S. Koo, S.Y. Kim, Appl. Environ. Microbiol. 69 (2002) 4534.
[61] M. Veiga-Da-Cunha, H. Santos, E.V. Schaftingen, J. Bacteriol. 175 (1993) 3941.