Tools and considerations to increase resolution of complex proteome samples by two-dimensional offline LC/MS

Application

Martin Vollmer
Patric Hörth
Edgar Nägele

Abstract

Off-line 2D-HPLC applying a continuous salt gradient in the first separation dimension significantly increases the number of identified proteins from complex samples due to higher chromatographic resolution compared to stepwise elution. In addition, it provides high flexibility for the investigator since micro fractions of the first dimension can be stored, re-analyzed or chemically modified prior to further analysis. To achieve optimal resolution, the applied 2D LC method strongly requires optimization of the two separation dimensions. Especially elution gradients, analysis time, stationary phase material, and column dimension need to be designed and assembled adequately to attain a considerable increase in overall peak capacity. This Application Note shows data which influences the variables and components. It can serve as a guideline for the investigator to adapt his separation method for a specific proteome sample in order to achieve maximum peptide sequence information by MS/MS analysis. This will result in a maximum of identified proteins after a database search.

Agilent Technologies

Introduction

The analysis and identification of single proteins from highly complex proteomic samples with 10^4 to 10^5 different protein species demand analytical techniques with tremendous resolving power. Efficient separation prior to detection by MS and MS/MS is especially crucial to give the MS enough time for successful identification of low abundance proteins, qualitative and quantitative changes in the expression patterns and for detection of posttranslational modifications. Basically two different approaches are under development to address this ambitious goal. These are 2D gel electrophoresis (2D GE)[1] and two- or multidimensional HPLC (2D LC)[2,3]. While 2D GE is still unsurpassed in resolution it still suffers from some important drawbacks that can be overcome by using liquid phase separation. 2D GE suffers from a lack of automation and reproducibility. In addition, hydrophobic membrane proteins, basic and acidic proteins and very large and small proteins are poorly resolved or recovered. Since many of the promising disease markers or drug targets are characterized by these features 2D LC technology provides a good alternative[4].

For 2D LC separations on the peptide level strong cation exchange (SCX) and reversed phase chromatography (RP) were combined[5]. After protein digestion the SCX separation step is usually performed "online" with the second separation step (RP). Peptides from protein digests are eluted from the SCX by sequential salt plugs with increasing ionic strength. After each salt step peptides are concentrated on an enrichment column and transferred to an analytical RP column where they are separated with a gradient of increasing organic solvent concentration[3,6]. Recently, we demonstrated significantly increased peak capacity by using "offline" 2-D LC where the SCX chromatography is performed by applying a linear salt gradient with subsequent fraction collection using the novel Agilent 1100 Series micro fraction collection system[7]. This technique is especially promising for highly complex samples originating from total cell lysates or body fluids while the highly automated online technique provides sufficient resolution for pre-enriched or sub cellular fractions. With the Agilent 1100 micro fraction collection system[8], the Agilent Nanoflow Proteomics Solution[9] and the recently developed Bio SCX Series II stationary phase which is specially designed for peptide separations, we demonstrate that overall peak capacity can be significantly augmented further by optimizing gradient length, analysis time, stationary phase selection and column dimension. These considerations can also serve the investigator when developing the optimal separation method for his specific proteome sample.

Experimental

Equipment

Agilent 1100 Series micro fraction collection system:

- Capillary pump with micro vacuum degasser, thermostatted micro well-plate autosampler, thermo-statted column compartment with micro 2-position/6-port valve, DAD with 500 nL cell, thermostatted micro fraction collector (figure 1 A).

