Electronic Supplementary Material (ESI) for Analyst.

A simple dual online ultra-high pressure liquid chromatography system (sDO-UHPLC) for high throughput proteome analysis.

Hangyeore Lee,[a] Dong-Gi Mun,[a] Jingi Bae,[a] Hokeun Kim,[a] Se Yeon Oh,[a] Young Soo Park,[b] Jae-Hyuk Lee[c] and Sang-Won Lee[a]*

[a]Department of Chemistry, Research Institute for Natural Sciences, Korea University, Seoul 136-701, South Korea.
[b]Department of Pathology, University of Ulsan College of Medicine, Asan Medical Center, Seoul 138-873, South Korea.
[c]Department of Pathology, Chonnam National University Medical School, Gwangju 501-746, South Korea.

*Corresponding author:

Sang-Won Lee: 1, 5-ka, Anam-dong, Seongbuk-gu, Seoul 136-701, South Korea; Tel: +82-2-3290-3137; Fax: +82-2-3290-3121; Email: sw_lee@korea.ac.kr

Experimental methods

Peptide sample preparation.

Enolase protein was dissolved in 50 mM ammonium bicarbonate and digested for 24 h at 37 °C using trypsin at a trypsin-to-protein ratio of 1:50 (w/w). The peptides were dried and reconstituted in LC solvent A (0.1 % formic acid in water) to concentrations of 10 ng/μL, 1 ng/μL, 100 pg/μL and 10 pg/μL.

The gastric tissues (cancer and adjacent normal) from a patient were pulverized into fine powders by an automated tissue pulverizer (CP02 Cryoprep, Covaris, Woburn, MA). The pulverized tissues were lysed in a lysis buffer containing 4 % sodium dodecyl sulfate (SDS) and 0.1 M Tris-HCl (pH 7.6) by focused-ultrasonicator (S220, Covaris, Woburn, MA). The tissue lysate was reduced in the SDT buffer containing 4 % SDS and 0.1 M dithiothreitol (DTT) followed by a slightly modified filter-aided sample preparation (FASP) digestion method[1]. Briefly, the reduced proteins on a 30 k membrane filter unit (Microcon device, YM-30, Millipore, MA) were washed using 8 M urea solution and then alkylated with 500 mM iodoacetamide (IAA) solution. After the alkylation, the proteins were first washed with 8 M urea and subsequently washed using 50 mM ammonuim bicarbonate. The trypsin was added to the protein sample at a trypsin to protein ratio of 1:50 and incubated at 37 °C overnight. After the first digestion, a second digestion was carried out by adding additional trypsin (1:100 ratio) and incubating at 37 °C for 6 h. Final peptide concentrations from the two tissues were then measured using the BCA assay (BCA Protein Assay Kit, Pierce). For LC-MRM experiments, the peptides from normal tissue were used as matrix.

The resultant peptides were labeled using 4-plex iTRAQ reagent (AB Sciex, Foster City, CA) according to manufacturer instructions. The two adjacent normal peptide samples (100 μg each) were labeled with 114 and 116 iTRAQ reagents, respectively and the two cancer peptide samples (100 μg each) were labeled with

115 and 117 iTRAQ reagents, respectively. The four iTRAQ-labeled peptide samples were pooled and concentrated to 200 μL and the concentrate was immediately subjected to the basic pH reverse-phase fractionation[2].

Peptide fractionation.

The fractionation was performed using an Agilent 1260 HPLC system (Agilent, Palo Alto, CA) equipped with an Xbridge C18 analytical column (4.6 mm × 250 mm, 130 Å, 5 μm, Waters, Milford, MA) and a guard column (4.6 mm × 20 mm, 130 Å, 5 μm). A 115 min gradient was used for the peptide separation at a flow rate of 500 μL/min: at 0% solvent B for 10 min, from 0% to 5% solvent B over 10 min, from 5% to 35% solvent B over 60 min, from 35% to 70% solvent B over 15 min, 70% solvent B for 10 min, from 70% to 0% solvent B over 10 min. Solvent A and B were 10 mM TEAB in water (pH 7.5) and 10 mM TEAB in 90% ACN (pH 7.5), respectively. The eluent was collected into a 96-well plate by a fraction collector (G1364C, Agilent, Palo Alto, CA) every 1 min. The 96 fractions were non-contiguously concatenated into 24 fractions as previously described[3]. The 24 fractions were dried using SpeedVac and stored at -80 °C until LC-MS/MS experiments.

LC-MS/MS experiments.

The iTRAQ-labeled peptides from two pairs of gastric tissues were analyzed by the sDO-UHPLC system coupled to a bench top orbitrap mass spectrometer (Q-Exactive, Thermo Electron, San Jose, CA) by using the same electrospray interface describe above. A 180 min gradient (1%-40% solvent B over 160 min, 40%-80% over 5 min, 80% for 10 min and 1% for 5 min, 300 nL/min) was used for each of the 24 fractions. The spray voltage was maintained at 2.4 kV. The mass spectrometer was operated in data-dependent MS/MS mode with a m/z range of 400-2,000 Th. MS precursor scans were acquired at an automated gain control (AGC) target value of 1.0×10^6 and a resolution of 70,000 with a maximum ion injection time (IT) of 20 ms. Up to 10 most abundant ions in a precursor scan were selected for fragmentation using higher energy collisional dissociation (HCD) at the collision energy of 30 and the isolation width of 1.6 Th. The intensity threshold of initiating fragmentation was set to 1.7×10^4. MS/MS spectra were acquired at AGC target value of 1.0×10^6 and a resolution of 17,500 with a maximum IT of 60 ms.

LC-MS/MS data analysis.

The bipartite graph analysis was performed using in-house software for protein grouping[4, 5]. After bipartite graph analysis, the representative protein of each protein group was selected as follows. The protein with the largest number of matched peptides was considered as representative of the protein group. For the multiple proteins having the same number of peptides, the protein with higher sequence coverage was selected. When

