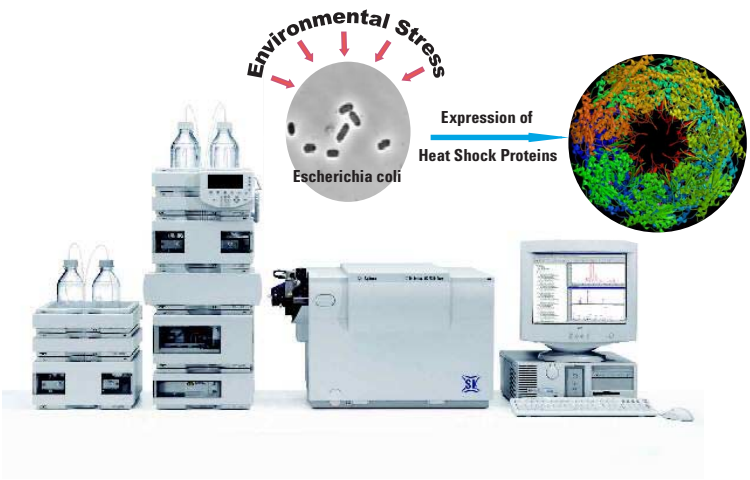
Proteome profiling of *E. coli:* Effect of heat-shock conditions on protein expression pattern

Application

Edgar Nägele
Martin Vollmer
Patric Hörth



Abstract

A variety of stress conditions induce the synthesis of heat-shock proteins (HSPs) in *E.coli.* Using the Agilent Nanoflow Proteomics Solution, comprehensive peptide maps had been constructed for *E. coli*, grown under heat-shock and non-heat-shock conditions. According to a standard method, digested bacterial samples were separated in a two-dimensional manner by combining strong cation exchange (SCX) and reversed-phase chromatography (RP). Peptides were eluted online into the Agilent 1100 Series LC/MSD ion trap mass spectrometer and further analyzed by automatic data dependent MS/MS fragmentation. Bacterial proteins originating from the samples were identified by data base search using the NCBI Database. This Application Note presents the results of the analysis,which led to the identification of several hundred different proteins expressed under heat-shock and non-heat-shock conditions. Several HSPs were identified which occurred predominantly in the sample exposed to high temperature.

Agilent Technologies

Introduction

Exposure of bacterial cells to stress conditions, such as high temperature, starvation, chemicals, UV light and phage infection result in the induction of a variety of heat-shock proteins (HSPs). In *E. coli*, increase of temperature from 30 °C to 42 °C is followed by a rapid induction of HSPs followed by an adaptation period where the rate of HSP synthesis decreases to reach a new steady-state level[1,2]. Major HSPs act either as molecular chaperones or as proteases. Chaperones including DnaK (HSP 70), DnaJ, GroEL, and GroES are crucial for cell survival since they play an important role in promoting refolding and preventing protein aggregation under harmful environmental conditions. Proteases such as Lon and Clp assist in the proteolysis of irreversibly damaged polypeptides. The change in protein composition under heat-shock is mediated by the sigma factor of the RNA polymerase. During heat-shock, sigma 32 (an alternative sigma factor) leads to high level expression of HSPs.

In order to construct complex peptide maps under heat-shock and non-heat-shock conditions for *E.coli* and to identify known HSPs in a background of hundreds of *E.coli* proteins, two-dimensional nano LC/MS/MS was done using the fully automated Agilent Nanoflow Proteomics Solution[3]. The advantages for 2D LC/MS over 2D gel electrophoresis /MS are speed, automation, reproducibilty and flexibility. In addition, 2D LC was shown to be superior for hydrophobic (membrane), low abundant, very large and very small proteins, respectively. In recent reports it has been demonstrated that the combination of various orthogonal HPLC techniques coupled to MS analysis is at least equivaltent to 2D gel electrophoresis[4,5].