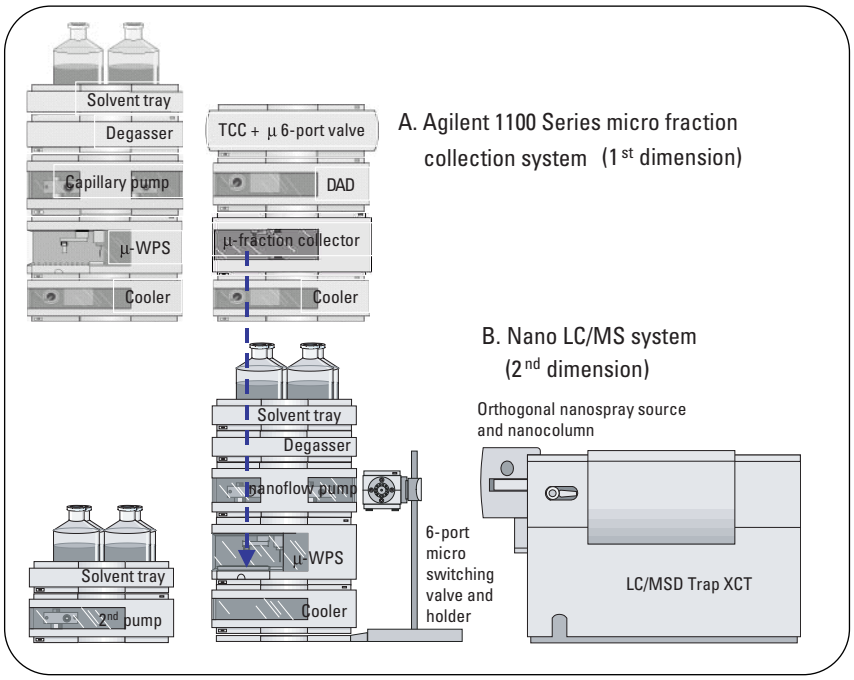

Figure 1
Agilent 1100 Series micro fraction collection system and Agilent Nanoflow Proteomics Solution

Agilent 1100 Nanoflow Proteomics Solution:

- Nanoflow pump with micro vacuum degasser, thermostatted micro well-plate autosampler, micro 2-position/6-port switching valve box with holder, quaternary pump, LC/MSD Trap XCT with orthogonal nanospray source (figure 1B).

Software:

- ChemStation A10.01
- Spectrum Mill MS Proteomics Workbench

The typical workflow for an offline 2D LC proteome analysis is summarized as follows:
In the first LC separation dimension tryptic peptides are loaded onto the SCX column. Peptides are then eluted with a linear salt gradient with subsequent collection of fractions in the micro liter range with the 1100 Series micro fraction collection system (figure 1A). The well-plates containing all collected fractions are then transferred manually to the second dimension – the Agilent Nanoflow Proteomics Solution (figure 1B). After re-injection the peptides from each fraction are concentrated on a short RP enrichment column located between two ports of the micro 6-port valve. This enrichment column is then switched into the nanoflow path, the concentrated peptide fraction is eluted, further separated by a gradient with increasing organic solvent on an analytical nanobore column and sprayed directly into the MS.

Chromatographic conditions: Offline 2D LC

1st dimension (SCX chromatography): Agilent 1100 Series micro fraction collection system

Solvents: A: 5 % AcN +0,03 % formic acid, B: 5 % AcN, 500 mM NaCl + 0.03 % formic acid.
Flow: 5 µL/min.
Gradients: as discussed in the section "Results and discussion"
Column: Agilent BioSCX Series II, 0.30 x 35 mm, 3.5 µm particles (or other dimension as indicated)
Autosampler: 5 µL sample injection volume
Time-based fraction collection mode: 1st fraction 0 - 5 min, 2nd to 13th fraction each 3 min, liquid contact control mode, fraction size 15 µL, cooling 4 °C.

2nd dimension (RP Chromatography): Agilent 1100 Nanoflow Proteomics Solution

Nanoflow pump:
Solvents: A = Water + 0.1 % formic acid, B = ACN + 0.1 % formic acid
Gradient: 0 min 5 %B, 5 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: 300 nL/min
Columns: Zorbax 300SB C18, 75 µm x 150 mm, 3.5 µm particles
Zorbax 300SB C18 or C8, 0.3 x 5 mm, 5 µm particles (or other dimensions as indicated section "Results and discussion)
Micro valve: Enrichment column switch: 5 min
Autosampler: 25 µL sample injection volume (15 µL sample + 10 µl H_2O)

MS conditions: LC /MSD Trap XCT

Ionization mode: Positive nanoelectrospray with Agilent orthogonal source
Drying gas flow: 5 L/min
Drying gas temperature: 300 °C
Vcap: typically 1800-2000 V
Skim 1: 30 V
Capillary exit offset: 75 V
Trap Drive: 85 V
Averages: 1 or 2
ICC: On
Maximum accumulation time: 150 ms
Smart Target: 125,000
MS scan range: 300-2200