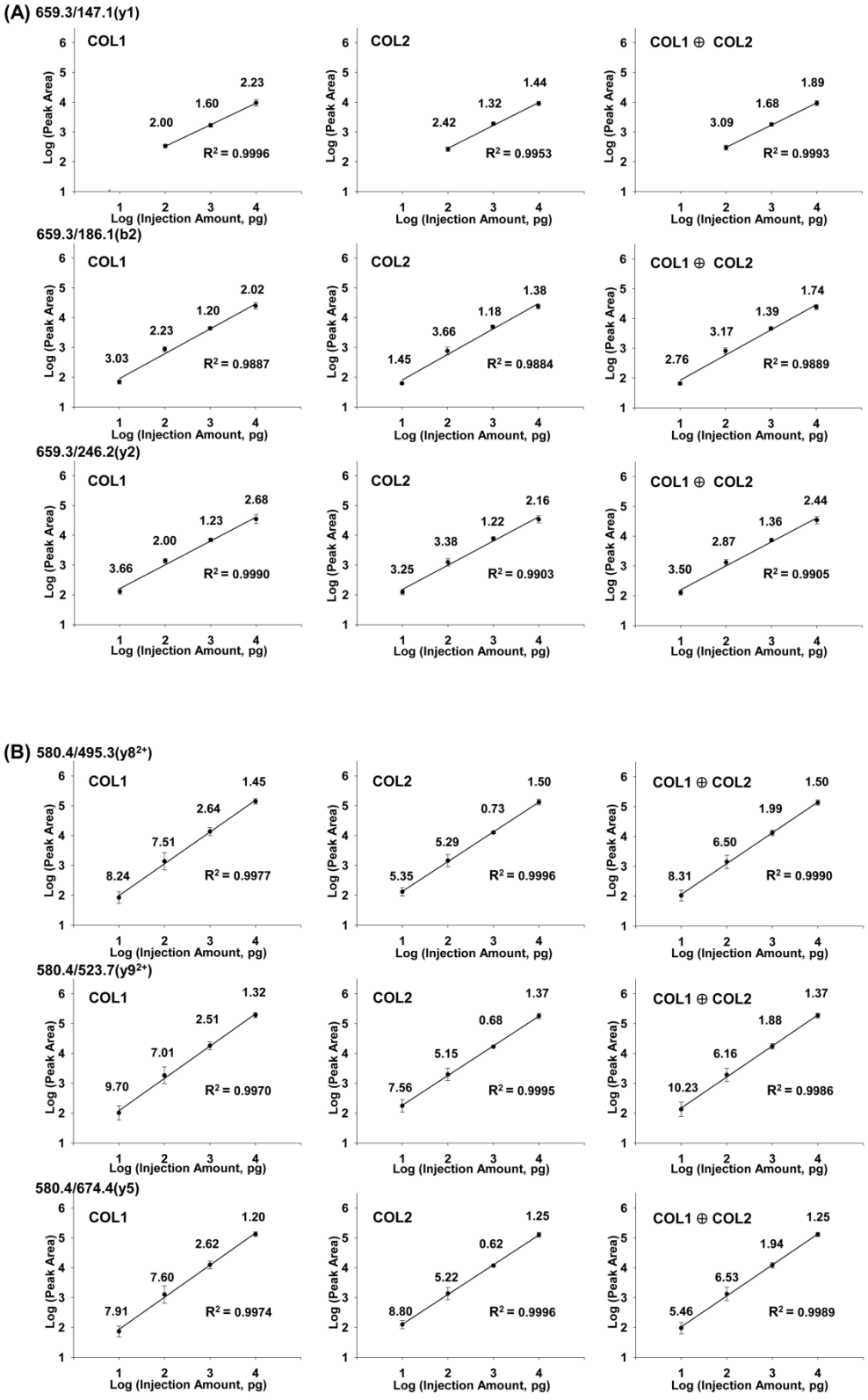
a protein was identified with unique peptide(s), it was also considered the representative protein irrespective of the number of peptides. Two or more peptides were required for protein group identification, while proteins with a single peptide were also included in the final protein groups when its specEvalue of MS-GF+ was less than or equal to $10^{-10\,6}$.

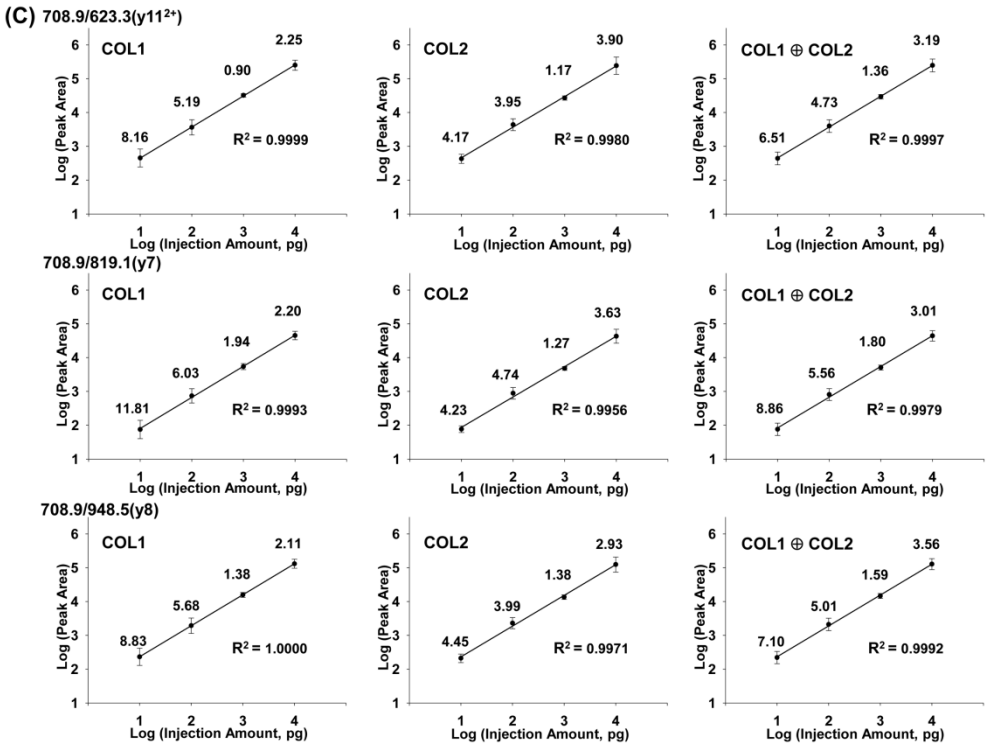
The identified protein groups were mapped to the Ensembl 71 mapping table to identify protein coding genes. For proteins not mapped in Ensembl 71 mapping table, neXtProt (http://nextprot.org, release June 2013) was used for additional mapping. For the gene mapping analysis, the representative proteins of a single peptide were neglected.

