Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Chromatography A, 1082 (2005) 25–32
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Bacterial response to eukaryotic cells
Analysis of differentially expressed proteins using nano liquid chromatography–electrospray ionization tandem mass spectrometry

C. Alpert [a,*], W. Engst [a], A. Guehler [a], T. Oelschlaeger [b], M. Blaut [a]

[a] *German Institute of Human Nutrition Potsdam-Rehbruecke, Department of Gastrointestinal Microbiology, A.-Scheunert-Allee 114–116, D-14558 Nuthetal, Germany*
[b] *Wuerzburg University, D-97070 Wuerzburg, Germany*

Available online 11 April 2005

Abstract

Host–bacteria interactions have mostly been investigated with regard to the host response or to activities of pathogenic bacteria. In contrast, we aim to identify reactions of non-pathogenic bacteria that result from their contact with host cells of the gastrointestinal tract. In a proteomic approach, the response of non-pathogenic human *Escherichia coli* bacteria on gut epithelial cells (rat IEC-6) was investigated in an in vitro co-culture model. For this purpose, a sensitive analytical procedure was developed based on the identification of two-dimensional polyacrylamide gel electrophoresis separated proteins by online nanoLC–electrospray ionization MS/MS using a quadrupole time-of-flight tandem mass spectrometer for accurate mass determination. We demonstrate here the efficiency of this technique by the identification of a total of 43 differentially expressed proteins, out of which 25 were up-regulated and 18 were down-regulated. They represent a wide range of molecular weight and different metabolic and physiological functions.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Host–bacteria interaction; Proteomics; NanoLC–MS/MS

1. Introduction

The human digestive tract is a complex ecosystem, where several hundred bacterial species interact with each other, food components and the host [1]. The gut microbiota fulfills a number of functions in host physiology but is also presumed to contribute to the development of diseases, such as inflammatory bowel disease, colon cancer and allergies. While the influence of the bacteria on the host has been repeatedly demonstrated, the host may also have an influence on the bacteria. At present, knowledge about these interactions is very scarce. Host–bacteria interactions have mostly been investigated with regard to the host response or to activities of pathogenic bacteria. In contrast, we aim to identify reactions of non-pathogenic bacteria, which result from their contact with host cells, as they occur in the gastrointestinal tract by means of the detection of differentially expressed proteins.

In a proteomic approach, the response of non-pathogenic human *Escherichia coli* cells to gut epithelial cells (rat IEC-6) was investigated using an in vitro co-culture model system, two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) for protein separation and an online nanoLC–electrospray ionization tandem mass spectrometry (ESI-MS/MS) technique for subsequent identification of differentially expressed proteins. Meanwhile, 2D-PAGE is an established method for the analysis of proteomes from various sources [2]. Usually, samples are taken from a single type of tissue or a mono-bacterial culture, where easily defined changes or signals such as varying pH, salt, temperature or nutrient composition are applied. In contrast, in our experiments, the presence or absence of gut epithelial cells may represent a complex set of signals, ranging from competition for some nutrients and the secretion of metabolites to the simple presence of epithelial surface structures.

* Corresponding author. Tel.: +49 33200 88 440; fax: +49 33200 88 407.
E-mail address: alpert@mail.dife.de (C. Alpert).

0021-9673/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2005.03.094

In many proteomics projects, analysis of differentially displayed protein profiles is done by peptide mass fingerprinting using matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) or de novo peptide sequencing using ESI-MS/MS [3]. Recent developments of nanoLC–ESI-MS/MS systems permit sensitive analysis of proteins isolated from spots of interest including the separation of peptide mixtures obtained from tryptically digested proteins by nanoLC followed by sequencing of peptides by ESI-MS/MS and subsequent protein identification by databank searches based on this sequence information [3–5].

Protein identification by MALDI-MS is also an excellent technique in routine proteomic analysis, but it is limited to samples from organisms, for which complete genomic information is available. Since there is no genomic information available at present for the wild-type strain which we used in these experiments, and which therefore might express proteins with up to now unknown sequences and functions, we developed an analytical procedure based on nanoLC followed by ESI-MS/MS for the characterization of differentially expressed proteins. In addition, we expect to eventually apply these methods to other human gut bacteria, which have not yet been sequenced, as well as to in vivo derived samples.

2. Materials and methods

2.1. Biological material and techniques

2.1.1. Bacterial strain and cell line

The *E. coli* strain UNC and the rat epithelial cell line IEC-6 were a kind gift of Dirk Haller (Technical University Munich). *E. coli* was initially isolated from a patient with Crohn's disease by the Clinical Microbiology Laboratory of the University of North Carolina Hospitals, Chapel Hill, NC, USA and described by Rath et al. [6,7].