Automatic MS/MS:
Peptide scan mode (standard-enhanced for MS and ultra scan for MS/MS)
Number of parents: 3 or 4
Averages: 2
Fragmentation amplitude: 1.3 V
SmartFrag: On, 30-200 %
Active exclusion: On, 2 spectra, 1 min
Prefer +2: on
MS/MS scan range: 100-1800
Ultra scan on
ICC target: 125,000

Sample preparation

Cell disruption was performed of cooled resuspended lyophilized *E.coli* cells (in 50 mM AmBic, 8M urea) in a bead beater (0.1 mm glass beads), followed by reduction with 1 mM DTT, 37 °C, alkylation with 10 mM iodoacetamide 30 min RT, ultrafiltration for buffer exchange and tryptic digest with TPCK trypsin at 37 °C for 16 h. Digested peptides were lyophilized, washed in deionized water followed by a second lyophilization using a SpeedVac. Finally, the sample was dissolved in buffer A prior to analysis.

Results and discussion

SCX gradient comparison

Four gradients differing in length on the first separation dimension (SCX) were compared in order to approach an optimum gradient for the resolution of a high number of peptides for a 2D LC run by still keeping the total analysis time in an acceptable time range. The first series of experiments comprises a digest of 10 bovine proteins, the second series a digested *E. coli* lysate. Both were separated on the Agilent Bio SCX Series II column

(0.3 x 35 mm, 3.5 µm particles) especially designed for peptide and protein separation. Linear salt gradients (0 – 100 mM NaCl) from 15 to 120 min were programmed followed by steep short gradient up to 500 mM NaCl (tables 1 and 2). Micro fractions of 15 µL (3 min) were deposited in cooled 96 conical shaped well-plates by the liquid contact control mode of the Agilent 1100 Series micro fraction collector. Conditions for re-injection onto the second dimension and the second dimension itself were kept constant.

The number of identified proteins clearly demonstrate that for relatively simple mixtures like the artificial 10-protein mix a short

Run	SCX-gradient (min) 1st part (0-100 mM NaCl)	SCX-gradient (min) 2nd part- (100-500 mM NaCl)	Number of SCX fractions	Number of peptides identified	Number of proteins detected
1	15	3	7	**30**	**10**
2	30	6	13	**22**	**8**
3	60	12	25	**22**	**8**
4	120	24	49	**23**	**9**

Table 1
Off-line 2D LC separation of 10 tryptic proteins

Run	SCX-gradient (min) 1st part (0-100 mM NaCl)	SCX-gradient (min) -2nd part- (100-500 mM NaCl)	Number of SCX fractions	Number of peptides identified	Number of proteins detected
1	15	3	7	**486**	**182**
2	30	6	13	**769**	**268**
3	60	12	25	**942**	**312**
4	120	24	49	**1081**	**339**