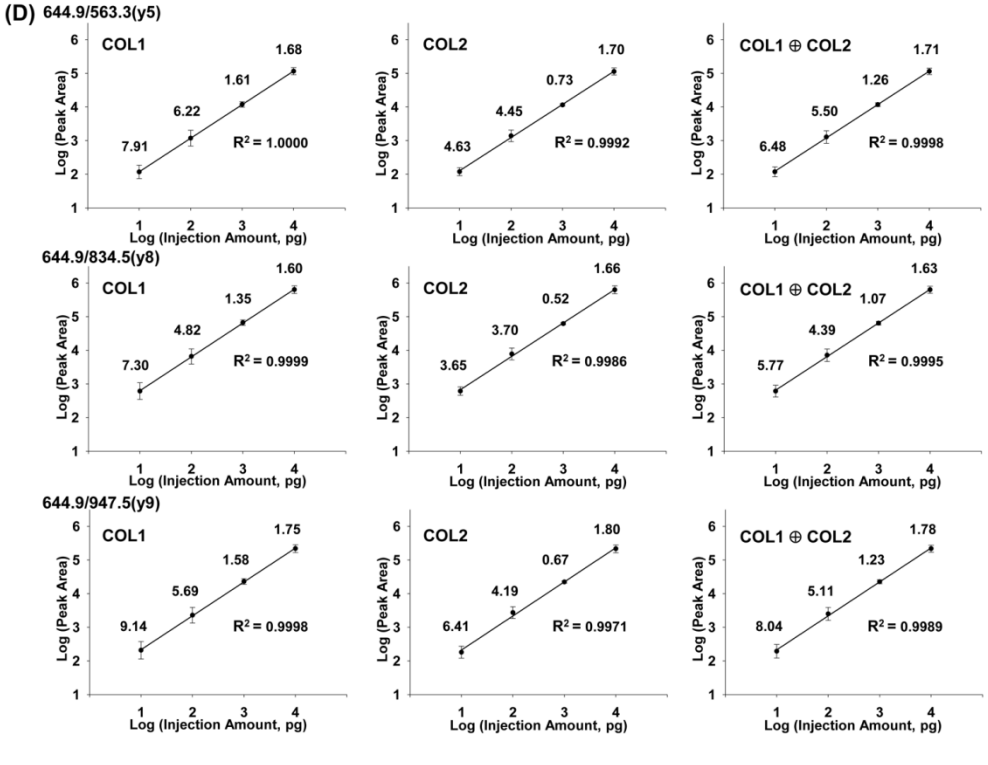
Table S1. Summary of targeted peptides of enolase 1 from yeast (Saccharomyces cerevisiae, P00924) and scheduled MRM parameters

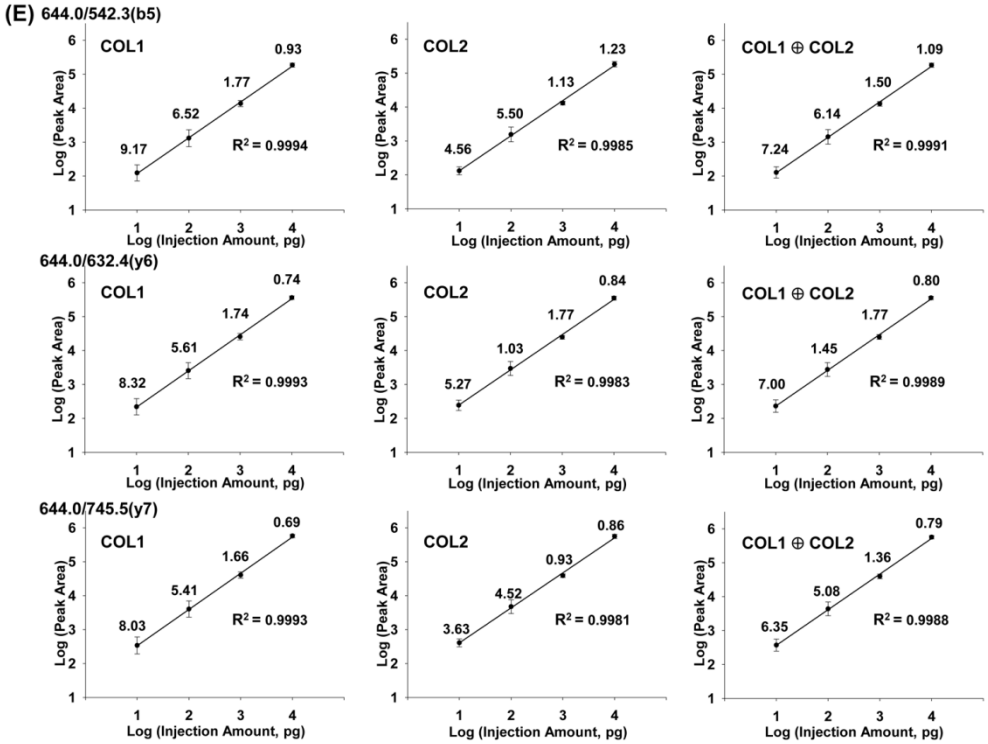
Peptide	MRM transition		RT (min)	Collision energy (V)
	Q1	Q3		
ANIDVK	659.3	147.1 (y1)	49.4	20
	659.3	186.1 (b2)		
	659.3	246.2 (y2)		
IGSEVYHNLK	580.4	495.3 ($y8^{2+}$)	63.5	14
	580.4	523.7 ($y9^{2+}$)		
	580.4	674.4 (y5)		
GNPTVEVELTTEK	708.9	623.3 ($y11^{2+}$)	84.5	25
	708.9	819.1 (y7)		
	708.9	948.5 (y8)		
VNQIGTLSESIK	644.9	563.3 (y5)	87.0	20
	644.9	834.5 (y8)		
	644.9	947.5 (y9)		
NVNDVIAPAFVK	644.0	542.3 (b5)	104.3	17
	644.0	632.4 (y6)		
	644.0	745.5 (y7)		
TAGIQIVADDLTVTNPK	878.6	1002.5 (y9)	111.4	29
	878.6	1073.5 (y10)		
	878.6	1172.6 (y11)		
AVDDFLISLDGTANK	789.9	489.9 (y5)	122.4	25
	789.9	661.0 (b6)		
	789.9	918.5 (y9)		
SGETEDTFIADLVVGLR	911.5	771.5 (y7)	148.2	31
	911.5	842.5 (y8)		
	911.5	955.6 (y9)		