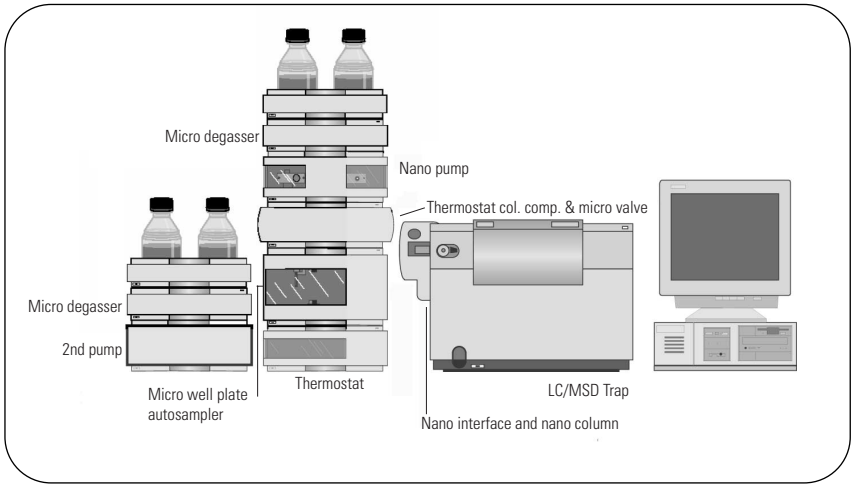


Figure 1
Instrumental setup of the Agilent Nanoflow Proteomics Solution

Instrumentation and methods

Instrumentation

The Agilent Nanoflow Proteomics Solution (figure 1) included the following:

- Agilent 1100 Series nanoflow pump with micro vacuum degasser
- Agilent 1100 Series thermostatted micro well-plate autosampler
- Agilent 1100 Series themostatted column compartment with micro 2-position/6-port valve or the Agilent micro 2-position/6-port switching valve box
- Agilent 1100 Series LC/MSD Trap SL equipped with a nanoelectrospray ion source (Bruker Daltonics)
- Agilent ChemStation A09.03 and Ion Trap software 4.1

Second pump:

- Agilent 1100 Series quaternary pump with micro vacuum degasser

Database search:

- Agilent Spectrum Mill MS Proteomics Workbench

The columns used were:

- Reversed phase (RP): ZORBAX 300 SB C18, 75 µm x 150 mm, particle size 3.5 µm (Agilent part number: 5065-9911).
- Enrichment: ZORBAX 300 SB C18, 0.3 mm x 5 mm, 5 µm particles (Agilent part numbers: 5065-9915 and 5065-9913)
- Strong cation exchange (SCX): PolyLC Inc., Polysulfoethyl A 0.32 mm x 50 mm, 5 µm particles

Cultivation and sample preparation

E.coli K12 was grown at 30 °C in LB-Medium until mid-log phase. Half of the culture was harvested and the remaining culure was incubated at 42 °C for another 30 minutes. Cells were spun down in a Stratos Biofuge (Heraeus Intruments) for 15 minutes at 5000 rpm

at 4 °C. Pellets were resuspended in 50 mM ammonium bicarbonate and lysed with glass beads for 2.5 minutes in a bead beater (Hamilton Beach). Cell debris and beads were removed by centrifugation. The clear supernatant was subjected to protein concentration determination with the Coomassie Plus Protein Assay Kit (Pierce). Protein samples were reduced with 1 mM DTT (45 minutes at 37 °C), alkylated with 10 mM iodoacetamide (1 h, RT) and digested with TPCK trypsin (protein: trypsin 30:1) at 37 °C for 24 h. Quality of the digest was assessed by injecting an aliquot onto reversed phase chromatography. The digest was acidified to pH 3.0 with formic acid, desalted and concentrated by solid phase extraction using an Accubond C8 disposable column (Agilent Technologies). Peptides were eluted in 75 % acetonitrile, 0.1 % formic acid. The eluate was lyophilized to dryness using a SpeedVac Concentrater (Bachofer) and frozen until analysis was performed.

