ASMS 2005

IMPROVED PEPTIDE IDENTIFICATION AND PROTEIN COVERAGE FOR PROTEOMICS SAMPLE USING ALTERNATIVE 2D-HPLC (MS/MS) APPROACHES

Petra Olivova, Martin Gilar, Craig A. Dorschel and John C. Gebler
Life Sciences R&D, Waters Corporation

OVERVIEW

- Three off-line 2D-LC setups were compared including SCX-RP and two alternative systems: high/low pH RP-RP, and HILIC-RP.
- Orthogonality of 2D-LC separation was defined
- RP-RP (high/low pH) 2D-LC system provides the greatest number of identified peptides/proteins.
- RP-RP 2D-LC system provides the highest practical peak capacity
- Two methods of tandem mass spectrometry identification of peptides/proteins were compared: DDA, and a novel alternate scanning MS^E

INTRODUCTION

LC-MS methods are used for proteome research for disease markers and/or drug targets discovery. Due to the complexity of proteome samples and the wide protein dynamic range the analysis is difficult. An efficient LC separation and MS/MS analysis is required to enable the detection of lower-abundant proteins of interest in mixtures. One way to address the complexity of these samples is the application of multidimensional separation techniques like 2D-HPLC.

Currently the most common 2D-HPLC approach for separation of peptides is a combination of strong cation-exchange (SCX) and reversed-phase (RP) HPLC.

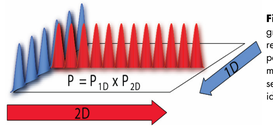


Figure 1: 2D-LC is expected to provide for a greater peak capacity (theoretical number of resolved components) than 1D-LC. A total 2D peak capacity can be calculated as multiplication of peak capacities in first and second dimensions only when the LC modes are ideally orthogonal.

The principles of a geometrical orthogonality are outlined in Figure 5. Each single bin containing a peptide is highlighted; each bin represent an area of an eluting peak. Sum of all bins represent a total area used for separation. In practice, a completely non-orthogonal system (Figure 5A) area coverage is 10%. Ideally orthogonal system (random distribution) coverage is ~64%. By constructing a 10 x 10 or 15 x 15 space, one can compare different data sets numbering 100 or 225 peptides, respectively. Practical peak capacity of 2D separation systems can be calculated from the knowledge of normalized area used for separation, and the peak capacities in both 1D-LC separation modes.

In this work we also measured retention times of 196 tryptic peptides (using 5 protein digests— MassPREP digestion standards, Waters) using 1D-LC-MS in order to construct the normalized peptide retention maps (Figure 6) and compare an orthogonality of 2D-LC systems. The geometrical approach for orthogonality description was developed, employing a surface area of retention plots utilized for 2D separation. The retention maps were normalized and plotted in a normalized 2D separation space (Figure 6). Reversed-phase (RP), hydrophilic interaction chromatography (HILIC), size exclusion chromatography (SEC), and strong cation exchange (SCX) LC modes were evaluated.

We applied two alternative strategies for 2D-HPLC separation. First uses a combination of two RP-RP modes (using different pH), the second employs hydrophilic interaction chromatography (HILIC) and RP-LC. All 2D-HPLC methods were run in off-line mode using a 17 protein digests containing ~1500 peptides.

A novel alternate scanning mode for MS/MS using alternate scans at low and elevated energy has been used and compared to a traditional DDA methods.

METHODS

Sample Preparation

The standard proteins (Sigma) were reduced, alkylated and digested with modified porcine Trypsin (Promega) according to a standard protocol and then mixed in dynamic range 0.03-10 pmol/uL. ~8 Fractions (5 minutes) were collected in each mode (SXC, HILIC, RP high pH), volume was reduced by evaporation to the volume of original sample injected on 1stD LC (no concentration prior to 2ndD LC) and analyzed by LC-MS/MS.

HPLC Columns for fractionation

PolySULFOETHYL Aspartamide SCX 200Å, 5um 1x150mm (PolyLC Inc.)
Atlantis HILIC, 5um 1x150mm (Waters)
RP BEH-C18, 3.5um 1x150mm (Waters)

LC-MS Conditions

Tryptic peptides were analyzed by LC-MS/MS on Q-Tof Ultima (Waters). ProteinLynx Global Server V2.2 software was used for the data analysis.

