Electronic Supplementary Material (ESI) for Molecular Omics.

Supporting Information

Comprehensive comparison of sample preparation workflows for proteomics

Weimin Zheng [a], Pengyuan Yang [a,b], Chuanyu Sun [c]*, and Yang Zhang [b]*

a. Department of Chemistry, Fudan University, Shanghai 200433, P. R. China

b. Department of Systems Biology for Medicine and Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, Shanghai 200032, P. R. China;

c. Department of Urology, Huashan Hospital, Fudan University, Shanghai, 200040, China

*Correspondence: Chuanyu Sun: zhugexianglong@163.com;

Yang Zhang: zhangyang@fudan.edu.cn.

Table of contents

.

.

The optimal protocol for protein preparation

1. human testis and epdidiymis were cut into pieces and washed with 2 mL PBS 3-4 times to remove blood on the ice. The precleaned tissues were ground to powder with nitrogen and take weight.
2. The powder was separately suspended into urea/thiourea lysis buff (buffer A: 7 M urea, 2M thiourea) with the mass ratio of 1:5 (powder: buffer, w: w). Protease inhibitor cocktail was also added to the mix with 50:1 (sample: protease inhibitor, v: v).
3. The sample was homogenized for 30 min, followed by sonication (Qsonica, Newtown, CT, USA) for 5 s × 10 times at 80 W.
4. The samples were centrifuged for 30 min at 14,000g, and supernatants were collected in the fresh tube.
5. Protein concentrations were determined using the 2-D Quant kit according to the manufacturer's instructions, and then samples were stored at -80 °C before use.
6. 200 μg extracts were transferred to a 1.5 mL tube. Then the proteins were diluted by 50 mM ABC to an approximate volume of 200 μL (the concentration of proteins was 1 μg/μL).
7. The proteins were added by 400 mM stock DTT to a final concentration of 10 mM DTT and incubated for 60 min for reduction.
8. Then proteins were added by 1 M stock IAA to a final concentration of 25 mM IAA, and alkylation at 37 °C for 30 min in the dark.
9. 6.7 μg of trypsin digestion was added to performed overnight at 37 °C (enzyme: protein ratio 1:30, w/w). The digestion was stopped by adding 0.5 μL of 20% TFA until the concentration of TFA was 0.05%.
10. Then the peptide mixture was desalted using Sep-Pak C18 columns according to the manufacturer's instructions. Briefly, the peptides were loaded on the column and washed with 1 mL of 0.1% TFA. The peptides were firstly eluted using 500 μL of 50% ACN, 0.1% TFA, then eluted by 500 μL of 80% ACN, 0.1% TFA. The solution was brought to dryness and resuspended in 10 μL of 0.1% FA, and 2% ACN.
11. For hi-pH RPLC: the solvents consisted of 1% ACN with 10 mM ammonium formate as mobile phase A, and 90% ACN with 10 mM ammonium formate as phase B. Solvents A and B were adjusted to pH 10.0 with ammonia.
12. 10 μL digested proteins were injected into the C18 column (2.1 x 250 mm, X Bridge BEH300, Waters Corporation, Milford, MA, USA) with a 100 μL loop connected to a Nano ACQUITY UPLC system (Waters Corporation, Milford, MA, USA).
13. The flow rate was 600 uL/min with the following gradient: 1- 35% B from 2 min to 22 min, 35-80% B over 2 min, held at 80% B for 2 min, then reduced to 1% B over 2 min. The peptide elution profile was monitored using UV absorbance at 214 nm.
14. 32 initial fractions were collected from 2.5 min to 20 min, then the 32 fractions were pooled to 16 from beginning to end. The details of how to collect the 32 initial fractions and pool to 16 fractions were shown in Table S8.
15. Pooled fractions were concentrated using a SpeedVac, then each fraction was resuspended in 10 μL of 2% ACN with 0.1% FA.
16. The fractions were performed using an ultra-performance liquid chromatography system ACQUITY UPLC system (Waters Corporation, Milford, MA, USA) coupled

with an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Mobile phase A was composed of water and 0.1% (v/v) FA, while mobile phase B was composed of ACN and 0.1% (v/v) FA.

17. 2 μL of samples were loaded onto the C18 Nano Trap Column (Waters Corporation, 20 mm ×180 μm, 5 μm) for 3 min at 2 μL/min.
18. Then samples were analyzed on a C18 column (Thermo Fisher Scientific, 150 mm×75 μm, 3 μm) and eluted for 120 min at 300 nL/min using the following linear gradient: from 2% to 25% B in 100 min, from 25% to 50% B in 9 min, a 3 min wash at 90% B, from 90 to 2% in 0.01 min, followed by 5.99 min post-run equilibration at 2% B.
19. Survey scans of peptide precursors from 400 to 1,600 m/z were performed at 120,000 resolution with a 10^6 ion count target. The 15 most abundant ions in each MS scan were automatically selected and fragmented in HCD mode to achieve high mass accuracy in MS/MS spectra. For MS/MS analysis, the isolation window was set as 2.0 Da, normalized collision energy as 30.0, activation time as 0.1ms, and the starting mass as 100.0 Da.
20. Raw files were searched using MaxQuant version 2.0.1.0 (Martinsried, Bavaria, Germany), using the Andromeda search engine against the human database (Version Jun 2021, 20396 entries) selecting trypsin digestion. Search settings included up to two mis-cleavages, carbamidomethyl cysteine (monoisotopic mass change, + 57.02 Da) as a fixed modification. Oxidation of methionine and acetyl of the protein N-terminus were specified as variable modifications. The instrument selected was the Orbitrap. The main search peptide tolerance and MS/MS tolerance were set to an initial mass tolerance of 4.5 ppm and 20 ppm, respectively. Minimal peptide length was set to 7 amino acids and a maximum of two mis-cleavages was allowed. Peptide identifications and protein identifications were accepted with an FDR <1.0%.