Table 2
Off-line 2D LC separation of a cytosolic tryptic digest from *E. coli*

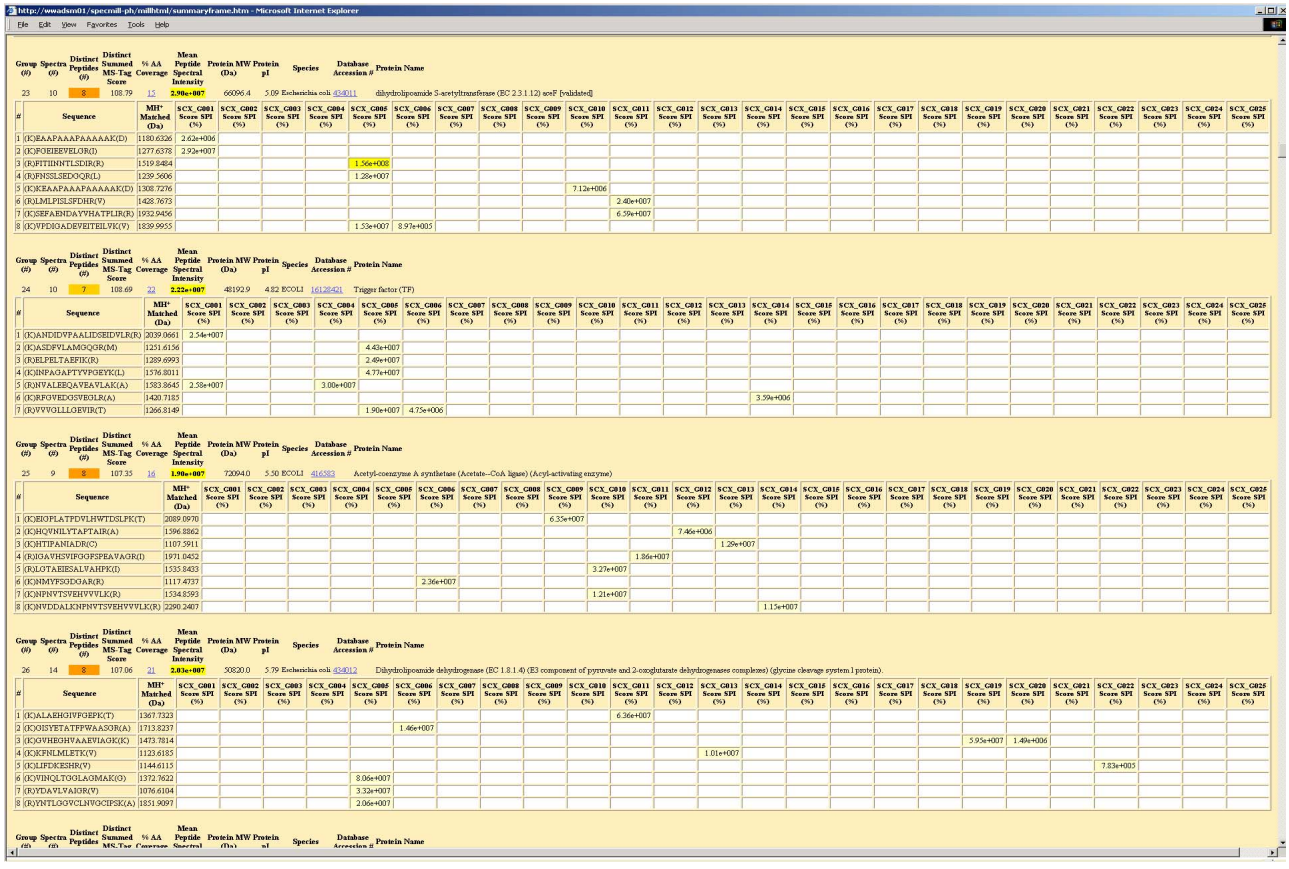

Figure 2
Distribution of single peptides from selected proteins in different SCX fractions

gradient clearly provides sufficient resolution and no further increase is achieved by a more shallow gradient. This certainly also holds true for low complexity samples originating, for example, from a highly resolved 2D gel spot or from isolated protein complexes. For a complex protein expression analysis, however, longer gradients result in a significant increase of identified proteins as shown for the total *E. coli* lysate (tables 1 and 2). However, this increase in numbers of identified peptides and proteins results in much longer total analyses time. The increase in detected peptides for the *E. coli* sample was especially pronounced between the 15 min and the 30 min gradient experiment (37 % more peptides identified). More shallow gradients (60 min or 120 min gradient) increase the number of identified peptides by another 19 % and 13 %. If total analysis is of importance the 30-min NaCl gradient seems to be a good compromise. In this case analysis time from the first dimension and the serial separation of the obtained 13 fractions by subsequent nanoflow LC/MS add up to a total analysis time to 22 h for the off-line 2D LC run. If further resolution is required, for example, for the identification of low abundant proteins, longer gradients (60 min or 120 min) should be conducted alternatively. This however, extends the total analysis time of the off-line 2D LC procedure by another 100 % and 200 %, respectively. An indication for the quality of the overall SCX separation is the fact that even in the shallower gradients single peptides elute in only one or two distinct fractions (figure 2). Comprehensive MS/MS fragmentation pattern, also for the