RP column: NanoEase Atlantis dC18, 3um 0.3x150mm (Waters)
RP flow: 5 ul/min
Solvent A: 0.1% Formic Acid in water
Solvent B: 0.1% Formic Acid in 80% ACN
Temperature: 40°C
Gradient: 0-45 min 0-56% B; 1% ACN/min

A. SCX mode, pH 2.6

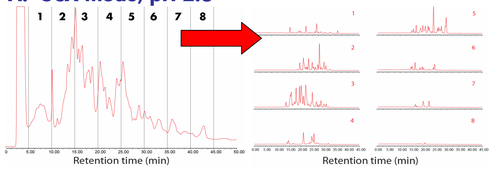


Figure 2: LC-MS analysis of fractions

SCX-LC separation of 17 protein digest (~1500 peptides). Fraction collection intervals is indicated. Conditions: Gradient 0-250mM NaCl in 50 min ; 25% ACN; 10mM Phosphate buffer pH 2.6. Flow rate 50 ul/min. Column temperature 40°C. UV 210nm.

B. HILIC mode, pH 4.5

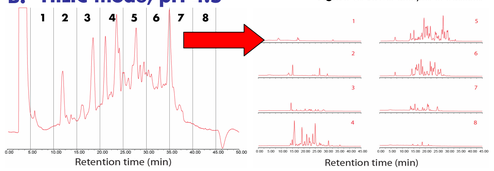


Figure 3: LC-MS analysis of fractions

HILIC separation of 17 protein digest (~1500 peptides). Fraction collection intervals is indicated. Conditions: Gradient 90-40% ACN in 50 min; 10mM Ammonium Formate pH 4.5. Flow rate 50 ul/min. Column temperature 40°C. UV 280nm.

C. RP mode, pH 10.0

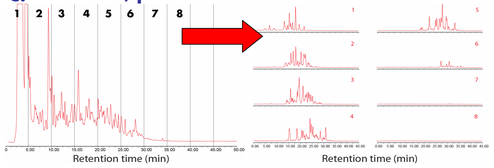


Figure 4: LC-MS analysis of fractions

RP-LC separation of 17 protein digest (~1500 peptides). Fraction collection intervals is indicated. Conditions: Gradient 0-60% ACN in 60 min; 10mM Ammonium Formate pH 10. Other conditions as SCX mode.

RESULTS AND DISCUSSION

Figure 5: Geometrical approach to orthogonality description. 10 x 10 separation space, total 2D peak capacity=100. 100 data points were plotted in 2D space. **(A)** example of completely non-orthogonal 2D separation; 10% area is used. **(B)** idealized orthogonal separation, covering 100% area. **(C)** completely orthogonal system with randomized retention data plot. In average, 64% area is used in ideally orthogonal 2D-seaparation **(B)**. The normalized area of 0.1-0.64 represents 0-100% orthogonality.

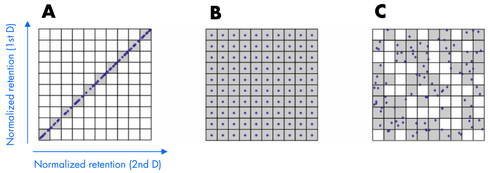
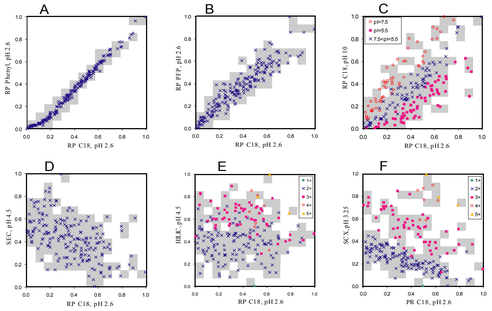


Figure 6: Peptide retention (196 tryptic peptides) was measured in 1D-LC-MS; 2D retention plots were then constructed in a 2D-LC normalized space: **(A)** Phenyl x C_{18}, **(B)** Pentafluorophenyl x RP C_{18}, **(C)** C_{18}, pH 10 x C_{18}, pH 2.6, **(D)** SEC x C_{18}, **(E)** HILIC x C_{18}, **(F)** SCX x C_{18}. The 2D-space was divided into 14 x 14 bins, total peak capacity is 196. Bins containing peptides are highlighted in grey. Greater surface coverage indicates a greater orthogonality (see Table 1).