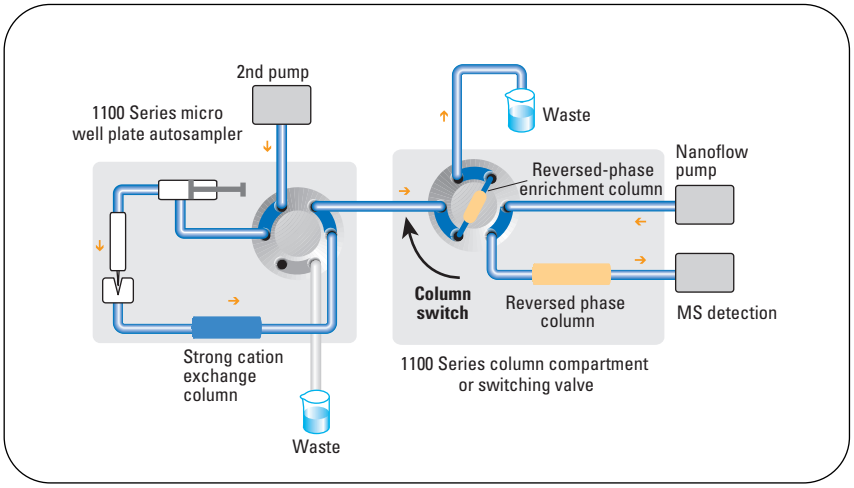


Figure 2
Flow paths, columns and valve connections for 2 D LC

Chromatography

The principle of 2D nano LC is illustrated in figure 2. For the first dimension, 8 µl of redissolved digest in mobile phase A (approximately 50 µg total peptide) was injected onto the SCX column. The column was directly connected to the needle seat of the micro well-plate sampler (figure 2). Mobile phase A was pumped from the second pump (quaternary pump) through the autosampler and SCX column. The column outlet was connected to the 2-position/6-port valve in the autosampler which allows separate cleaning and reconditioning of the SCX column. Flowthrough of non-binding peptides was directed to the 2-position/6-port valve in the column compartment and bound on a short C18 enrichment column, which was mounted in between two ports of this valve. During the first dimension, the enrichment column was in line with the SCX column and the flowthrough of the enrichment column was directed to waste. With this configuration, salts and other non-binding contaminants are removed prior to MS analysis.

In the second flow path, mobile phase was directed from the nanoflow pump through the valve in the column compartment to the RP nanocolumn. The nanocolumn is interfaced directly to the sprayer needle (PicoTip™, New Objective) at the nanoelectrospray source to minimizing post-column peak broadening prior to MS analysis. For the second dimension (RP separation), the valve in the column compartment was switched (figure 2). Here, the enrichment column was transferred into the nanoflow path and flow in the column was reversed which resulted in backflushing. Increasing concentration of organic solvent eluted the concentrated sample from the enrichment column and further separation was achieved on the analytical reversed-phase nano-column. For the next step, the enrichment column was switched back into the flow path of the SCX column. Elution from SCX was obtained stepwise by injecting increasing concentrations of ammonium formate from the autosampler (20 µl portions, 5 column volumes). Salt concentrations of 20, 40, 60, 80, 100, 150, 200, 300, 500, 1000 mM were used to elute peptides from the SCX column. After each salt injection the enrichment column was switched in the nanoflow path and peptides are further separated by RP chromatography directly coupled to the mass spectrometer[6] Table 1 shows the method for 2D nano-LC/MS.

Mass spectrometry and data analysis
The outlet of the RP nanocolumn was connected online to the nanoflow electrospray ion source of the Agilent LC/MSD Trap SL. Eluting peptides were directly analyzed by data dependent MS/MS to obtain mass and sequence information for database analysis. From the obtained MS/MS spectra aquired during all LC runs high quality peptide spectra were filtered with Spectrum Mill software by tag length and signal to noise ratio. Spetra from the same precursor within a time window were merged by the Spectrum Mill software. Subequent MS/MS database search was performed using NCBI database. Individual runs for the heat-shock and non-heat-shock conditions were performed repeatedly and only search results indicated with significant scores were taken into consideration. In addition, mass spectra, ion series of fragmentation patterns and sequence coverage from identified heat-shock proteins were inspected manually and checked for plausibility.