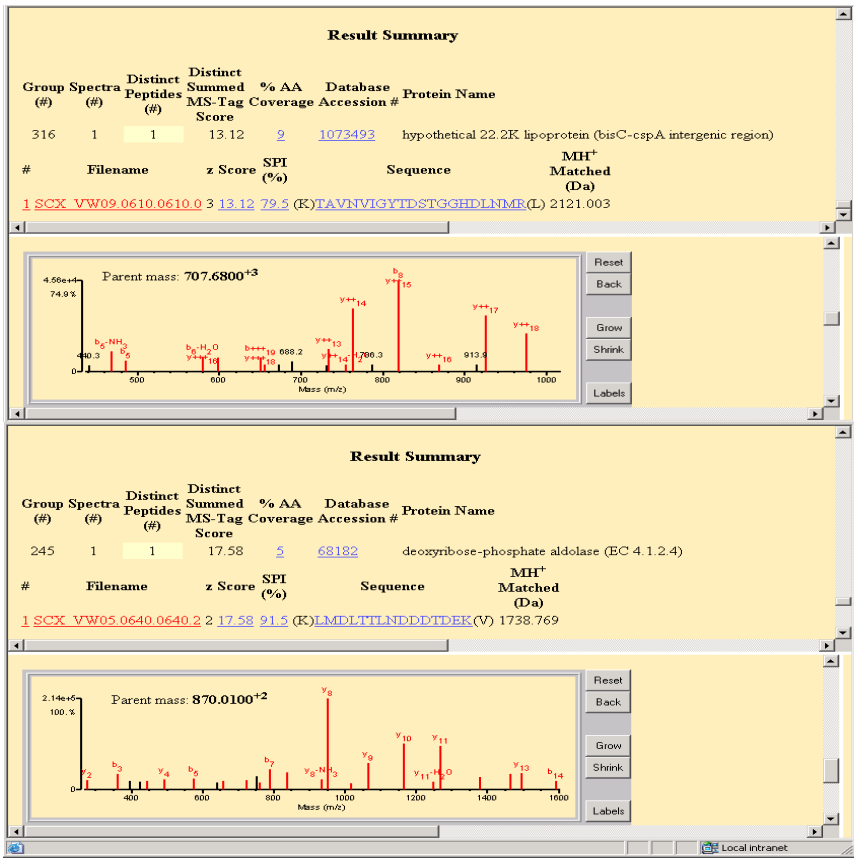

Figure 3
MS/MS ion fragmentation pattern from peptide precursors that originate from proteins just above the level of significance that was set for the experiments.

Column	SCX-dimension (diameter x length in mm)	Flow rate SCX (µl/min)	Fraction size (µl)	Number of peptides identified	Number of proteins identified
1	0.3 x 5	5	15	**167**	**65**
2	0.3 x 35	5	15	**1082**	**422**
3	0.75 x 50	20	60	**490**	**179**
4	0.75 x 150	20	60	**527**	**169**

Table 3
Comparison of different SCX column dimensions for proteomic applications. Data reflect positively identified peptides and proteins from off-line 2D LC runs with a tryptic *E.coli* lysate.

peptides just above the significance threshold, indicate the quality and reliability of the overall analysis (figure 3).

SCX column dimensions

Maximum detection of proteins from a complex sample requires proper combination of column sizes of the 1st and 2nd LC separation dimensions. Nanobore columns (in our case 75 µm inner diameter) in combination with nano electrospray provide by far highest sensitivity and were therefore chosen for the RP separation. In order to avoid overloading from the first dimension

on the one hand and to provide enough material for the detection of lower abundant proteins on the other hand, four SCX column sizes were tested (0.75 x 150 mm, 0.75 x 50 mm, 0.3 x 35 mm, 0.3 x 5 mm). Gradients and flow rates were adjusted roughly according to column dimension in order to achieve good separation. Total analysis time was kept constant for all experiments. *E. coli* lysates were separated in the off-line mode as described earlier. Results demonstrate that using the medium size SCX column (0.3 x 35 mm) in combination with the 75 µm inner diameter nanobore RP column results clearly in the highest number of identified proteins and therefore seems most suitable for the setup of the two dimensional separation of complex proteomic samples (table 3) as described before. The small number of identified peptides with the shorter column (0.3 x 5 mm) is probably due to limited capacity of this column size, the lower amount of identified peptides with the large diameter SCX columns (0.75 mm) could be explained by peak broadening with these columns in comparison to the 0.3 mm i.d. column.