2.1.2. Characterization of E. coli UNC

The bacteria were characterized with respect to pathogenicity markers by multiplex PCR using total bacterial DNA as a template. Five sets of six reactions each were run in order to detect the following markers of pathogenic potential: pool 1—*ibeA*, *papEF*, *kpsMT* III, *fimH*, *papA*, PAI; pool 2—K1, *papG* allele III, *iutA*, *sfa/focDE*, *bmaE*, *fyuA*; pool 3—*papC*, *kpsMT* II, *papG* allele I, *nfaE*, *rfc*, *hlyA*; pool 4—*papG* allele II, *traT*, *focG*, *cdtB*, *cva*C, *gafD*; pool 5—K5, *sfaS*, *cnfI*, *ava/draBC*, *papG* II/III, *papG* I [8]. For positive controls, markers were amplified from DNA of strains J96 and IA2 [9,10] as well as strains U5, V15, PM9, V27, P42 and U8 (strain collection of the Institute for Molecular Biology of Infectious Diseases, University Würzburg). Corresponding reactions with water added instead of template DNA were used as negative controls.

2.1.3. Cell culture

IEC-6 epithelial cells from the rat small intestine were cultured in D-MEM (Gibco 41965-039; 4.5 mg/l glucose) containing 0.1 U/ml insulin (Gibco 13007-018), 5% (v/v) fetal calf serum (Sigma 120K3359) and 1% (v/v) antibiotic–antimycotic (Gibco 15240-096) at 37 °C in T75 cell culture flasks (No. 9076, Biochrom, Berlin, Germany). The atmosphere in the incubator was maintained at 5% (v/v) CO_2 and 60% relative humidity.

2.1.4. Bacterial culture

For maintenance purposes, bacteria were cultured aerobically in LB broth or on LB agar (International Diagnostics Group, UK) at 37 °C, while bacterial stocks were kept as glycerol cultures at −20 °C. For experiments, bacteria were pre-cultured aerobically overnight at 37 °C in D-MEM media (4500 mg/l glucose, Gibco).

2.1.5. Co-culture experiments

For co-cultures, T75 flasks containing a confluent lawn of IEC-6 cells were washed twice with 1× PBS (pH 7.4) (Biochrom) and bacteria from pre-cultures resuspended in fresh D-MEM were added at a ratio of 100:1 bacteria to epithelial cells. Co-cultures were incubated at 37 °C in an atmosphere enriched with 5% CO_2. Bacteria were withdrawn after 4 h of co-incubation by draining the liquid media from the culture flask. For controls, bacteria were cultured in the absence of IEC-6 cells.

2.1.6. 2D-PAGE

Bacteria were harvested by centrifugation, washed in 0.5 ml of 100 mM Tris–1 mM EDTA (pH 7.5), 1 mM PMSF, 100 μg/ml chloramphenicol, centrifuged and resuspended in 0.5 ml wash buffer in a 1.5 ml reaction tube. Cells were subsequently disrupted by shaking for 15 min on a Uniprep speedsetting 3 (UniEquip, Martinsried, Germany) with 1.2 ml glass beads (Roth, 0.25–0.5 mm diameter). Cell debris was removed by centrifugation for 15 min at 14 000 × *g*. The resulting supernatant was recovered and the protein concentration determined according to the method of Bradford, using bovine serum albumin as a reference protein [11]. Aliquots containing 500 μg of protein were mixed with 62.5 U of Benzonase, concentrated on Microcon-3 filters (Amicon), added to a mix of DeStreak rehydration solution and IPG buffer prepared according to the instructions of the manufacturer and applied to 13 cm IPG strips (pH 4–7, linear gradient) (all Amersham Biosciences) in a total volume of 250 μl for passive overnight rehydration in an Immobiline DryStrip reswelling tray. For isoelectric focusing, strips were mounted in an Immobiline strip tray on a Pharmacia LKB Multiphor II cooling plate. The focusing chamber was maintained at 20 °C, and the applied voltage was raised from 100 to 3500 V for a total of 17 650 V h. Focused proteins were denatured and reduced by shaking the IPG strip for 30 min in 5 ml sodium dodecyl sulfate (SDS) equilibration buffer (50 mM Tris–HCl (pH 8.8), 6 M urea, 30% glycerol, 2% SDS, 10 mg/ml dithiothreitol (DTT)). Denatured proteins were subsequently oxidized by incubation for 30 min in equilibration buffer containing 25 mg/ml iodoacetamide instead of DTT. Separation

in the second dimension was on 12.5% T acrylamide–0.1% SDS gels in a Bio-Rad Protean II XI cell. Proteins were detected by hot Coomassie staining: Coomassie R-350 in 10% acetic acid was heated to 90 °C and poured over the gels. After shaking for 10–30 min, the gels were destained in 10% acetic acid [2]. For documentation, the gels were scanned on a Bio-Rad GS-800 calibrated densitometer. Spots of interest were identified by direct comparison between control and co-incubation samples, cut from the gel and digested with trypsin by the method of Schrattenholz [12]. The peptide mixtures were dissolved in 15 μl of 0.1% formic acid and analyzed on the LC–MS system (usually 0.5–1 μl).