Results and discussion

The completion of the DNA sequence for the *E. coli* genome has revealed that this model organism encodes for about 4200 genes[7]. The protein composition however, is highly variable depending on many extra- and intra-cellular stimuli. Using the fully automated online 2D LC/MS system, up to 500 proteins were identified in a single 2D LC experiment. The analysis was done repeatedly and an average of 100-300 proteins were found in each relevant fraction. The proteins with the top scores are presented for the heat-shock and non-heat-shock subcultures (table 2). The peptide maps constructed for both conditions revealed that cytosolic as well as hydrophobic membrane spanning proteins could be detected by coupling liquid chromatography to nanoelectrospray ion trap MS analysis. In addition, very small, very large, acidic and basic proteins were identified. For these types of proteins, 2D gel electro-phoresis /MS is limited and 2D LC/MS is superior in identifying such proteins in a global proteome analysis[5, 8].

Detecting differences in expression patterns is one of the most challenging fields of interest in proteomic analyses. In order to evaluate the capacity of the Agilent Nanoflow Proteomics Solution, experiments were performed on *E. coli* cells grown under control (30 °C) and heat-shock conditions (42 °C). The analysis was focused on the proteins that are related to heat-shock envirmonmental stress response. In total, 5 different proteins related to heat-shock response were detected only in the heat-shock *E. coli* bacteria culture (table 3). Addtionally one cold shock protein was identified in the low temperature control group (table 1). The identified heat shock proteins GroEL, GroES and DnaK play an important role in regulation of the heat-shock response and are the most prominent among them. The picture on the first page shows a crystal structure model of the

Method for 2D nano-LC/MS

Columns:	
SCX:	PolyLC Inc., PolySulfoethyl Aspartamide, 0.32 mm x 50 mm, 5 µm
Enrichment column:	Zorbax 300SB C18, 0.3 mm x 5 mm, 5 µm
Analytical column:	Zorbax 300SB C18, 75 µm x 50 mm, 3.5 µm
Nanoflow pump:	
Solvent:	A = 0.1 % formic acid in water; B = 0.1 % formic acid in acetonitrile
Gradient:	0 min 5 %B, 10 min 5 %B, 12 min 15 %B, 72 min 55 %B, 74 min 75 %B, 75 min 75 %B, 75.01 min 5 %B Stop time 90 min, Post time 10 min
Flow:	450 nL/min
Micro valve:	
Enrichment column switch:	0 min in-line with SCX; 10 min in-line with nanocolumn; 85 min in-line with SCX
Autosampler:	
Injection volume:	20 µL sample, 20 µL salt solution
Salt steps injected on SCX ($HCOONH_4$):	20, 40, 60, 80, 100, 150, 200, 300, 500, 1000 mM
Quaternary pump:	
Solvent:	0.1 % formic acid and 3% acetonitrile in water
Flow:	0 min 0.1 mL, 0.5 min 0.05 mL, 0.51 min 0.01 mL, 8 min 0.01 mL, 8.01 min 0.005 mL, 9.09 min 0.005 mL, 10.00 min 0.000 mL, 85 min 0.000 mL, 85.01 min 0.005 mL

MS conditions:

Source: positive nanoelectrospray
Drying gas flow: 3 L/min
Drying gas temp: 225 °C
Skim 1: 40 V
Capillary exit: 135 V
Trap drive: 80 V
ICC: on, target 40,000; Max. accu. time: 150 ms
Averages: 4

Automatic MS/MS:
Number of precursors: 2
Isolation width: 1.15 V
Preferred charge state: +2
SmartFrag: 0n, 30-200 %

Table 1
Method for 2D nano-LC/MS.