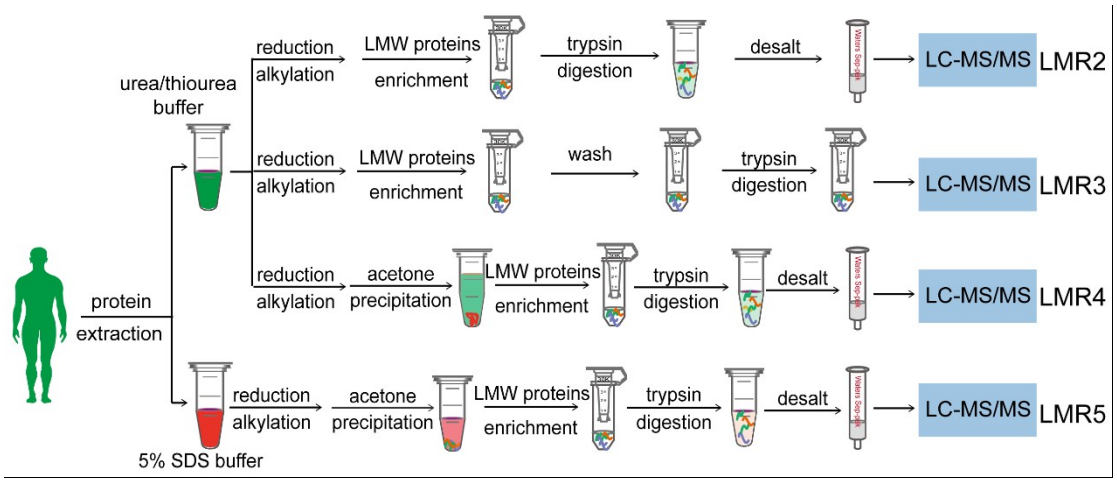


Figure S1. Experimental design for comparing LMW protein enrichment strategies. The protein extraction proteins digestion and desalting were performed as described in Figure 1. The enrichment of LMW proteins was performed using 30 kDa filters. The detailed processes were described as follows: LM-R2, urea/thiourea buffer and the in-solution digestion; LM-R3, urea/thiourea buffer, and the FASP method; LM-R4, urea/thiourea buffer, with acetone precipitation; LM-R5, 5% SDS buffer, with acetone precipitation.