2.2. NanoLC–ESI-MS/MS

Data were acquired using a Waters CapLC system (Milford, MA, USA) and a Waters Micromass quadrupole time-of-flight (QTOF) Ultima API hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Manchester, UK) fitted with a nanospray ion source.

The CapLC system consisted of a Waters CapLC pump, autosampler with cooled sample tray and a stream select module attached directly to the nanoflow interface. This module allowed the preconcentration and online desalting of the sample prior to separation. It was configured with a 0.30 mm × 5 mm Waters Symmetry300 C_{18} OPTI-PAK trap column and a 150 mm × 0.075 mm Waters Atlantis NanoEase C_{18} analytical column.

Sample loading and desalting was conducted at a flow rate of 20 μl/min with aqueous 0.1% formic acid. After 8 min, the stream select valve was switched to connect the precolumn to the separating column. A flow rate of 200 nl/min was adjusted by splitting the pump delivery flow rate of 5.5 μl/min.

For elution and separation of the peptides, an increasing organic solvent concentration from 5 to 95% B was applied (solvent A: aqueous 0.1% formic acid/acetonitrile (95:5, v/v); solvent B: aqueous 0.1% formic acid/acetonitrile (20:80, v/v)). All solvents were of HPLC gradient grade. The gradient program started with a linear gradient from 5 to 60% B within 38 min. Then, the ratio was changed within 2 min to 95% B and 5% A. This ratio was maintained for 4 min before returning to the initial conditions.

For ESI-MS and MS/MS analysis, the mass spectrometer was operated in the positive ion mode with a potential of 1800 V applied to the spray tip. Fused silica online PicoTip emitters with 10 μm I.D. (New Objective, Woburn, MA, USA) were used. MS/MS data acquired by means of the MassLynx 4.0 software (www.waters.com) using an automated data-dependent switching between the MS and MS/MS mode based upon ion intensity, mass and charge state (data directed analysis (DDA)). The instrument was set to perform a MS survey scan of 1 s in the *m*/*z* range of 400–1600. Charge state recognition was used to select only doubly, triply and quadruply charged precursor ions for the MS/MS experiments, which also includes the automated selection of the collision energy based on charge and mass. A maximum of three precursor masses were defined for concurrent MS/MS acquisition from a single MS survey scan. The intensity thresholds were set to 30 counts/s (MS to MS/MS) and 2 counts/s (MS/MS to MS), respectively. MS/MS fragmentation spectra were collected from *m*/*z* 50 to 1600.

2.3. Data processing and protein identification

Protein identification was achieved using ProteinLynx-GlobalServer2.1 software (www.waters.com) for processing of the MS/MS data and subsequent databank searching. The MaxEnt3 algorithm was used for deconvoluting the data to the single charge state and deisotoping. After converting to a XML format, the dataset was searched against the SwissProt version 44 databank using a peptide tolerance of 100 ppm and a fragment tolerance of 0.1 Da. Carbamidomethylation of cysteine, methionine oxidation and phosphorylations (STY) were set as possible modifications. Trypsin was set as the digest reagent. One missed cleavage was accepted.

3. Results and discussion

3.1. Characterization of E. coli UNC

E. coli UNC did not display any pathogenic traits in the experiments of Rath et al. [6,7]. However, since the intent of the project was the investigation of the effects of gut epithelial cells upon apathogenic bacteria, the strain was further characterized by multiplex polymerase chain reaction (PCR). In five reaction pools, a total of 30 different loci coding for potentially pathogenic traits were tested for their presence on the genome of *E. coli* UNC (data not shown). Only three determinants were detected: *fimH*, which codes for pili of the type 1; *fyuA*, which codes for an outer membrane receptor of a *Yersinia*-type iron transport system; and finally the capsule encoding locus *kpsMT* II. This lack of virulence genes in *E. coli* UNC is strong evidence for the apathogenicity of this strain and corroborates the observations of Rath et al. We therefore assume that the signaling between *E. coli* UNC and IEC-6 cells should reflect the interactions of a commensal with host cells.

3.2. Protein identification using nanoLC–ESI-MS/MS analysis

The results of the protein analyses are summarized in Tables 1 and 2.

Using online nanoLC–ESI-MS/MS analysis for peptide separation and peptide sequencing in combination with a databank searching approach, it was possible to characterize all selected "2D spots of interest" representing proteins that changed their expression level upon incubation with gut epithelial cells.