LC conditions: All columns were 150 x 2.1 mm, except SEC (YMC diol, 60 Å, 750 x 4.6 mm). Flow rate was 0.2 mL/min, column temperature 40 °C, except for SCX (30 °C). Mobile phase A was 0.2% FA, pH 2.6 or appropriate ammonium formate buffer (pH is indicated in the figures); mobile phase B was acetonitrile. For RP systems (**A,C**) the gradient was 0.8% acetonitrile per min for 50 minutes. SEC column was used with a mobile phase containing 20% acetonitrile and 40 mM ammonium formate. HILIC gradient started from 90% acetonitrile, peptides were eluted with 0.8% water gradient per minute. SCX column (PolyLC Polysulfoaspartamide, 200 Å) was run with constant 25% acetonitrile content in mobile phase. Gradient was from 40 mM to 300 mM of ammonium formate, pH 3.25. SCX was compatible with a direct LC-MS detection.

Table 1: Orthogonality of investigated 2D-LC systems. RP-RP high/low pH system has comparable orthogonality to SCX-RP. Greatest orthogonality was observed for HILIC-RP system. Blue font indicate the 2D-LC systems using RP C_{18} as the second dimension.

	Atlantis C18 pH 2.6	Phenyl pH 2.6	XTerra C18 pH 10	HILIC pH 4.5	SEC pH 4.5	SCX, pH 3.25 pH 3.25	PFP pH 3.25
Atlantis C18, pH 2.6	0	13	53	69	58	54	31
Phenyl, pH 2.6	13	0	56	69	54	55	32
XTerra C18, pH 10	53	56	0	59	60	55	49
HILIC, pH 4.5	69	69	59	0	52	54	65
SEC, pH 4.5	58	54	60	52	0	49	56
SCX, pH 3.25	54	55	55	54	49	0	54
PFP, pH 3.25	31	32	49	65	56	54	-1

Protein Expression Analysis vs. DDA - MS/MS acquisition

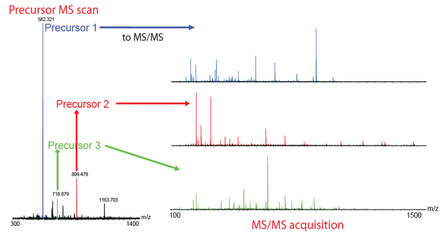


Figure 7: Data Directed Analysis (DDA). DDA employs an MS survey scan to identify abundant components (precursors) and then select these individual precursors for MS/MS fragmentation. The survey/analysis cycle is repeated throughout the run.

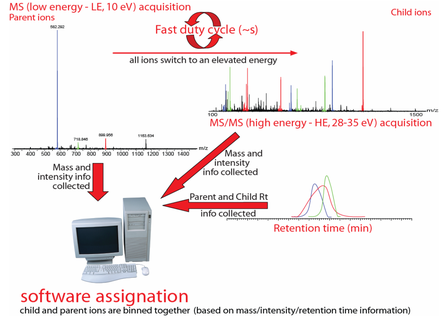


Figure 8: Protein Expression Analysis (MS^E). PE Analysis utilizes a fast cycle a low and elevated potential to collect parent and child ions. The child and parent ions are binned together via a accurate mass/intensity/retention time information. Data are then deconvoluted via software. This mode of data acquisition has an increased duty cycle over traditional DDA methods as well as a typical 5-50 fold increase in sensitivity.

Table 2: Practical peak capacity of selected 2D-LC systems. The RP-RP high/low pH system (Fig5.C) provides the greatest practical peak capacity.