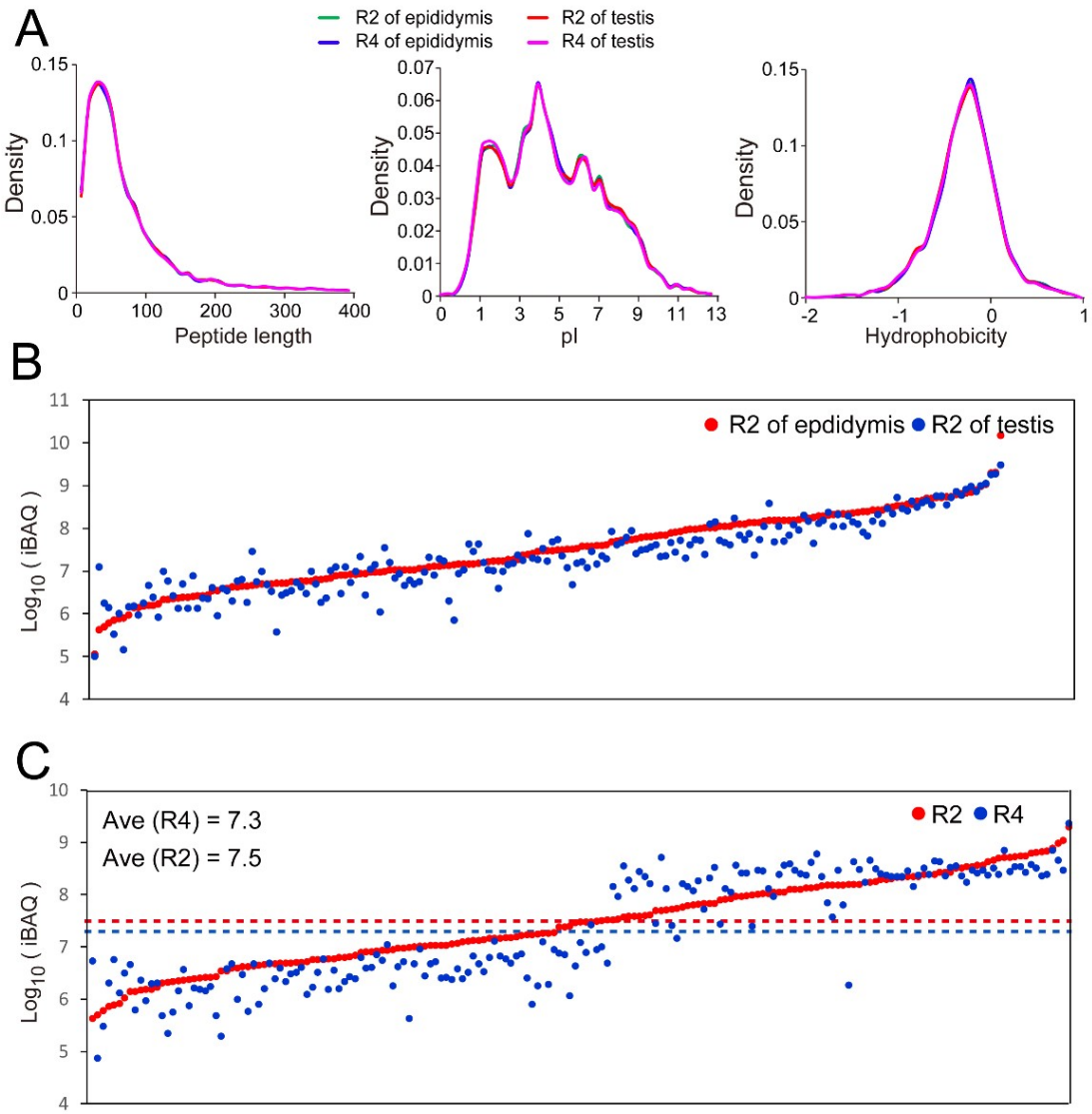


Figure S2. Evaluation of the influence of acetone precipitation on protein identification by LC-MS. (A) The distribution of the whole proteome of R2 and R4 according to peptide length, pI, and hydrophobicity; (B) The