Forty-three bacterial proteins were clearly identified with a maximal possible score of 11.96 for the used database with

Table 1
List of up-regulated proteins, i.e. proteins which appeared to be more expressed in the presence of IEC-6 cells

Accession number	Identified protein	MW	p*I* (cal.)	Coverage (%)	Peptides matched
P02408	50S Ribosomal protein L10	17725	9.3	44	7
P08837	PTS system, glucose-specific IIA component (PTGA)	18109	4.7	46	9
P33633	Protein yfiD (YFID)	14275	5.1	50	5
P02392	50S Ribosomal protein L7/L12 (RL7)	12156	4.6	29	3
P14178	Pyruvate kinase 1 (KPY1)	50697	6.0	35	21
P23843	Periplasmic oligopeptide protein precursor (OPPA)	60861	6.4	7	5
P00391	Dihydrolipolyl dehydrogenase (DLDH)	50525	6.1	29	15
P00822	ATP synthase alpha chain (ATPA)	55187	6.0	4	2
P00574	RNA polymerase alpha subunit (RPOA)	36489	5.0	27	10
P02990	Elongation factor Tu (EF-Tu) (P43)	43155	5.4	5	2
P08324	Enolase (ENO)	45495	5.4	6	2
P11665	Phosphoglycerate kinase (PGK)	40961	5.1	47	18
P11604	Fructose-bisphosphate aldolase class II (ALF)	38991	5.8	22	6
P00479	Aspartate transcarbamylase (PYRB)	34274	6.5	35	10
P62707	Phosphoglyceromutase (GPMA)	28407	6.1	33	7
P12758	Uridine phosphorylase (UDP)	27010	6.1	12	3
P21155	SAICAR synthetase (PUR7)	26977	5.1	25	6
P05838	Stringent starvation protein A (SSPA)	24158	5.3	15	3
P30859	Arginine-binding periplasmic protein 1 precursor (ART1)	26912	5.9	6	2
P26427	Sulfate starvation-induced protein 8 (SSI8) (AHPC)	20617	5.1	40	7
P17288	Inorganic pyrophosphatase (IPYR)	19588	5.1	33	6
P39174	Sulfate starvation-induced protein 7 (SSI7)	29021	6.5	19	5
P02364	30S Ribosomal protein S10 (RS10)	11728	10.0	37	4
P02419	50S Ribosomal protein L18 (RL18)	12761	10.7	15	2
P02378	30S Ribosomal protein S20 (RS20)	9475	11.4	22	2

p*I* (cal) = calculated p*I* according to protein sequence in the databank.

158 011 entries (100% probability). The numbers of matched peptides for a unique protein vary from 2 to 21, resulting in sequence coverage of 4–50%, respectively (median 22%). The molecular masses of these proteins cover a wide range from 9.5 up to 95.5 kDa. Moreover, the two-dimensional gel characteristics are clearly related to the (theoretical) characteristics of the identified proteins, such as molecular weight and isoelectric point (p*I*). As an example, the identification of glyceraldehyde-3-phosphate dehydrogenase (GAPDH A) in gel spot 8 (see Fig. 1) is demonstrated in detail (Figs. 1–3 and Table 3). Fig. 2 shows the base peak intensity chromatograms of the survey and MS/MS functions of the tryptically digested proteins of gel spot 8. Three different collision energies were used for fragmentation of the tryptic peptides eluting between 19 and 40 min. The different collision energies were set automatically by the software, dependent on mass and charge state of the precursor ion. Fig. 3A shows the unprocessed MS/MS spectrum with the precursor ion mass of 838.37.

Table 2
List of down-regulated proteins, i.e. proteins which appeared to be less expressed in the presence of IEC-6 cells

Accession number	Identified protein	MW	p*I* (cal.)	Coverage (%)	Peptides matched
P16174	Ribosome recycling factor (RRF)	20626	6.4	20	5
P10344	Glutamine-binding periplasmic protein precursor (GlnBP)	27173	8.4	38	8
P76116	Hypothetical protein yncE precursor (YNCE)	38589	9.5	23	7
P02997	Elongation factor Ts (EFTS)	30273	5.2	40	13
P06977	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH A)	35379	7.1	32	11
P09625	Thioredoxin reductase (TRXB)	34470	5.3	8	2
P02996	Elongation factor G (EFG)	77401	5.3	26	15
P03815	Heat shock protein (ClpB)	95525	5.5	14	11
P39184	Phosphate acetyltransferase (PTA)	76993	5.4	13	8
P09373	Formate acetyltransferase 1 (PFLB)	85172	5.9	22	19
P28302	Glutamate decarboxylase beta (DCEB)	52634	5.4	16	8
P15288	Aminoacyl-histidine dipeptidase (Carnosinase) (PEPD)	52751	5.3	7	3
P80063	Glutamate decarboxylase alpha (DCEA)	52665	5.3	22	11
P19935	TolB protein precursor (TOLB)	45927	7.7	26	9
P15046	Acetate kinase (ACKA)	43263	6.2	21	7
P00907	Carbamoyl-phosphate synthase small chain (CARA)	41389	6.4	6	2
P26606	Protein HdeB precursor 10K L protein (HDEB)	12035	5.9	33	4
P19245	ATP-dependent Clp protease proteolytic subunit (CLPP)	23172	5.7	7	2

p*I* (cal) = calculated p*I* according to protein sequence in the databank.