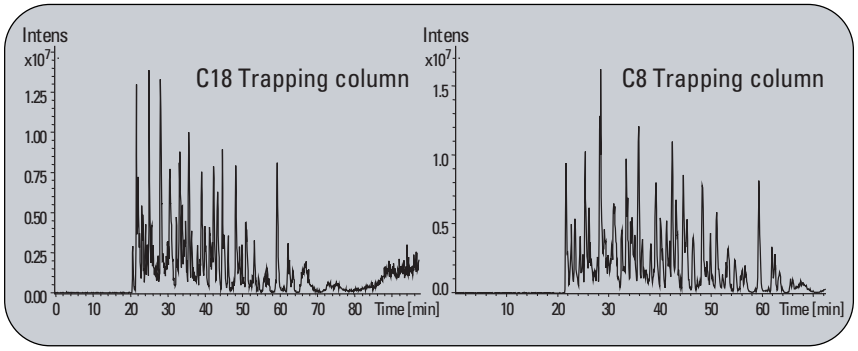

Figure 4
Base peak chromatogramms of nano LC/MS separations performed with C18 and C8 trapping columns

BSA peptides found with C8	Score	BSA peptides found with C18	Score	MH+ Matched (Da)
(K) DDPHACYSTVFDKLK(H)	4.20	(K) DDPHACYSTVFDKLK(H)	5.34	1796.816
(K) HLVDEPQNLIK(Q)	11.88	(K) HLVDEPQNLIK(Q)	15.84	1385.717
KQTALVELLK	4.70	(K) KQTALVELLK(H)	4.67	1142.715
(R) KVPQVSTPTLVEVSR(S)	13.97	(R) KVPQVSTPTLVEVSR(S)	14.76	1639.938
(R) LCVLHEKTPVSEK(V)	10.85	(R) LCVLHEKTPVSEK(V)	8.98	1548.805
(K) LGEYGFQNALIVR(Y)	5.12			1479.796
(K) LKPDPNTLCDEFK(A)	16.43	(K) LKPDPNTLCDEFK(A)	16.25	1577.752
(K) LVNELTEFAK(T)	8.88	(K) LVNELTEFAK(T)	13.58	1163.631
(R) NECFLSHKDDSPDLPK(L)	10.93	(R) NECFLSHKDDSPDLPK(L)	12.12	1902.854
(K) QTALVELLK(H)	12.54	(K) QTALVELLK(H)	7.64	1014.628
(R) RPCFSALTPDETYVPK(A)	12.54	(R) RPCFSALTPDETYVPK(A)	14.59	1881.906
(K) SLHTLFGDELCK(V)	13.26	(K) SLHTLFGDELCK(V)	12.77	1420.678
(K) YLYEIAR(R)	8.19	(K) YLYEIAR(R)	9.16	927.494
	133.49	(K) EACFAVEGPK(L)	12.37	1108.498
		(R) FKDLGEEHFK(G)	11.94	1249.622
		(R) ETYGDMADCCEKQEPER(N)	8.25	2119.789
		(K) EYEATLEECCAK(D)	6.20	1504.582
		(K) CCAADDKEACFAVEGPK(L)	5.77	1930.751
			180.23	

Table 4
BSA peptides identified from C8 and C18 trapping columns

C8 Trapped Peptides No.	Spectra No.	Score	Coverage % AA	C18 Trapped Peptides No.	Spectra No.	Score	Coverage % AA	Protein name
16	29	214,55	37	18	34	232,73	39	Serum albumin
18	25	199,05	35	17	31	222,73	37	Serotransferrin
9	14	118,62	16	9	16	117,69	15	Lactoperoxidase
6	8	66,05	15	7	10	89,09	10	Glutamate dehydrogenase
4	4	40,25	45	2	2	24,04	15	Hemoglobin
3	3	39,28	26	2	2	24,71	15	beta Lactoglobulin
2	2	22,99	13	4	4	33,79	32	Trypsinogen
1	1	7,58	3	1	2	13,28	3	Carbonic anhydrase
1	1	4,75	2	1	2	7,73	2	Carboxypeptidase A
60	**87**	**713**	**av 21.3**	**61**	**103**	**764**	**av. 18.7**	**total**