Protein name	Score	Identified peptides	Matched spectra	Sequence coverage [%]	Isoelectric point	MW (kDa)
A. Heat-shock conditions						
Elongation Factor TF	32,5	3	8	9	5,3	43
Heat-shock Protein Cpn70 DnaK	26,9	2	4	7	4.8	69
50S Ribosomal Protein	22,3	2	14	20	9,5	20
Malate Dehydrogenase	20,2	1	3	11	5,6	29
Ribosomal Protein L29	19,2	2	3	46	10,0	7
Trigger Factor (TF)	19,2	1	5	4	4,8	48
Nucleoside Diphosphate Kinase	18,1	1	1	16	5,5	15
3-Oxoacyl-[acyl-carrier-protein] Synthase II	16,6	2	2	12	5,7	43
Heat-shock Protein Cpn60 GroEL	16,4	2	7	10	4,8	57
Heat-shock Protein Cpn10 GroES	13,9	2	3	27	5,1	10
Outer Membrane Protein A	13,6	1	29	6	5,9	37
Seryl-tRNA Synthetase	13,2	1	1	6	5,3	48
5-Keto 4-deoxyuronate Isomerase	12,9	1	1	7	5,7	31
Ribosomal Protein L24	12,0	1	4	16	10,2	11
Transaldolase B	11,5	2	2	14	5,1	35
Glucose Transferase	9,9	1	14	6	9,3	37
B. Non-heat-shock conditions						
Elongation Factor TF	37,6	4	7	17	5,3	43
50S Ribosomal Protein	26,4	2	13	20	9,5	20
Trigger Factor (TF)	25,1	2	14	9	4,8	48
Ribosomal Protein L19	21,0	2	4	26	10,6	13
Outer Membrane Protein A	14,9	1	33	6	5,9	37
Nucleoside diphosphate Kinase	12,9	1	1	16	5,5	15
Seryl-tRNA Synthetase	11,3	1	1	6	5,3	48
Glycine Hydroxymethyltransferase	10,3	1	1	4	6,0	45
Transcriptioanl Regulator Protein	9,6	2	3	16	6,5	35
Cold Shock Protein cspA	9,5	2	2	31	5,6	7,4
Orotate Phosphoribosyltransferase	8,7	1	1	13	5,3	23
6-Phosphogluconate Dehydrogenase	8,4	1	1	3	5	51

Table 2
Comparison of *E.coli* proteomes of heat-shock. (A) Non-heat-shock (B) condition. Presented are the proteins with the 20 top scores obtained by data analysis with Spectrum Mill software.

Heat Shock Protein	Function	Score	Identified peptides	Matched spectra	Sequence coverage [%]	Isoelectric point	MW (kDa)
Hsp70 DnaK	Chaperonin	26.9	2	4	7	4,8	69
Hsp60 GroEL	Chaperonin	16,4	2	7	10	4,8	57
Hsp10 GroES	Chaperonin	13,9	2	3	27	5,1	10
ClpB	ATP dep. Protease	4,4	1	1	2	5,3	95
htpG	Chaperonin	3,8	1	1	2	5,1	71

Table 3
Heat-shock proteins detected in the heat treated *E. coli* bacteria culture.

asymmetric chaperonin complex GroEL/GroES/(ADP)7 aquired by X-ray diffraction (www.rcsb.org/pdb). It has been reported previously that both DnaK and GroEL might be also detectable at normal temperature and that for both proteins differences are mainly due to cellular concentration and degree of phosphorylation[2]. Further experiments including quantitative tools like ICAT or GIST[9, 10] and selection for phosphorylated peptide species could give deeper insights in this area.

The other proteins, which were related to heat-shock response were also only detected in the heat- shock sample. These included the chaperonin protein htpG and protease ClpB. Selected mass spectra from GroEL, GroES and DnaK peptides are shown in figure 3. In figure 4, the base peak chromatograms (BPC) of selected fractions from the 2D LC of the non-heat shock *E. Coli* reference sample are shown (figure 4A) together with a chromatogram of the 40 mM fraction which shows the BPC and the MS/MS BPC (figure 4B). The scores, number of identified peptides and resulting sequence coverage (tables 2 and 3) together with the mass spectra and manually checked ion series of the fragmentation products formed by MS/MS (figure 3), demonstrate representatively the high quality of the data for all the identified proteins.

The fact that the identified heat-shock proteins were repeatedly found only in the heat- shock sample gave evidence that their expression was significantly impacted by the change in environmental condition.

Conclusion

Two-dimensional nano-LC/MS analysis of the *E. col*i proteome described showed that the Nanoflow Proteomics Solution is capable of analyzing complex "real life" proteome samples containing about 4000 proteins. More than 500 proteins were successfully identified. Furthermore, it can be concluded that 2D nano-LC coupled with electrospray ion trap MS/MS is an excellent tool to identify a subset of peptides in a complex global proteome analysis.