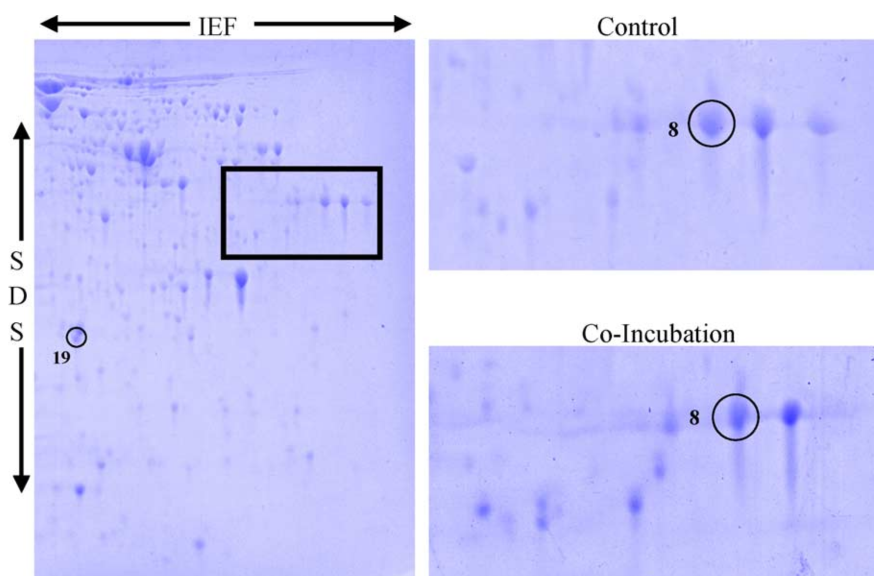

Fig. 1. Example of a 2D-PAGE with a control sample. The upper right part represents the region surrounding spot 8. The lower right is a similar region from a co-incubation sample. The corresponding region is indicated by the frame in the left panel. Also marked is spot 19.

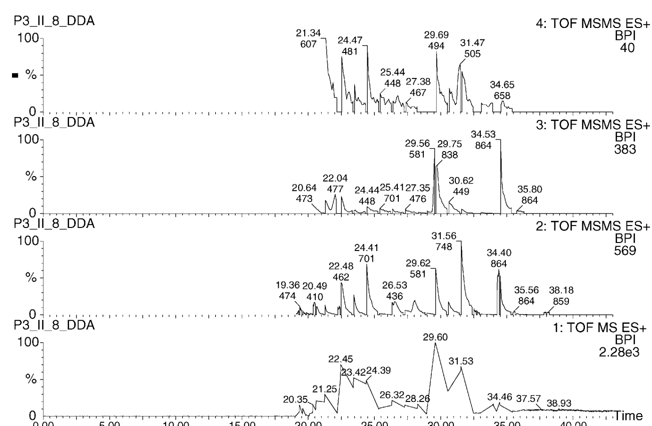

Fig. 2. Base peak intensity chromatograms of the survey and MS/MS functions of tryptic digest mixture of spot 8. Time scale in min. Peak annotation: scan set mass. The different MS/MS acquisitions represent different collision energies dependent on mass and charge state of the precursor ions.

Taking into account that only 1/30 of the available sample was submitted to the nanoLC–ESI-MS/MS analysis, this result illustrates that the applied nanoLC–ESI-MS/MS analysis is very sensitive and well suited for our investigations on host–bacteria interactions in vitro and should also be applicable to other in vitro test systems or in vivo studies. The result of the sequence interpretation of the spectrum of Fig. 3A using ProteinLynxGlobalServer2.1 including MaxEnt Lite for deconvoluting and deisotoping of the raw data is shown in Fig. 3B. The determined sequence was matched to the sequence section 307–320 of GAPDH A.