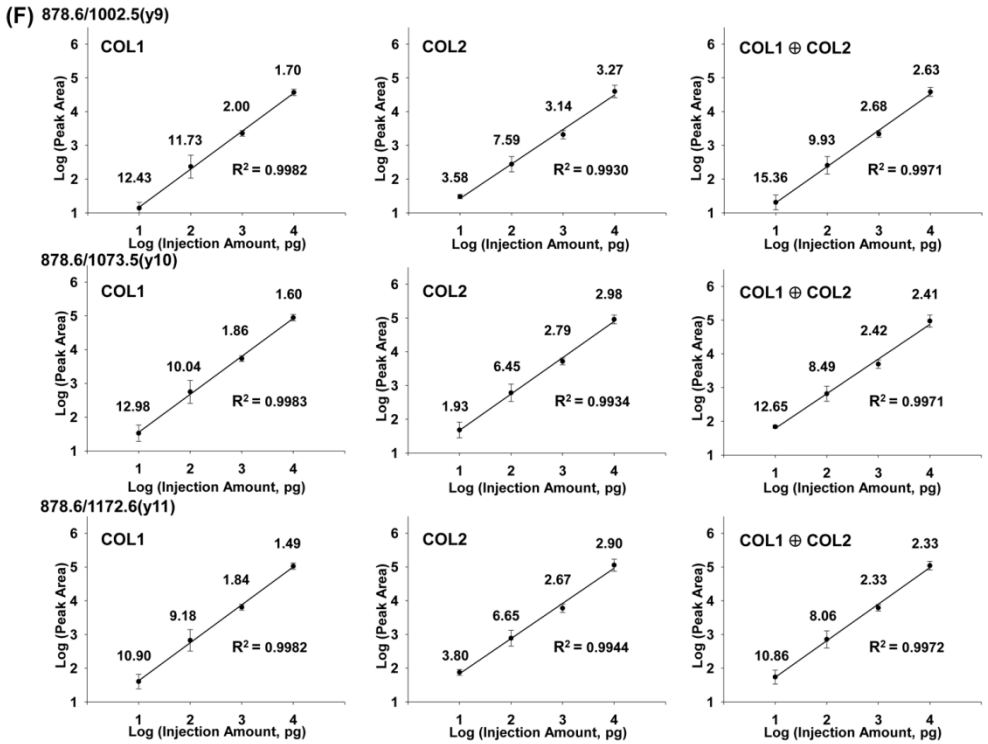
Fig. S1. Serial dilution curves of twenty-four transitions from LC-MRM experiments in matrix free condition using transition data of COL1 only (3 LC-MRM experiments), COL2 only (3 LC-MRM experiments), and using the transition data of both COL1 and COL2 (6 LC-MRM experiments). The numbers on the curves are the coefficient of variation (CV). : (A) ANLDVK, (B) IGSEVYHNLK, (C) GNPTVEVEITTEK, (D) VNQIGTLSESIK, (E) NVNDVIAPAFVK, (F) TAGIQIVADDLTVTNPK, (G) AVDDFLISLDGTANK, (H) SGETEDTFIADL VVGLR

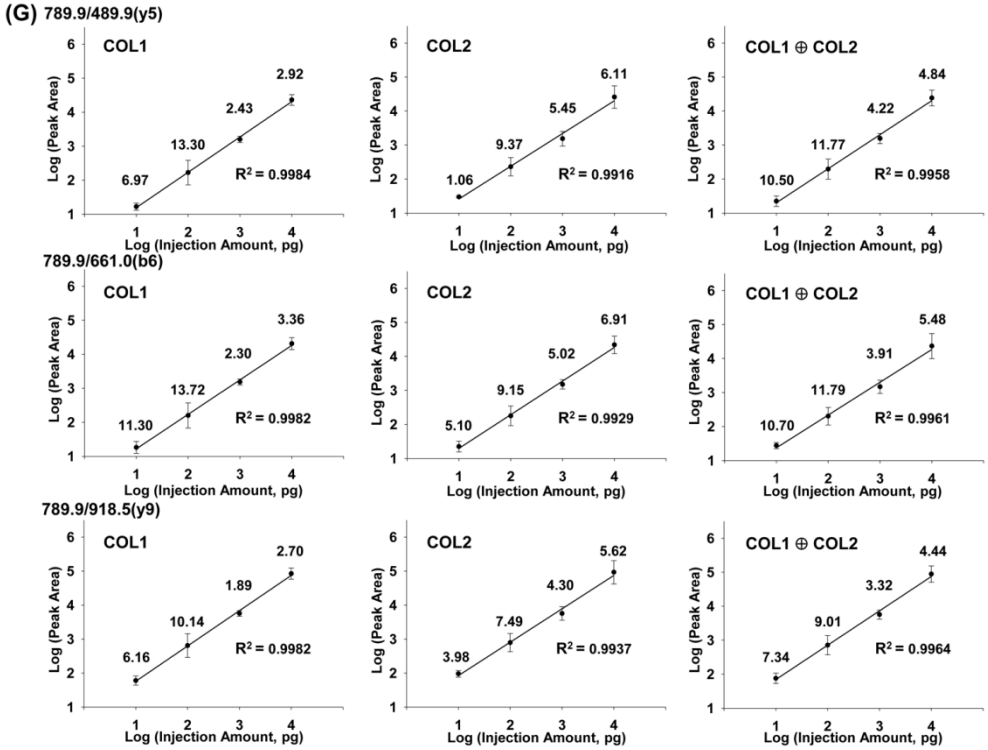



(C) 708.9/623.3(y11^{2+})
COL1
2.25
0.90
5.19
8.16
R^2 = 0.9999
COL2
3.90
1.17
3.95
4.17
R^2 = 0.9980
COL1 ⊕ COL2
3.19
1.36
4.73
6.51
R^2 = 0.9997
708.9/819.1(y7)
COL1
2.20
1.94
6.03
11.81
R^2 = 0.9993
COL2
3.63
1.27
4.74
4.23
R^2 = 0.9956
COL1 ⊕ COL2
3.01
1.80
5.56
8.86
R^2 = 0.9979
708.9/948.5(y8)
COL1
2.11
1.38
5.68
8.83
R^2 = 1.0000
COL2
2.93
1.38
3.99
4.45
R^2 = 0.9971
COL1 ⊕ COL2
3.56
1.59
5.01
7.10
R^2 = 0.9992
Log (Peak Area)
Log (Injection Amount, pg)


(D) 644.9/563.3(y5)
COL1
1.68
1.61
6.22
7.91
R^2 = 1.0000
COL2
1.70
0.73
4.45
4.63
R^2 = 0.9992
COL1 ⊕ COL2
1.71
1.26
5.50
6.48
R^2 = 0.9998
644.9/834.5(y8)
COL1
1.60
1.35
4.82
7.30
R^2 = 0.9999
COL2
1.66
0.52
3.70
3.65
R^2 = 0.9986
COL1 ⊕ COL2
1.63
1.07
4.39
5.77
R^2 = 0.9995
644.9/947.5(y9)
COL1
1.75
1.58
5.69
9.14
R^2 = 0.9998
COL2
1.80
0.67
4.19
6.41
R^2 = 0.9971
COL1 ⊕ COL2
1.78
1.23
5.11
8.04
R^2 = 0.9989
Log (Peak Area)
Log (Injection Amount, pg)


(E) 644.0/542.3(b5)
COL1
0.93
1.77
6.52
9.17
R² = 0.9994
COL2
1.23
1.13
5.50
4.56
R² = 0.9985
COL1 ⊕ COL2
1.09
1.50
6.14
7.24
R² = 0.9991
644.0/632.4(y6)
COL1
0.74
1.74
5.61
8.32
R² = 0.9993
COL2
0.84
1.77
1.03
5.27
R² = 0.9983
COL1 ⊕ COL2
0.80
1.77
1.45
7.00
R² = 0.9989
644.0/745.5(y7)
COL1
0.69
1.66
5.41
8.03
R² = 0.9993
COL2
0.86
0.93
4.52
3.63
R² = 0.9981
COL1 ⊕ COL2
0.79
1.36
5.08
6.35
R² = 0.9988
Log (Peak Area)
Log (Injection Amount, pg)