Table 5
Identified peptides from 10 protein mix using either C8 or C18 trapping columns

Trapping column materials and dimensions

A. Zorbax SB C8 vs. C18 material

Trapping or enrichment columns are used to efficiently concentrate a large volume sample either injected directly or transferred from the first dimension. In addition, enrichment allows sample loading and washing off salt components with high flow rates while sample desorption takes place with nanoflow rates in the analytical nanoflow path. This has the advantage that relatively large columns in the first dimension can be combined with nanobore columns in the second separation dimension. In order to find out if C8 or more hydrophobic C18 RP material is more suited in combination with a C18 analytical nanocolumn for complex peptide separations both materials were tested and compared with a BSA digest (100 fmol) and a 10-protein mix (100 fmol each protein) respectively. Higher peptide scores and higher number of identified peptides for the BSA digest (table 4) and less pronounced for the 10 protein mix (figure 4, table 5) indicate that due to higher hydrophobicity the C18 column is slightly superior to the C8 trapping column.

B. Trapping column dimensions

Since enrichment columns are the interface between first and second dimension it was crucial to investigate the influence of the enrichment column dimension in order to provide a column with enough loading capacity to trap all material from the first dimension (without overloading the following nanocolumn) and to effectively concentrate the sample in a narrow band. Therefore three different column sizes (0.3 x 2 mm, 0.3 x 5 mm, and 0.5 x 5 mm) differing in length and diameter were investigated. *E. coli* sample was injected directly in different amounts (0.1, 1 and 10 µL) onto the C18 trapping columns. With the 2-mm column significantly lower amounts of peptides are identified for 0.1 µl extract, however, all columns show comparable numbers for higher sample load, indicating that our standard trapping dimension (0.3 x 5 mm) fits properly into a 2D LC system with 0.3 x 35 mm on the SCX and 0.075 x 150 mm on the RP dimension, respectively (table 6).

E. coli (µl)	0.3 x 2 mm peptides identified	0.3 x 2 mm proteins identified	0.3 x 5 mm peptides identified	0.3 x 5 mm proteins identified	0.5 x 5 mm peptides identified	0.5 x 5 mm proteins identified
0.1	31	12	55	15	76	27
1	146	51	157	56	152	63
10	166	65	165	69	151	67

Table 6
Identified peptides and proteins from *E.coli* extract using different dimensions of trapping columns

Gradient Slope %B/min	*Distinct MS-Tag score*	*Distinct Peptides*	*Number of Spectra*	*% AA Coverage*
2	141.27	14	40	27
1	265.05	23	139	44
0.5	174.68	15	129	26

Table 7
Identified peptides from BSA from a 10 protein mixture applying RP gradients with different slope