The validated databank search result of identification of the top score protein of gel spot 8, GAPDH A is summarized in Table 3. Validated results include only the top scoring peptides and MS/MS spectra that exhibit a y-ion sequence stretch of at least three amino acids. The search was conducted with a precursor ion tolerance of 100 ppm and a fragment tolerance of 0.1 Da. Using these criteria, 11 peptides were matched to the GAPDH A sequence. The RMS mass error for these

Table 3
Database search result for gel spot 8 protein GAPDH A, adapted from ProteinLynxGlobalServer2.1, workflow result (protein view)

Mass submitted	Charge submitted	Experimental mass	Molecular weight	Delta (Da)	Delta (ppm)	Probability (peptide score)	Ladder score	Start	End	Sequence
447.727	2	893.438	893.449	0.011	11.91	90.2	78.4	53	60	FDGTVEVK
476.763	2	951.510	951.514	0.004	3.83	185.8	72.1	61	69	DGHLIVNGK
473.771	2	945.526	945.532	0.006	5.97	214.9	97.7	115	123	KVVMTGPSK
409.712	2	817.408	817.437	0.029	35.04	150.6	75.7	116	123	VVMTGPSK
476.230	2	950.444	950.453	0.009	9.09	87.1	67.6	124	131	DNTPMFVK
701.360	2	1400.704	1400.726	0.022	15.45	216.3	67.1	198	212	GASQNIIPSSTGAAK
748.411	2	1494.806	1494.841	0.035	23.17	210.6	71.2	232	245	VPTPNVSVVDLTVR
462.24	2	922.464	922.476	0.012	12.62	175.2	78.4	249	256	AATYEQIK
581.291	2	1160.566	1160.619	0.053	45.38	232.9	87.3	296	306	AGIALNDNFVK
838.360	2	1674.704	1674.753	0.049	29.04	328.5	86.3	307	320	LVSWYDNETGYSNK
554.831	2	1107.646	1107.665	0.019	16.83	166.6	93.9	321	330	VLDLIAHISK

Protein score: 11.962; probability: 100%; peptides matched: 11; coverage: 32.1%; molecular weight: 35 379; p*I*: 7.07.