relative quantitative value of ribosomal proteins in the testis and epididymis in the R2 route; (C) The relative quantitative value of ribosomal proteins in the epididymis.

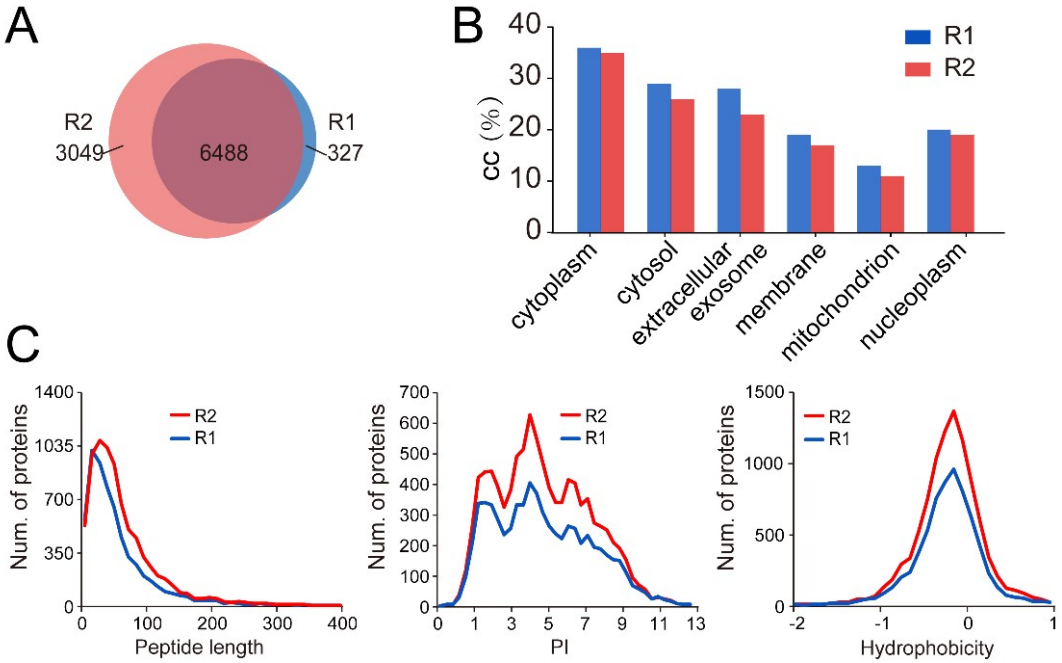
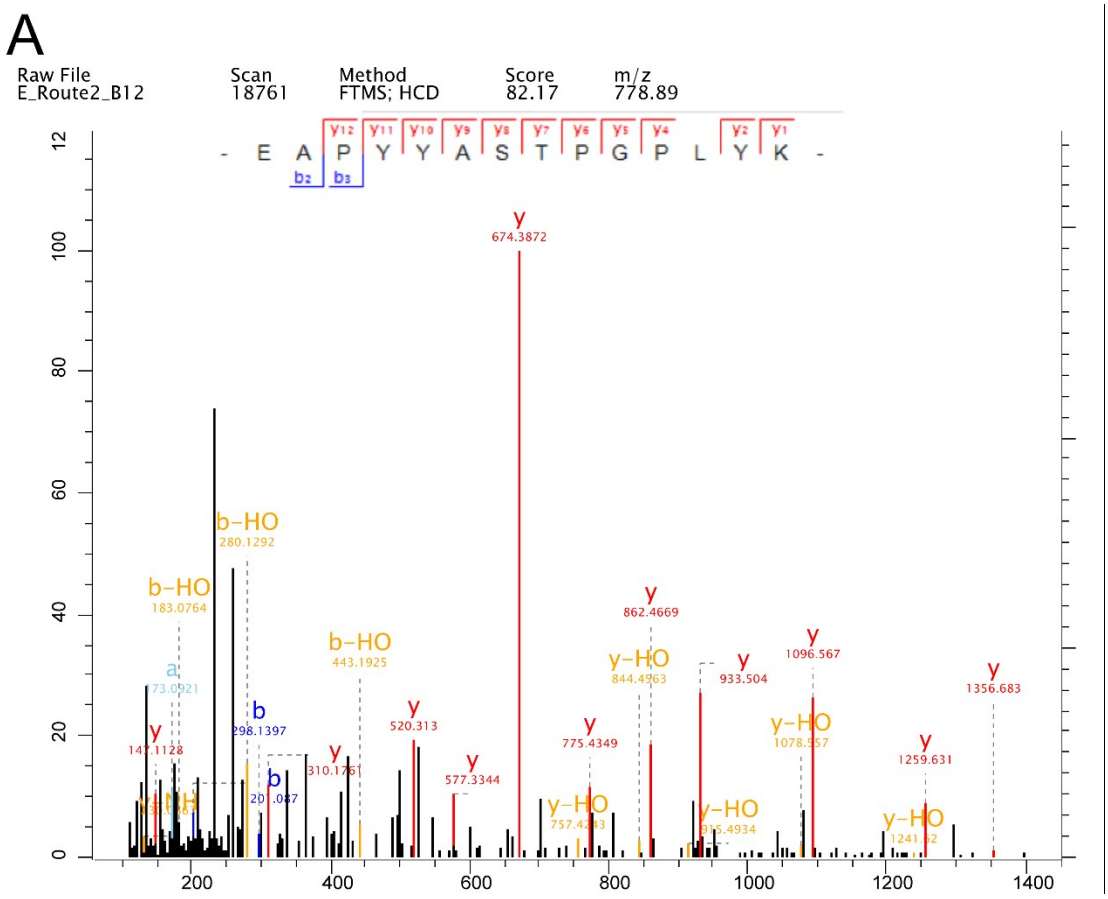