	Atl. C18, pH 2.6	XTerra C18, pH 10	SCX, pH 3.25	HILIC, pH 4.5	SEC, pH 4.5
peak width 4s (min)	0.35	0.35	0.6	0.4	1
useful tg (min)	40	40	30	31	12
1D peak capacity	115	115	51	79	13

	RP(10) x RP(2.6)	SCX x RP	RP(10) x SCX	HILICvsPR	SECvsRP
theor. P_{2D}	13291	5880	5880	9050	1499
area fraction	0.41	0.41	0.42	0.51	0.44
practical peak capacity	5425	2430	2460	4617	658

Protein Expression Results

Proteins	pmol/ul	max. number of peptides *	1D LC-MS Number of peptides	SCX	HILIC	RP high pH
b-Galactosidase (E.coli)	0.03	31	-	-	-	4
b-2 Microglobulin (human)	0.05	7	-	-	-	-
a-2 Macroglobulin (human)	0.1	62	28	13	18	26
b-Lactoglobulin A	0.1	9	5	1	-	1
Ovalbumin	0.3	18	-	-	1	2
Alkoholdehydrogenase (yeast)	0.5	30	16	12	11	16
Enolase (yeast)	0.5	30	18	11	15	15
Hemoglobin (bovine)	0.5	24	8	4	4	10
Phosphorylase b (rabbit)	0.5	82	35	17	15	12
Peroxidase (horseradish)	0.5	11	-	-	-	-
Glyceraldehyd-3-phosphate dehydrogenase (rabbit)	0.5	24	3	3	4	7
Carboxypeptidase A (bovine)	0.75	12	6	-	4	6
Catalase (bovine liver)	1.5	36	27	18	23	25
Ubiquitin (bovine)	2	8	4	3	1	2
Transferrin (bovine)	2	54	38	24	27	38
a-1 Acid glycoprotein (bovine)	7.5	14	11	7	2	8
BSA	10	60	49	33	39	49
Total:		512	248	146	164	221

Table 3: Off-line 2D LC-MS^E results for SCX, HILIC and RP high pH analysis of 17 protein digest (~1500 peptides).
* Maximum number of peptides represent peptides identified in a separate analysis of 2 pmol protein digest injection.

DDA Results

Proteins	pmol/ul	1D LC-MS Number of peptides	SCX	HILIC	RP high pH
b-Galactosidase (E.coli)	0.03	-	1	-	-
b-2 Microglobulin (human)	0.05	-	-	-	-
a-2 Macroglobulin (human)	0.1	5	4	6	9
b-Lactoglobulin A	0.1	1	3	1	3
Ovalbumin	0.3	1	1	1	3
Alkoholdehydrogenase (yeast)	0.5	9	10	6	12
Enolase (yeast)	0.5	9	6	5	8
Hemoglobin (bovine)	0.5	2	4	1	4
Phosphorylase b (rabbit)	0.5	10	6	6	12
Peroxidase (horseradish)	0.5	1	1	1	2
Glyceraldehyd-3-phosphate dehydrogenase (rabbit)	0.5	-	2	1	2
Carboxypeptidase A (bovine)	0.75	-	1	-	2
Catalase (bovine liver)	1.5	15	16	17	18
Ubiquitin (bovine)	2	1	2	1	3
Transferrin (bovine)	2	21	13	13	24
a-1 Acid glycoprotein (bovine)	7.5	7	5	4	8
BSA	10	43	21	24	36
Total:		125	96	87	146

Table 4: Off-line 2D LC-MS/MS (Data Directed Analysis) results for SCX, HILIC and RP high pH modes.

CONCLUSIONS

- A novel methods for 2D-HPLC separation of proteomic samples have been developed.
- A proposed high/low RP-RP 2D-HPLC approach allowed for a significantly increased number of identified peptides compared to HILIC-RP and SCX-RP modes.
- The advantages of RP-RP-HPLC are high peak capacity, high peptide recovery, and the use of salt free mobile phases.
- The performance of high/low pH RP-RP-HPLC compares well to current state-of-the-art SCX-RP-HPLC results and it could be used as an alternative to SCX-RP-HPLC.
- Novel MS/MS^E method doubled the number of identified peptides and provided a significantly better confidence for protein identification in complex sample compared to a traditional DDA mode.
- Novel MS^E method represents a promising tool for proteomic research.

Lit Code Number 720001223EN ©2005 Waters Corporation

TO DOWNLOAD A COPY OF THIS POSTER VISIT WWW.WATERS.COM/POSTERS