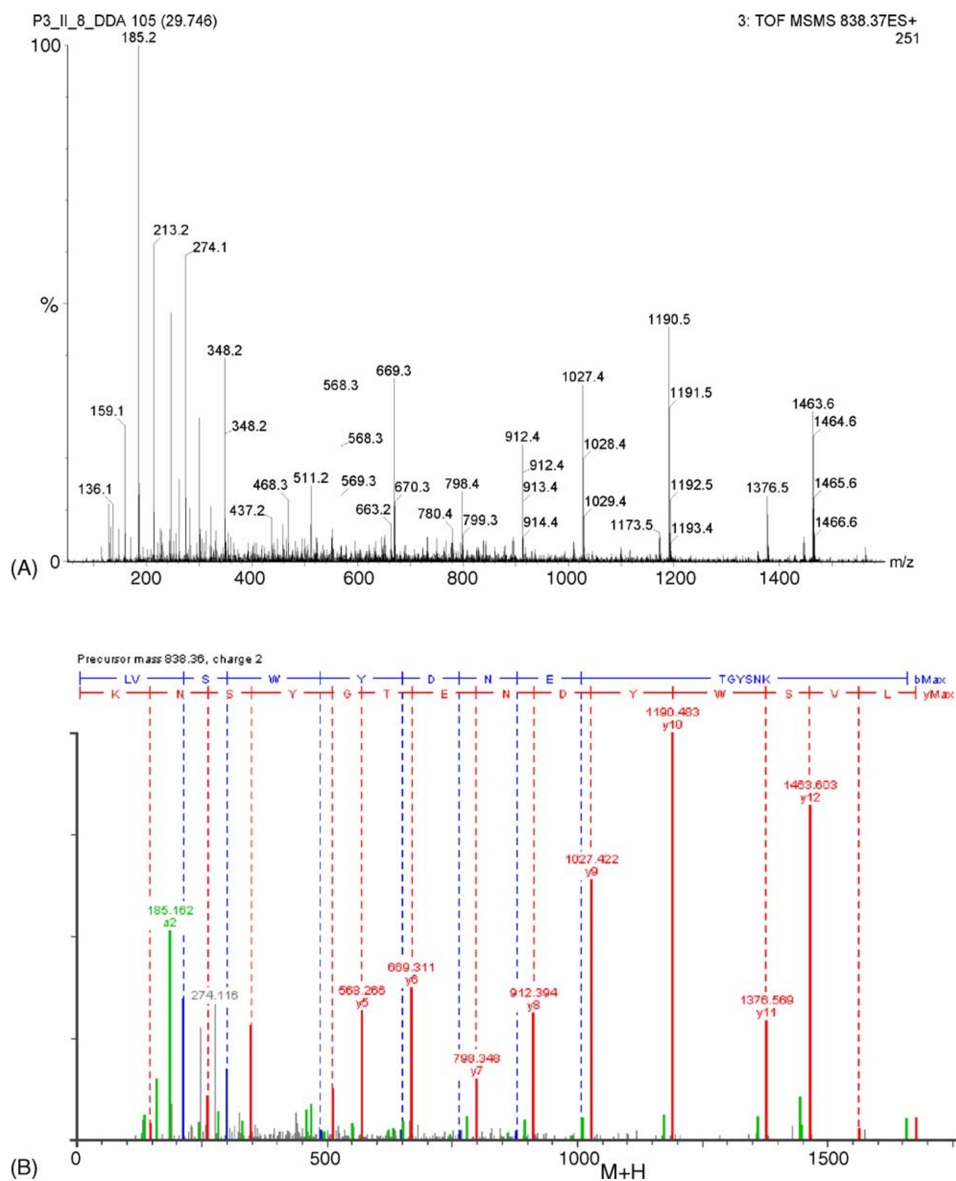

Fig. 3. Part A shows the unprocessed MS/MS spectrum of a tryptic peptide obtained from spot 8. Part B of the figure depicts the sequence interpretation by ProteinLynxGlobalServer2.1 of the processed MS/MS spectrum of the same peptide. The established sequence LVSWYDNETGYSNK was matched to the sequence section 307–320 of GAPDH A (see Table 3).

peptides was 20 ppm. The likelihood of identity of the sequences in the database and the sequence determined by MS/MS analysis is expressed by the ion probability (peptide score) based on the ProbSeq fragmentation model for interpretation of electrospray-tandem mass spectrometry data [13]. The number of the actually found b- and y-fragment ions as percentage of the number of theoretically expected b- and y-fragment ions is described by the ladder score. As shown in Table 3, more than 70% of the possible b- and y-fragments were actually found in the MS/MS spectra of the 11 matched peptides.

LC–ESI-MS/MS using data-dependent acquisition is a very efficient tool often used in proteomic studies, especially when the possible precursor ions are unknown. However, in data-dependent scanning, the most prominent peptide ions in a full-scan spectrum are selected with narrow mass windows for fragmentation to generate MS/MS spectra, whereas lower-abundance peptide ions may not be detected. Otherwise, in data-independent acquisition exists the problem of the impact of wide (m/z 10) isolation windows used for MS/MS on the identification of peptides by database searches [14]. Mass exclusion of already sequenced peptides maybe helpful to reduce such problems in data-dependent scanning.

Using the options of MassLynx 4.0 software of exclusion of peptide masses in data acquisition in combination with a subsequent merging of the results from different runs using ProteinLynxGlobalServer2.1, we were able to determine

SLINTKIKPF KNQAFKNGEF IEITEKDTEG RWSVFFFYPA DFTFVCPTEL GDVADHYEEL

QKLGVDVYAV STDTHFTHKA WHSSSETIAK IKYAMIGDPT GALTRNFDNM REDEGLADRA

TFVVDPQGII QAIEVTAEGI GRDASDLLRK IKAAQYVASH PGEVCPAKWK **EGEATLAPSL**

DLVGKI [1]

SLINTKIKPF KNQA**FKNGEF IEITEK**DTEG RWSVFFFYPA DFTFVCPTEL GDVADHYEEL

QKLGVDVYAV STDTHFTHK**A WHSSSETIAK** IK**YAMIGDPT GALTR**NFDNM R**EDEGLADR**A

TFVVDPQGII QAIEVTAEGI GRDASDLLRK IKAAQYVASH PGEVCPAKWK **EGEATLAPSL**

DLVGKI [2]

SLINTKIKPF KNQA**FKNGEF IEITEK**DTEG RWSVFFFYPA DFTFVCPTEL GDVADHYEEL

QKLGVDVYAV STDTHFTHK**A WHSSSETIAK** IK**YAMIGDPT GALTR**NFDNM R**EDEGLADR**A

TFVVDPQGII QAIEVTAEGI GRDASDLLRK IK**AAQYVASH PGEVCPAK**WK **EGEATLAPSL**

DLVGKI [3]

Fig. 4. Identification of AHPC in spot 19 by merging results of three DDA runs with ProteinLynxGlobalServer2.1. (1) First DDA (without exclude mass list); (2) second DDA using exclude mass list based on A (merge result 1 + 2); (3) third DDA using exclude mass list based on 1 and 2 (merge result 1 + 2 + 3). Data directed analyses were performed using the mass exclude function. Sequence coverage was increased from 8% after the first run (one peptide was matched) to 40% after the final run (seven peptides were matched, YAMIGDPTGALTR was matched with and without methionine oxidation). By means of this technique, the protein AHPC could be clearly identified in spot 19 besides the protein IPYR.

the sequence of additional low-abundant peptides, thereby increasing the sequence coverage of several identified proteins. The detected precursor ion masses were automatically stored as text file by MassLynx 4.0 in the raw data file. After this, it was possible to use this file for creating an exclude mass list in a new acquisition method. By the same scheme, we combined two stored files, in order to increase the number of excluded masses. For example, using this approach, the unambiguous identification of AHPC in gel spot 19 besides IPYR was possible (Fig. 4).