Figure S3. Comparison of R2 and R1 as one-dimensional separation. (A) Overlap of these two routes; (B) GO annotation of two routes and the top six enriched cellular components were displayed; (C) Distribution of identified proteins according to peptide length, pI, and hydrophobicity.


A
Raw File
E_Route2_B12
Scan
18761
Method
FTMS; HCD
Score
82.17
m/z
778.89
- E A P Y Y A S T P G P L Y K -
b-HO
280.1292
b-HO
183.0764
b-HO
443.1925
y
674.3872
y
862.4669
y-HO
844.4563
y
933.504
y
1096.567
y-HO
1078.557
y
1356.683
y
1259.631
y-HO
1241.62
y-HO
915.4934
y-HO
757.4243
y
775.4349
y
520.313
y
577.3344
y
310.1761
b
298.1397
b
201.087
a
73.0921
y
147.1128
0
20
40
60
80
100
12
200
400
600
800
1000
1200
1400

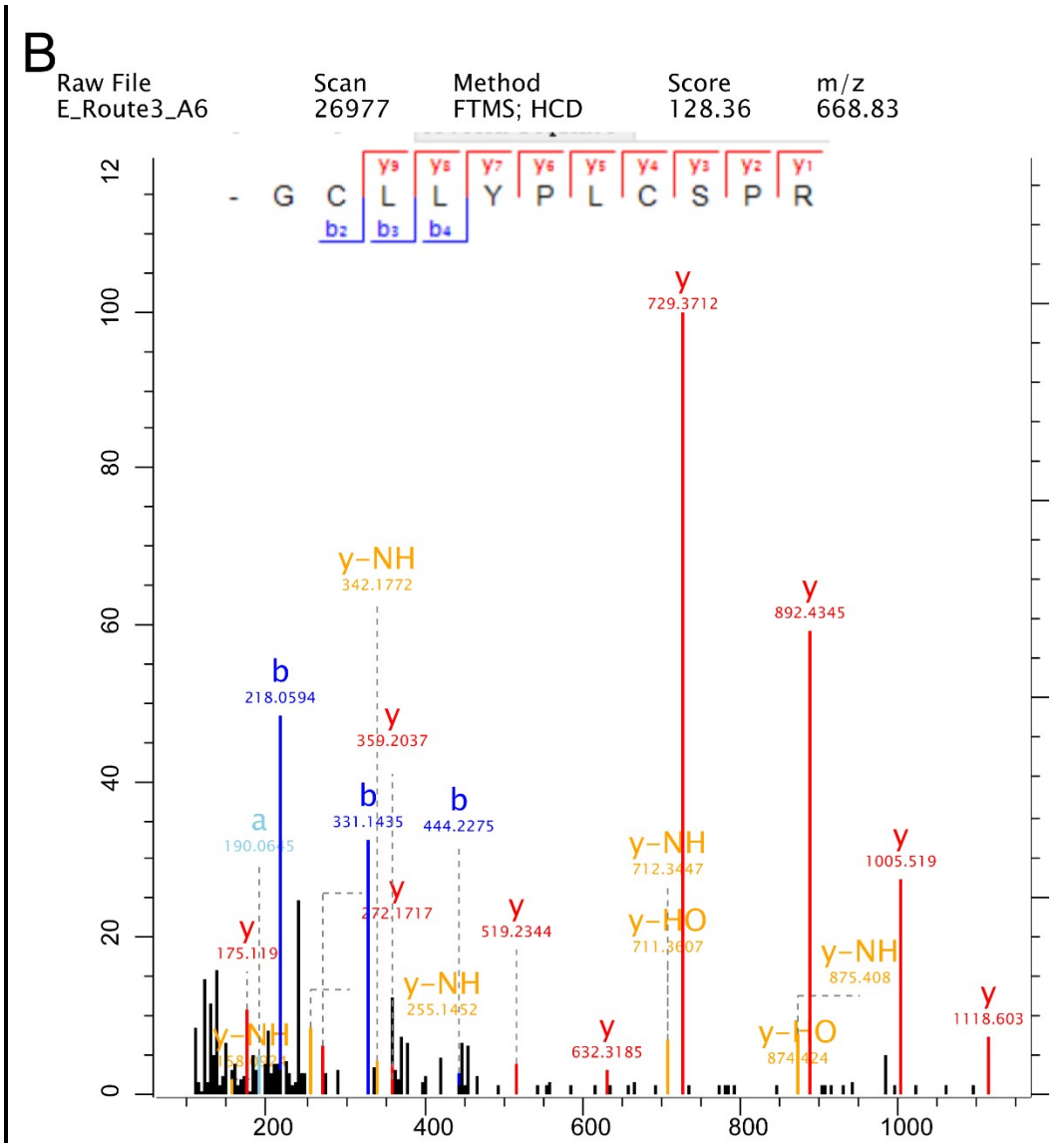
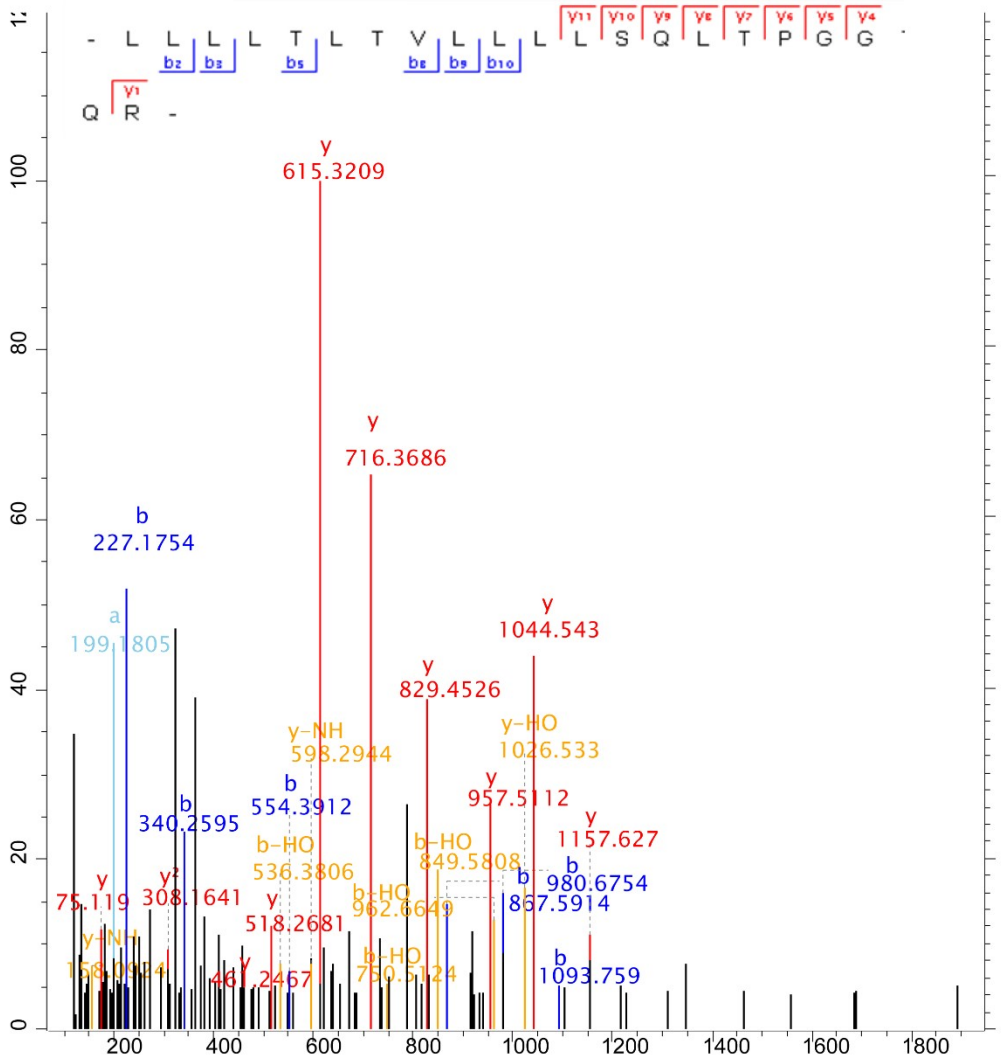


Figure S4. MS/MS spectrum of unique peptides (A) EAPYYASTPGPLYK and (B) GCLLYPLCSPR from candidate MP Protein TEDDM1 (Q5T9Z0) in this study.