(F) 878.6/1002.5(y9)
COL1
1.70
2.00
11.73
12.43
R² = 0.9982
COL2
3.27
3.14
7.59
3.58
R² = 0.9930
COL1 ⊕ COL2
2.63
2.68
9.93
15.36
R² = 0.9971
878.6/1073.5(y10)
COL1
1.60
1.86
10.04
12.98
R² = 0.9983
COL2
2.98
2.79
6.45
1.93
R² = 0.9934
COL1 ⊕ COL2
2.41
2.42
8.49
12.65
R² = 0.9971
878.6/1172.6(y11)
COL1
1.49
1.84
9.18
10.90
R² = 0.9982
COL2
2.90
2.67
6.65
3.80
R² = 0.9944
COL1 ⊕ COL2
2.33
2.33
8.06
10.86
R² = 0.9972
Log (Peak Area)
Log (Injection Amount, pg)


(G) 789.9/489.9(y5)
COL1
COL2
COL1 ⊕ COL2
6.97
13.30
2.43
2.92
R² = 0.9984
1.06
9.37
5.45
6.11
R² = 0.9916
10.50
11.77
4.22
4.84
R² = 0.9958
789.9/661.0(b6)
11.30
13.72
2.30
3.36
R² = 0.9982
5.10
9.15
5.02
6.91
R² = 0.9929
10.70
11.79
3.91
5.48
R² = 0.9961
789.9/918.5(y9)
6.16
10.14
1.89
2.70
R² = 0.9982
3.98
7.49
4.30
5.62
R² = 0.9937
7.34
9.01
3.32
4.44
R² = 0.9964
Log (Peak Area)
Log (Injection Amount, pg)