Summarizing our results, we demonstrate that our analytical approach based on a high-capacity sensitive nanoLC in connection with a very sensitive QTOF-tandem mass spectrometer with the ability of accurate mass determination is an excellent analytical tool for our studies of host–bacteria interactions in vitro which should be also applicable to other in vitro test systems or in vivo studies, where less sample material or even more complex mixtures have to be analyzed.

3.3. Biological interpretation

Among the proteins which appeared to be differentially expressed, representatives from different functional classes could be identified, e.g. the protein synthesis related ribosomal proteins S10, S20, L10, L18, L7/L12, the elongation factors Tu and Ts and the ribosome recycling factor (RRF); proteins related to energy metabolism like fructose-bisphosphate aldolase, enolase, phosphoglycerate kinase, phosphoglyceromutase and pyruvate kinase of the glycolysis, phosphate acetyltransferase, acetate kinase, formate acetyltransferase and ATP synthase alpha chain, as well as proteins responding to environmental stimuli such as the heat shock proteins ClpB and ClpP (ATP-dependent Clp protease proteolytic subunit), the acid response proteins YfiD, HdeB and glutamate decarboxylase or the stringent starvation protein A and the sulfate starvation-induced proteins SSI7 and SSI8 (Tables 1 and 2). Interestingly, the glucose-specific component IIA of the PTS system was also identified. This protein is not only involved in the PEP-dependent transport of glucose and mannose, but plays also a major role in catabolite repression and inducer exclusion, regulatory phenomena, which respond to the energetic state of the cell and substrate availability in the environment (for a review, see [15]).

Judging from the distribution of the up-regulated and the down-regulated proteins, it appears that the bacterial cells adjusted the expression of overall protein synthesis functions (reduced expression of EFTS and EFG, enhanced expression

of ribosomal proteins L10, S10, S20, L18) and the pyruvate utilizing part of the glycolytic pathway (phosphate acetyltransferase PTA, formate acetyltransferase PFLB, acetate kinase ACKA), while the expression of some other enzymes of the glycolytic pathway (ALF, ENO, PGK, KPY1) was enhanced.

4. Conclusions

We demonstrate the successful application of the combination of 2D-PAGE with nanoLC followed by ESI-MS/MS for the identification of proteins, which are differentially expressed in a co-incubation model system consisting of gut epithelial cells and bacteria. The sensitivity and accuracy of the developed analytical procedure should permit the extension of our investigations in the future to less well-characterized bacteria other than *E. coli*, for which no genomic data are available and also to more complex systems such as samples from the digestive tract of gnotobiotic mice, which have been mono-associated with the bacteria under investigation. In the long run, even a “meta proteomics” approach to investigate the reactions of several bacteria as found in the Schaedler altered microflora appears feasible [16].

References

[1] G.R. Gibson, Microbiol. Today 29 (2002) 4.
[2] R. Westermeier, T. Naven, Proteomics in Practice, Wiley–VCH, Weinheim, 2002.
[3] C. Dass, Principles Practice of Biological Mass Spectrometry, Wiley–Interscience, 2000.
[4] Y. Ishihama, J. Chromatogr. A 1067 (2005) 73.
[5] E. Gelpi, J. Mass Spectrom. 37 (2002) 241.
[6] H.C. Rath, H.H. Herfarth, J.S. Ikeda, W.B. Grenther, T.E. Hamm Jr., E. Balish, J.D. Taurog, R.E. Hammer, K.H. Wilson, R.B. Sartor, J. Clin. Invest. 98 (1996) 945.
[7] H.C. Rath, K.H. Wilson, R.B. Sartor, Infect. Immun. 67 (1999) 2969.
[8] J.R. Johnson, A.L. Stell, J. Infect. Dis. 181 (2000) 261.
[9] K. Mitsumori, A. Terai, S. Yamamoto, O. Yoshida, FEMS Immunol. Med. Microbiol. 21 (1998) 261.
[10] A.E. Stapleton, M.R. Stroud, S.I. Hakomori, W.E. Stamm, Infect. Immun. 66 (1998) 3856.
[11] M.M. Bradford, Anal. Biochem. 72 (1976) 248.
[12] A. Schrattenholz, Methoden der Proteomforschung, Spektrum Akademischer Verlag, Heidelberg, 2001.
[13] J. Skilling, R. Denny, K. Richardson, P. Young, T. McKenna, I. Campuzano, M. Ritchie, Comp. Funct. Genomics 5 (2004) 61.
[14] D.C. Liebler, Nat. Methods 1 (2004) 16.
[15] J.W. Lengeler, in: C.C. Lin, A.S. Lynch (Eds.), Regulation of Gene Expression in *Escherichia coli*, R.G. Landes, Austin, TX, 1995, p. 231.
[16] F.E. Dewhirst, C.C. Chien, B.J. Paster, R.L. Ericson, R.P. Orcutt, D.B. Schauer, J.G. Fox, Appl. Environ. Microbiol. 65 (1999) 3287.