A

Raw File	Scan	Method	Score	m/z
2D_E_Route2_A2	33270	FTMS; HCD	76.03	788.83



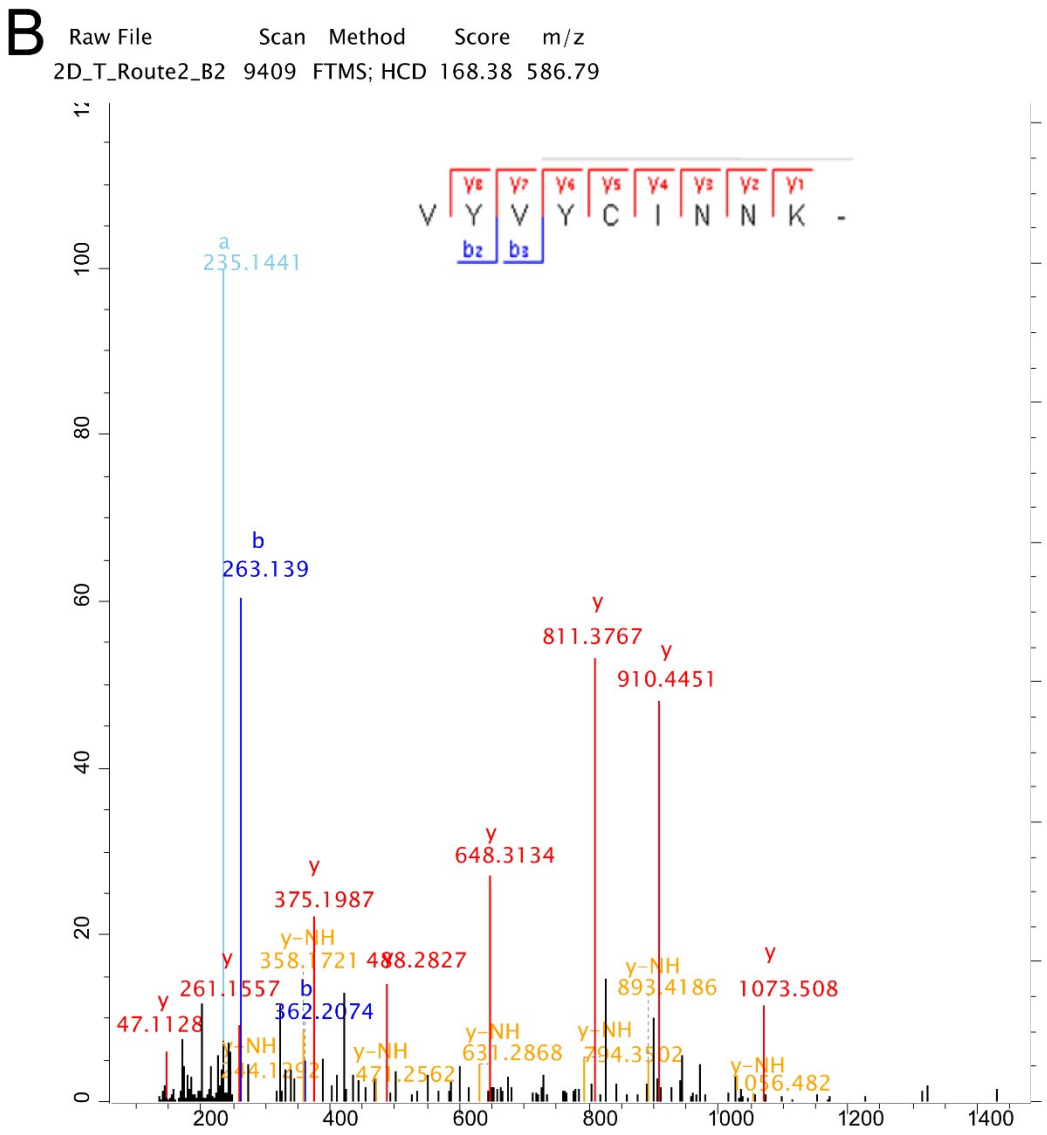
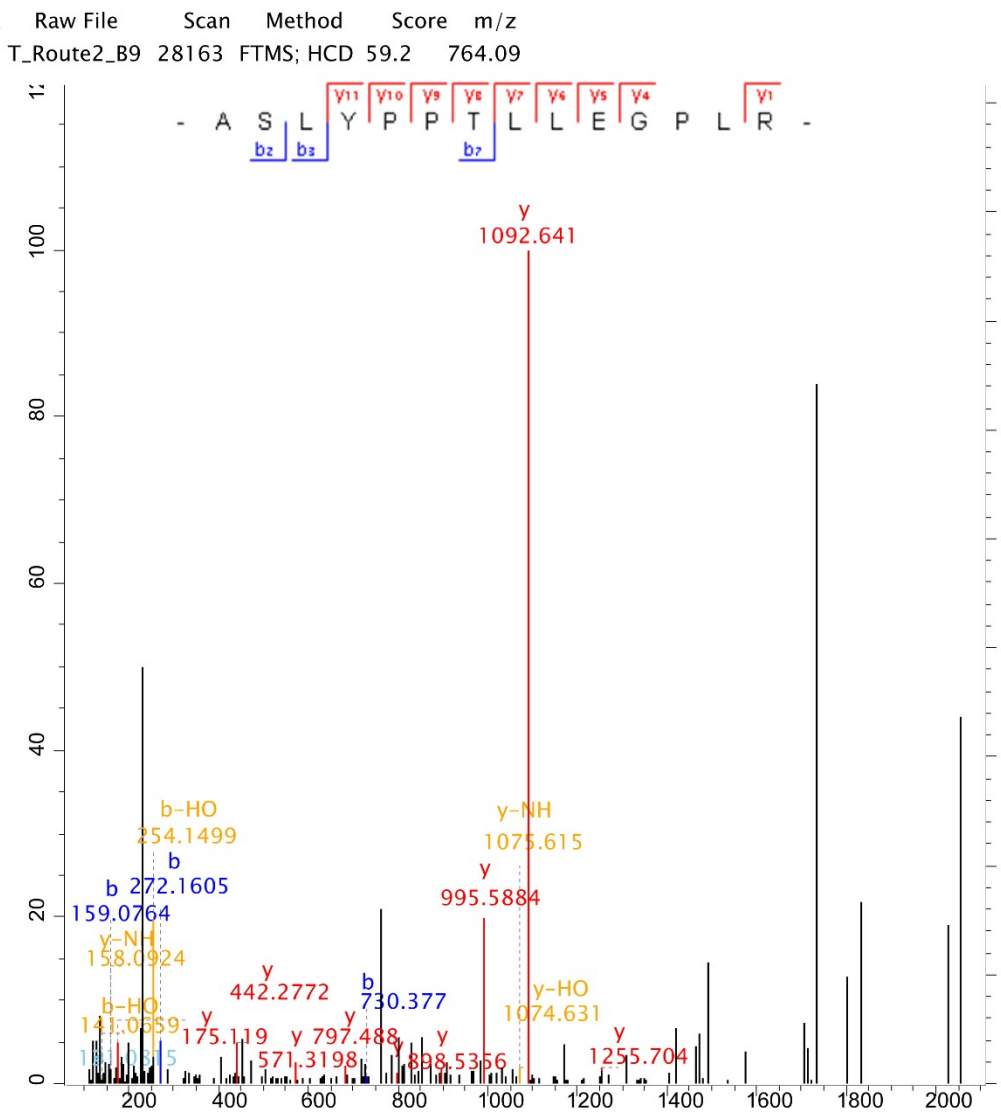