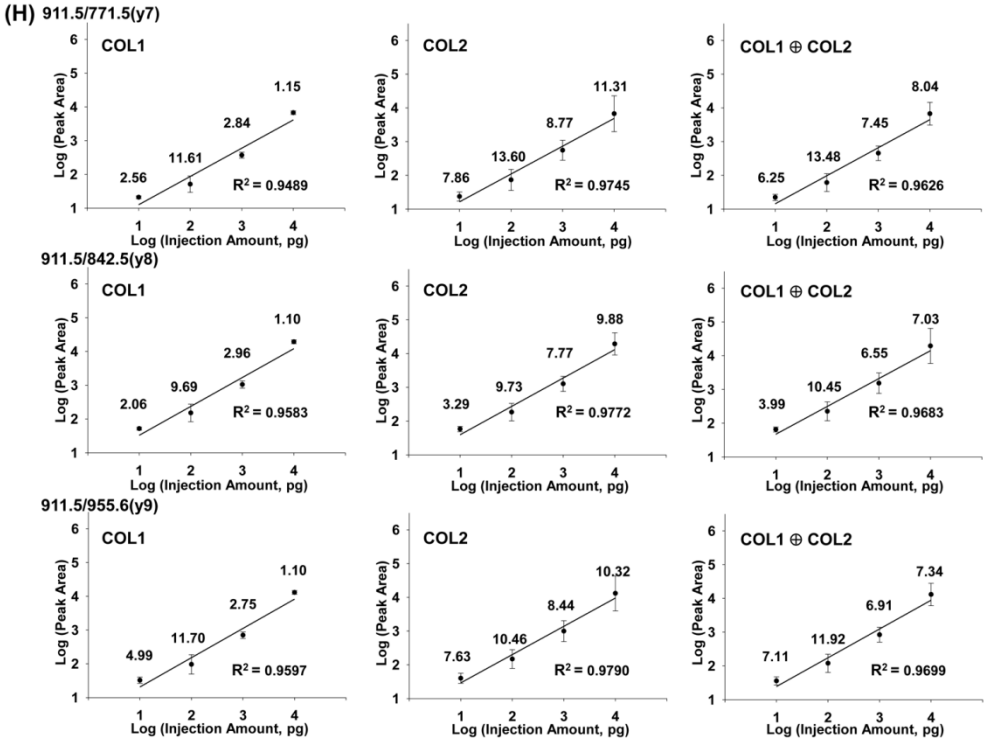
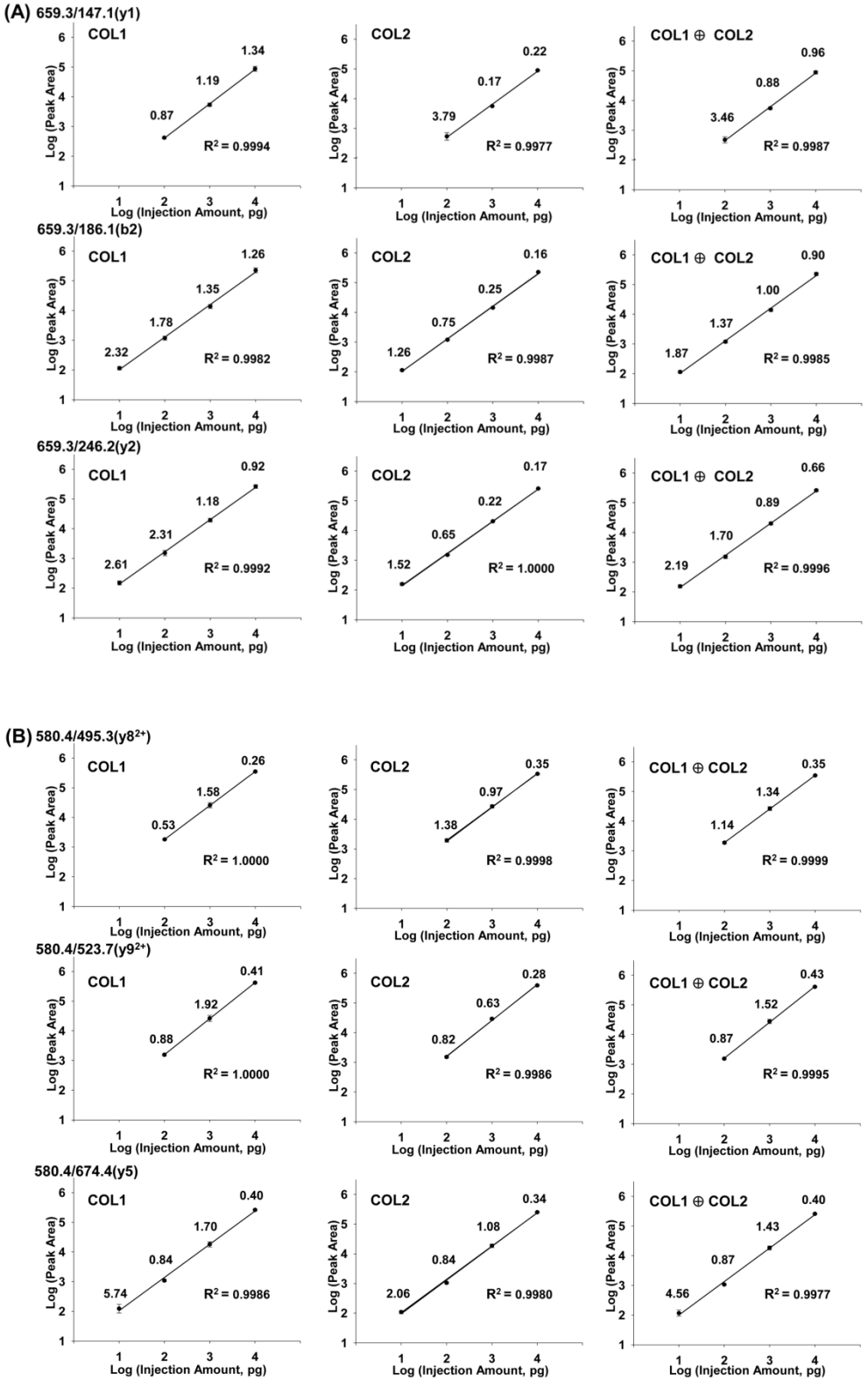
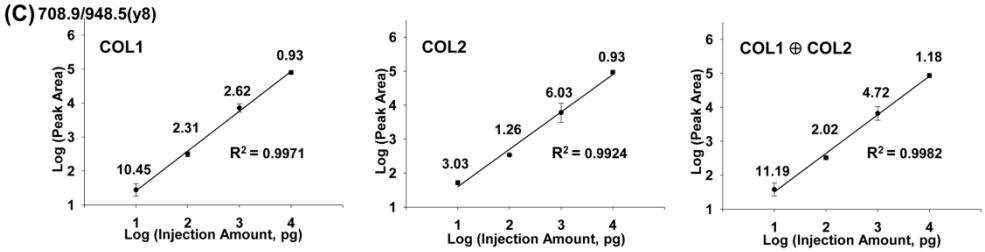
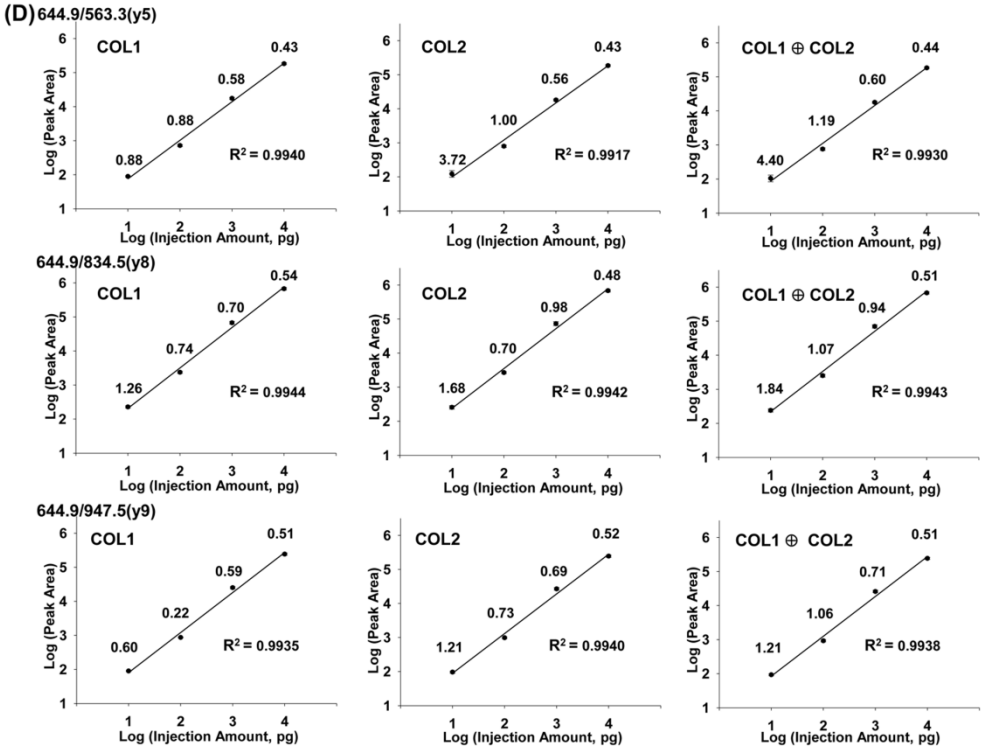
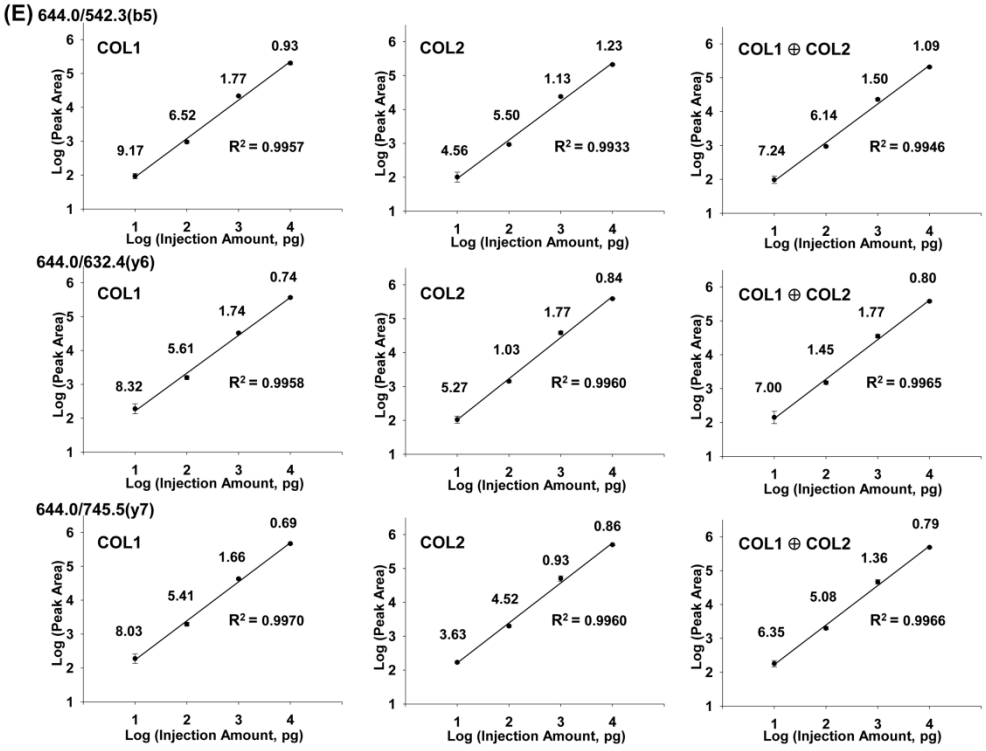