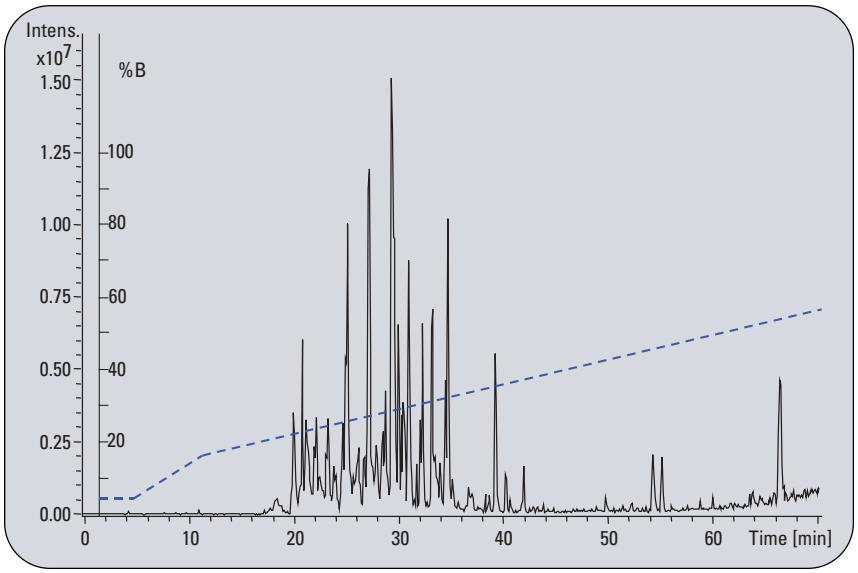

Figure 5
BPC and RP gradient (1 % AcN/min) from a 10 protein mixture

Comparison of RP gradients (the second dimension) on the separation of proteomic samples

In multidimensional chromatography, which uses RP chromatography in the second dimen-

sion directly prior to MS analysis it is most important that this separation step is effectively optimized. RP chromatography has a very high resolving power, therefore optimizing the separation has great impact on the resulting protein identification. In order to optimize the nanobore reversed phase chromatography a mixture of ten digested bovine proteins was used. The digest was separated with gradients delivering slopes ranging between 2 % and 0.50 % acetonitrile per minute in a concentration range between 5 % and 65 % acetonitrile. To find out the relation between gradient steepness and number of identified peptides, acquired spectra, scores and sequence coverage of BSA (as one component in the mixture) were compared for the applied gradients respectively (table 7). Maximum number of identified peptides were clearly obtained with a slope of about 1 % acetonitrile per minute. Steeper gradients might suffer from a lack of resolution while more shallow gradients lead to more flat and broader peaks eluting into the ion trap MS instrument too fast or below the limit of detection, respectively. The base peak chromatogram demonstrates that the majority of the peptides elutes between 20 and 50 minutes between 15 % and 45 % acetonitrile (figure 5). The optimized gradient therefore starts with an initial steeper ramp-up of acetonitrile to 15 % followed by a shallow part between 15 % and 65 % acetonitrile in which the majority of peptides elute.

Conclusion

The combination of strong cation exchange chromatography with nanobore RP chromatography in two-dimensional off-line and on-line HPLC for proteomic samples requires a careful design of gradients and column dimensions in order to achieve maximum resolution. For high complexity samples increase in gradient length on the SCX dimension clearly increases the number of identified proteins. This, however, also increases total analysis time significantly. The selection of 0.3 x 35 mm SCX columns containing BioSCX series II material in combination with 0.3 x 5 mm enrichment columns with Zorbax C18 SB material and a 0.075 mm x 150 mm nanobore column with the same material proves to be an excellent choice to perform off-line and on-line 2D LC in the Agilent system configuration described.

References

1.
Rabilloud T. *Proteomics 2: 3-10.* **2002.**

2.
Washburn M.P. Wolters D. Yates III, *J.R. Nat. Biotechnol., 19 242-247*, **2001.**

3.
a) Nägele E. Vollmer M. Hörth P. *J. Chromatogr. A, 1009: 197-205*, **2003.**
b) "Two-dimensional nano-LC/MS for complex proteomics application" *Agilent Application Note, Publication number 5988-9287EN*, **2003.**

4.
Hanash S., *Nature, 422: 226-232* **2003.**

5.
Wang S. Zhang X., Regnier F.E., *J. Chromatogr. A 2002, 949 153.* **2002.**

6.
a) Vollmer M., Nägele E., Hörth P., *J. Biomol. Techniques 14: 128-135*, **2003.**
b) "Two-dimensional nano LC/MS of *E. Coli* proteome grown on different carbon sources" *Agilent Application Note, Publication number 5988-9862EN*, **2003.**

7.
"Improved protein identification: Off-line multidimensional LC/MS as an effective tool for proteomics research", *Agilent Application Note, Publication number 5988-9913EN*, **2003.**

8.
"Agilent Micro-Fraction Collection System", *Agilent Brochure, Publication Number 5988-9614EN*, **2003.**

9.
"Agilent Nanoflow Proteomics Solution", *Publication Number 5988-3612EN*, **2002.**

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

www.agilent.com/chem

Published October 1, 2003
Publication Number 5989-0213EN

Agilent Technologies