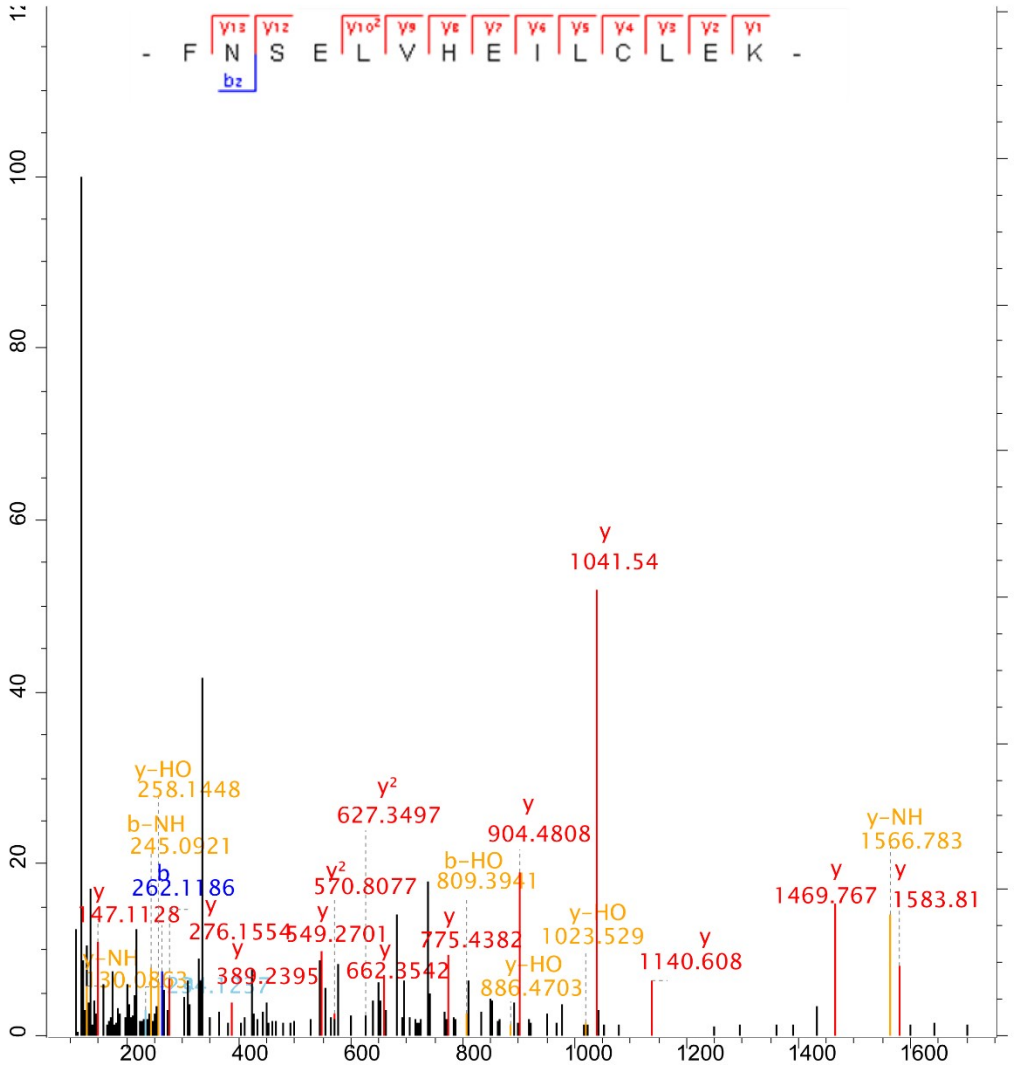
Figure S5. MS/MS spectrum of unique peptides (A) LLLLTLTVLLLLSQLTPGGTQR and (B) VYVYCINNK from candidate MP Protein DEFB123 (Q8N688) in this study.