(H) 911.5/771.5(y7)
COL1
COL2
COL1 ⊕ COL2
2.56
11.61
2.84
1.15
R² = 0.9489
7.86
13.60
8.77
11.31
R² = 0.9745
6.25
13.48
7.45
8.04
R² = 0.9626
911.5/842.5(y8)
2.06
9.69
2.96
1.10
R² = 0.9583
3.29
9.73
7.77
9.88
R² = 0.9772
3.99
10.45
6.55
7.03
R² = 0.9683
911.5/955.6(y9)
4.99
11.70
2.75
1.10
R² = 0.9597
7.63
10.46
8.44
10.32
R² = 0.9790
7.11
11.92
6.91
7.34
R² = 0.9699
Log (Peak Area)
Log (Injection Amount, pg)

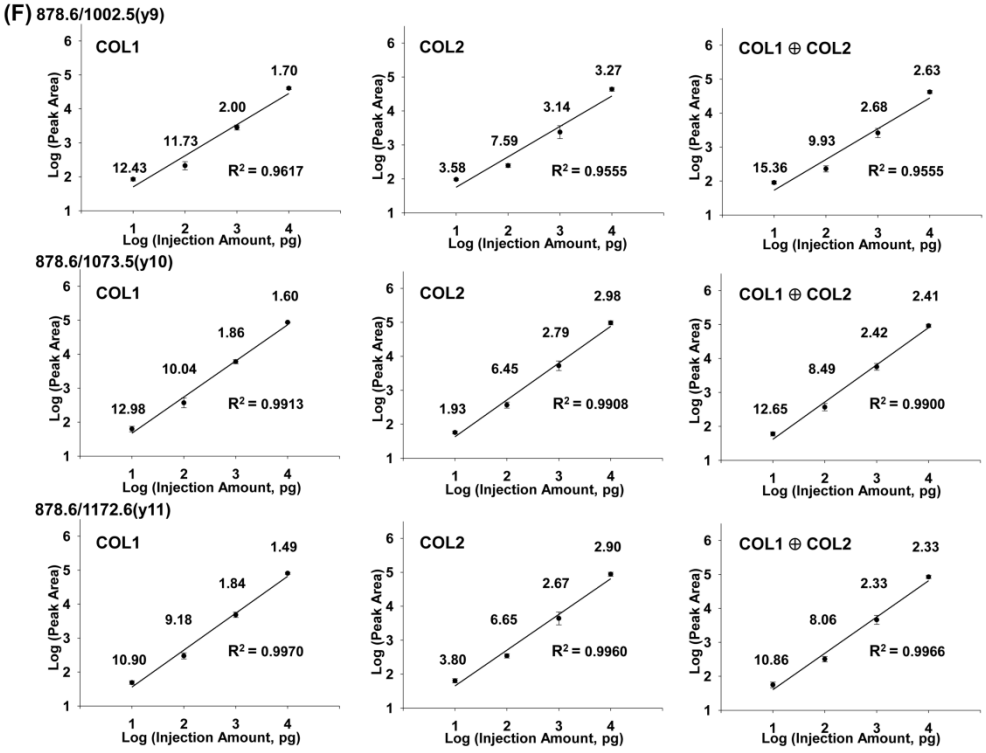
Fig. S2. Serial dilution curves of nineteen transitions from LC-MRM experiments in the presence of gastric tissue peptides as sample matrix using transition data of COL1 only (3 LC-MRM experiments), COL2 only (3 LC-MRM experiments), and using the transition data of both COL1 and COL2 (6 LC-MRM experiments). The numbers on the curves are the coefficient of variation (CV). : (A) ANLDVK, (B) IGSEVYHNLK, (C) GNPTVEVEITTEK, (D) VNQIGTLSESIK, (E) NVNDVIAPAFVK, (F) TAGIQIVADDLTVTNPK, (G) AVDDFLISLDGTANK




(C) 708.9/948.5(y8)
COL1
10.45
2.31
2.62
0.93
R² = 0.9971
COL2
3.03
1.26
6.03
0.93
R² = 0.9924
COL1 ⊕ COL2
11.19
2.02
4.72
1.18
R² = 0.9982
Log (Peak Area)
Log (Injection Amount, pg)


(D) 644.9/563.3(y5)
COL1
0.88
0.88
0.58
0.43
R² = 0.9940
COL2
3.72
1.00
0.56
0.43
R² = 0.9917
COL1 ⊕ COL2
4.40
1.19
0.60
0.44
R² = 0.9930
644.9/834.5(y8)
COL1
1.26
0.74
0.70
0.54
R² = 0.9944
COL2
1.68
0.70
0.98
0.48
R² = 0.9942
COL1 ⊕ COL2
1.84
1.07
0.94
0.51
R² = 0.9943
644.9/947.5(y9)
COL1
0.60
0.22
0.59
0.51
R² = 0.9935
COL2
1.21
0.73
0.69
0.52
R² = 0.9940
COL1 ⊕ COL2
1.21
1.06
0.71
0.51
R² = 0.9938
Log (Peak Area)
Log (Injection Amount, pg)


(E) 644.0/542.3(b5)
COL1
0.93
1.77
6.52
9.17
R² = 0.9957
COL2
1.23
1.13
5.50
4.56
R² = 0.9933
COL1 ⊕ COL2
1.09
1.50
6.14
7.24
R² = 0.9946
644.0/632.4(y6)
COL1
0.74
1.74
5.61
8.32
R² = 0.9958
COL2
0.84
1.77
1.03
5.27
R² = 0.9960
COL1 ⊕ COL2
0.80
1.77
1.45
7.00
R² = 0.9965
644.0/745.5(y7)
COL1
0.69
1.66
5.41
8.03
R² = 0.9970
COL2
0.86
0.93
4.52
3.63
R² = 0.9960
COL1 ⊕ COL2
0.79
1.36
5.08
6.35
R² = 0.9966
Log (Peak Area)
Log (Injection Amount, pg)


(F) 878.6/1002.5(y9)
COL1
1.70
2.00
11.73
12.43
R² = 0.9617
COL2
3.27
3.14
7.59
3.58
R² = 0.9555
COL1 ⊕ COL2
2.63
2.68
9.93
15.36
R² = 0.9555
878.6/1073.5(y10)
COL1
1.60
1.86
10.04
12.98
R² = 0.9913
COL2
2.98
2.79
6.45
1.93
R² = 0.9908
COL1 ⊕ COL2
2.41
2.42
8.49
12.65
R² = 0.9900
878.6/1172.6(y11)
COL1
1.49
1.84
9.18
10.90
R² = 0.9970
COL2
2.90
2.67
6.65
3.80
R² = 0.9960
COL1 ⊕ COL2
2.33
2.33
8.06
10.86
R² = 0.9966
Log (Peak Area)
Log (Injection Amount, pg)