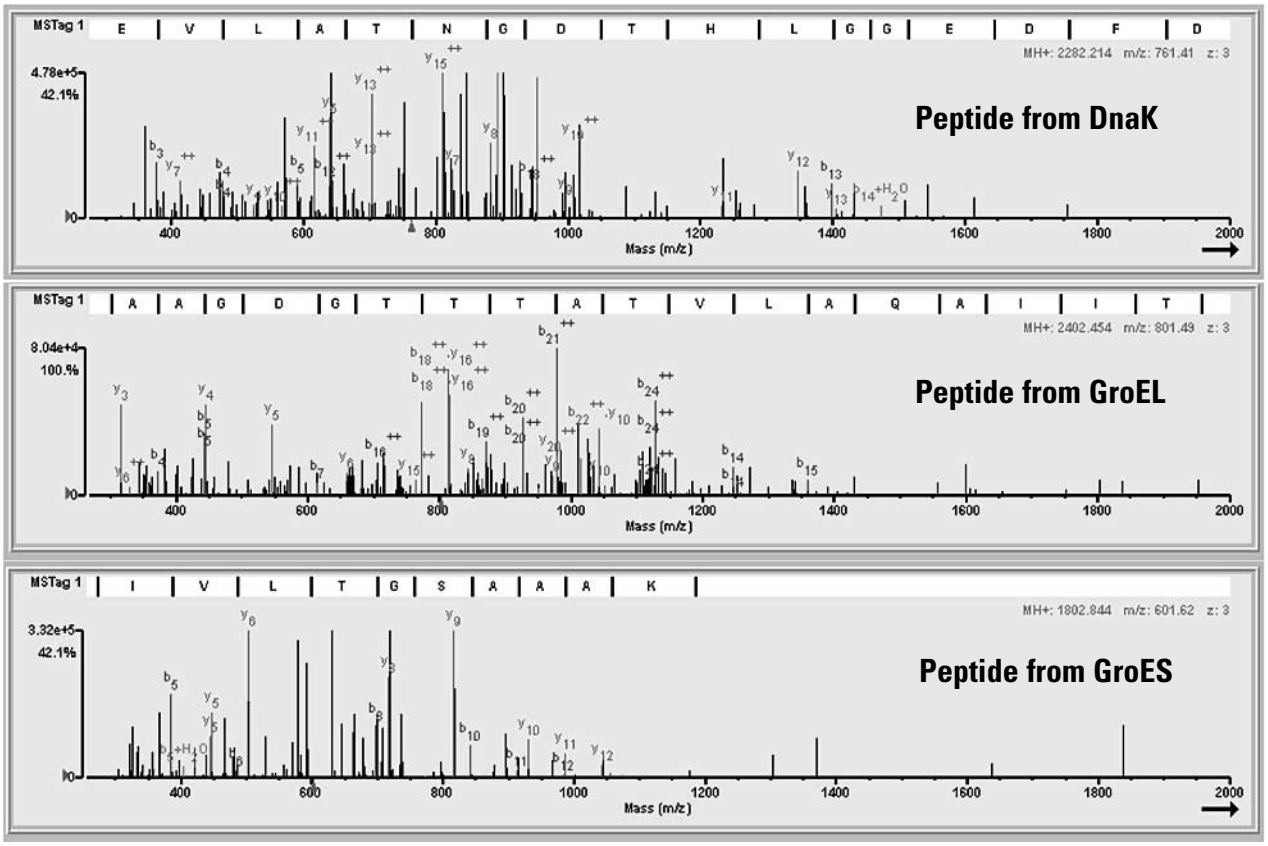


Figure 3
Fragmentation pattern of selected peptides from the heat-shock proteins GroEL, GroES and DnaK identified in the heat-shock *E. coli* sample.