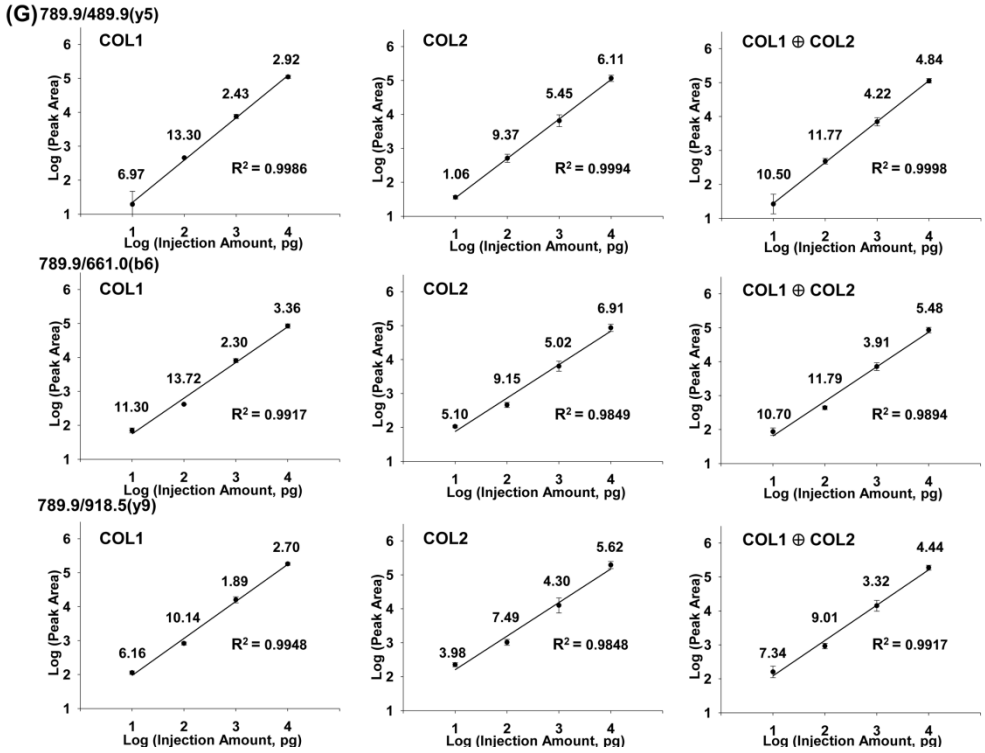
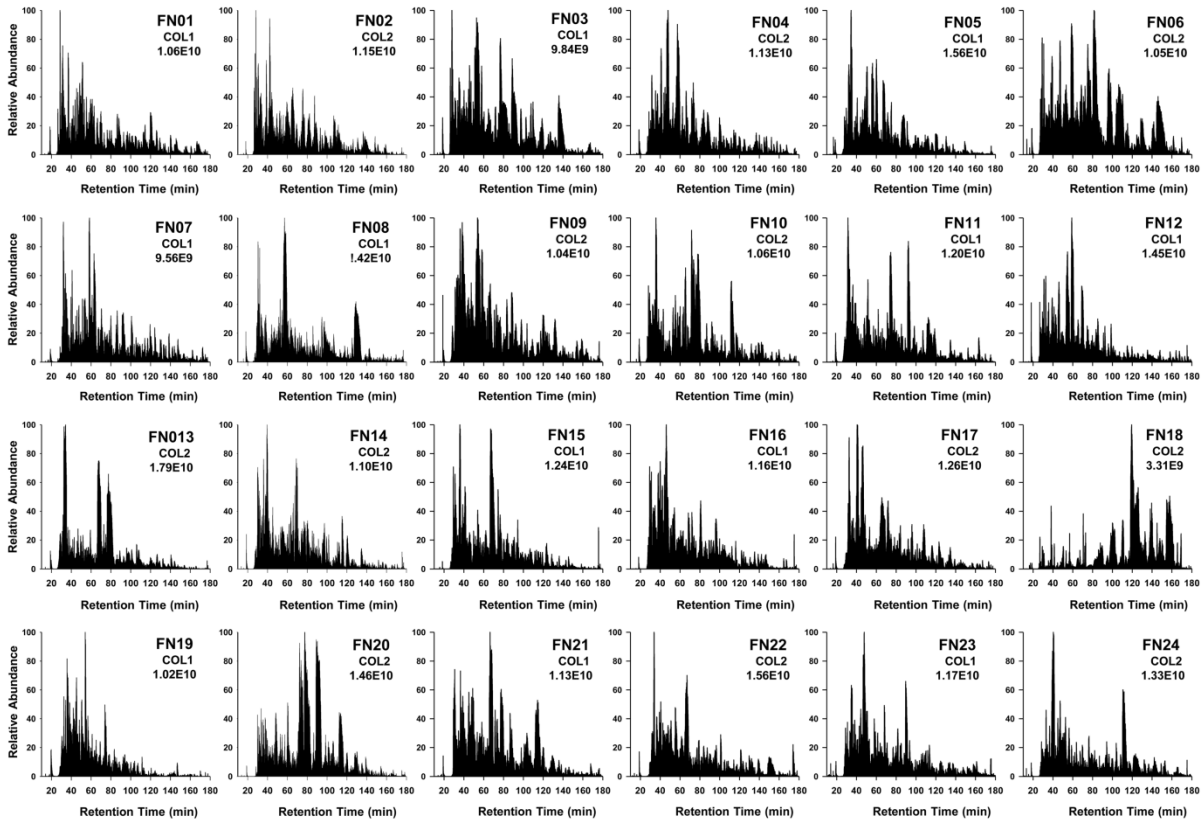


Fig. S3. Base peak chromatograms of twenty-four LC-MS/MS experiments for 24 fractions. Columns used and base peak intensities were indicated.



REFERENCES

1. J. R. Wisniewski, A. Zougman, N. Nagaraj and M. Mann, *Nat Methods*, 2009, **6**, 359-362.
2. Y. Wang, F. Yang, M. A. Gritsenko, Y. Wang, T. Clauss, T. Liu, Y. Shen, M. E. Monroe, D. Lopez-Ferrer, T. Reno, R. J. Moore, R. L. Klemke, D. G. Camp, 2nd and R. D. Smith, *Proteomics*, 2011, **11**, 2019-2026.
3. H. Li, K. B. Hwang, D. G. Mun, H. Kim, H. Lee, S. W. Lee and E. Paek, *Journal of proteome research*, 2014, **13**, 3488-3497.
4. B. Zhang, M. C. Chambers and D. L. Tabb, *Journal of proteome research*, 2007, **6**, 3549-3557.
5. S. J. Kim, S. Chae, H. Kim, D. G. Mun, S. Back, H. Y. Choi, K. S. Park, D. Hwang, S. H. Choi and S. W. Lee, *Molecular & cellular proteomics : MCP*, 2014, **13**, 811-822.
6. S. Kim and P. A. Pevzner, *Nature communications*, 2014, **5**, 5277.