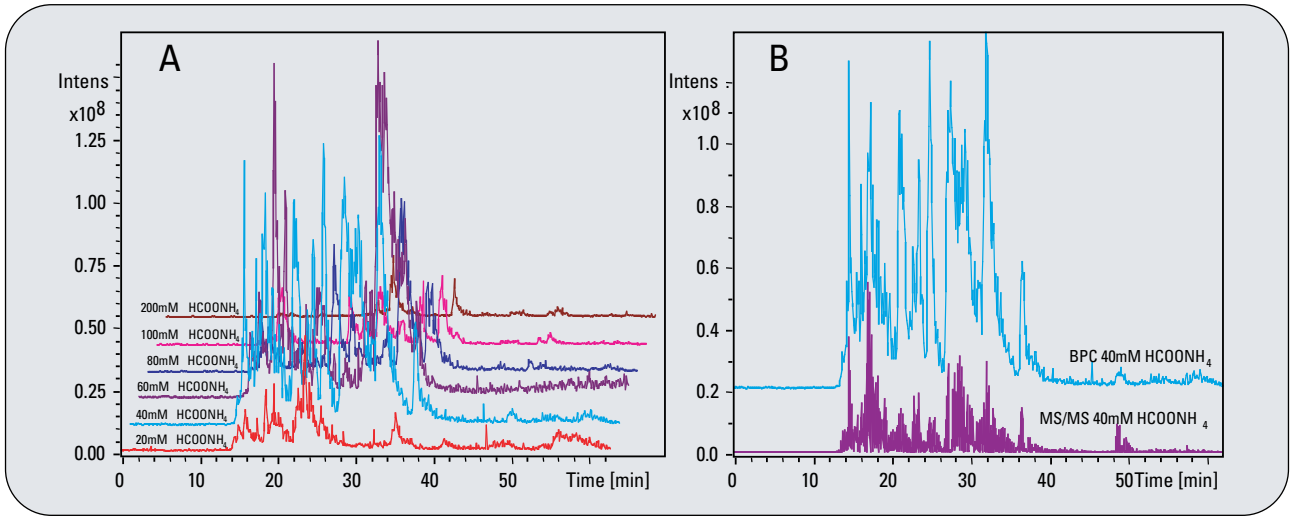


Figure 4
BPC chromatograms of selected fractions from 2D LC of *E. Coli* reference (A) and BPC chromatogram of 40 mM fraction with with MS/MS BPC.

References

1
"Molecular chaperones and folding catalysts regulation, Cellular function and mechanism.", *Bukau B, Harwood Academic Publishers, Amsterdam, p.690,* **1993.**

2.
"The heat shock response of Escherichia coli.", *Arsene F, Tomoyasu T, Bukau B., Int. J. of Food Microbiol. 55: 3-9,* **2000.**

3.
"Agilent Nanoflow Proteomics Solution", *Agilent product brochure* , Publication Number 5988-3621E~, **2002.**

4.
"An automated multidimensional protein identification technology for shotgun proteomics.", *Wolters DA, Washburn MP, Yates JR, Anal. Chem. 73: 5683-5690,* **2001.**

5.
"A proteomic view of the *Plasmodium falciparum* life cycle.", *Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M et al., Nature 419(6906): 520-6,* **2002.**

6.
"Two-dimensional nano LC/MS for complex proteomics applications.", *Vollmer M, Nägele E, Hörth P, Agilent Application Note publication number 5988-7510 EN,* **2002.**

7.
"The complete genome sequence of *Escherichia coli* K-12.", *Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V et al, .Science 277: 1453-1474,* **1997.**

8.
"Large-scale analysis of the yeast proteome by multidimensional protein identification technology.", *Washburn MP, Wolters D, Yates JR, Nature Biotechnol. 19: 242-247,* **2001.**

9.
"Quantitative profiling of differnentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry.", *Han DK, Eng J, Zhou H, Aebersold R, Nature Biotechnol. 19: 946-951,* **2001.**

10.
"Global internal standard technology for comperative proteomics.", *Chakraborty A, Regnier FE, J. Chromatogr A 949: 173-184,* **2002.**

Martin Vollmer is R&D Biochemist, Edgar Nägele is Application Chemist and Patric Hörth is R&D Chemist, all at Agilent Technologies GmbH, Waldbronn, Germany.

www.agilent.com/chem/1100

Printed June 1, 2003
Publication Number 5988-8629EN

Agilent Technologies