A


Raw File Scan Method Score m/z
T_Route2_B9 28163 FTMS; HCD 59.2 764.09
- A S L Y P P T L L E G P L R -
y11 y10 y9 y8 y7 y6 y5 y4 y1
b2 b3 b7
y
1092.641
b-HO
254.1499
b
b 272.1605
159.0764
y-NH
158.0924
y
442.2772
b
730.377
y-NH
1075.615
y
995.5884
y-HO
1074.631
b-HO
141.0559
y
175.119
y 797.488
y
571.3198
y
898.5356
y
1255.704
0
20
40
60
80
100
200
400
600
800
1000
1200
1400
1600
1800
2000

B

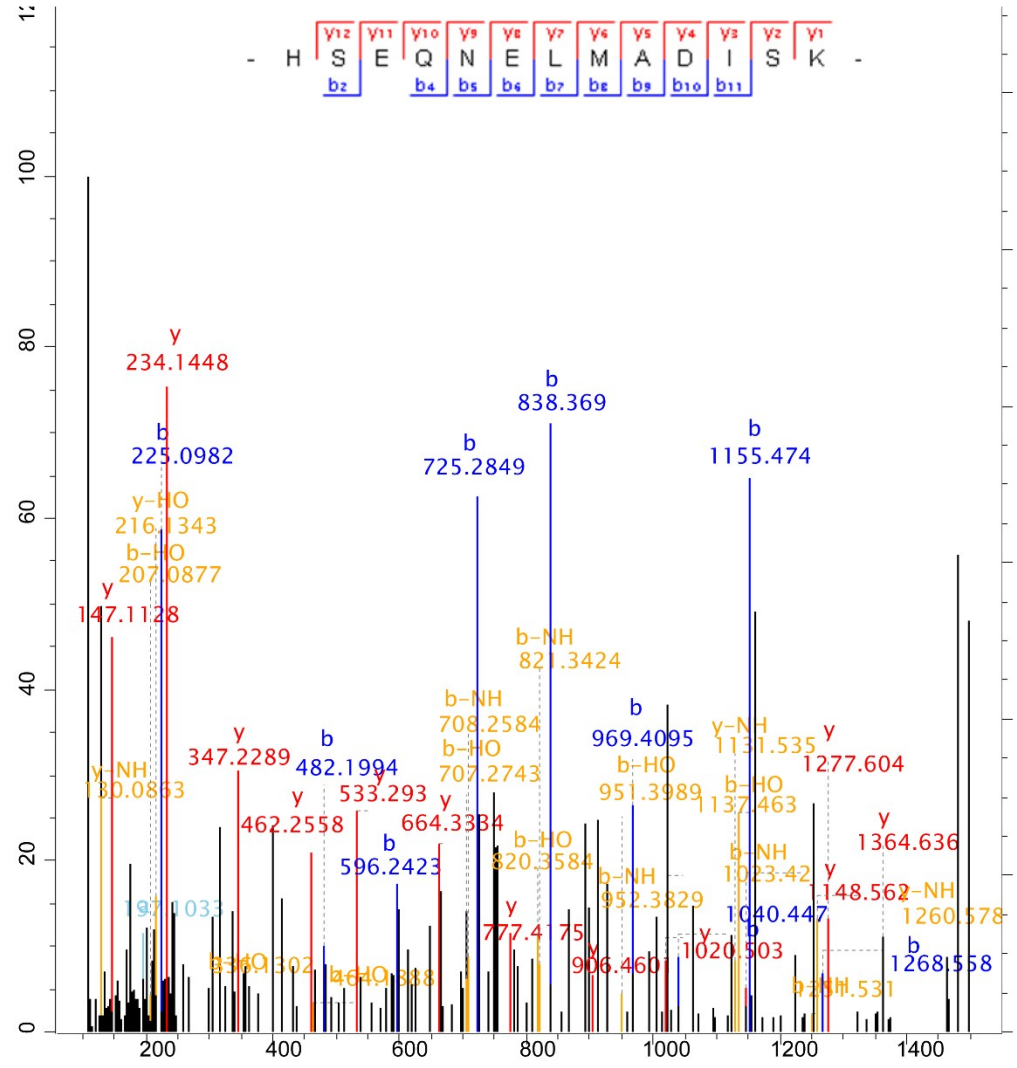
Raw File	Scan	Method	Score	m/z
T_Route2_A12	29002	FTMS; HCD	63.29	577.63



C

Raw File Scan Method Score m/z

T_Route2_A5 13135 FTMS; HCD 185.57 751.35



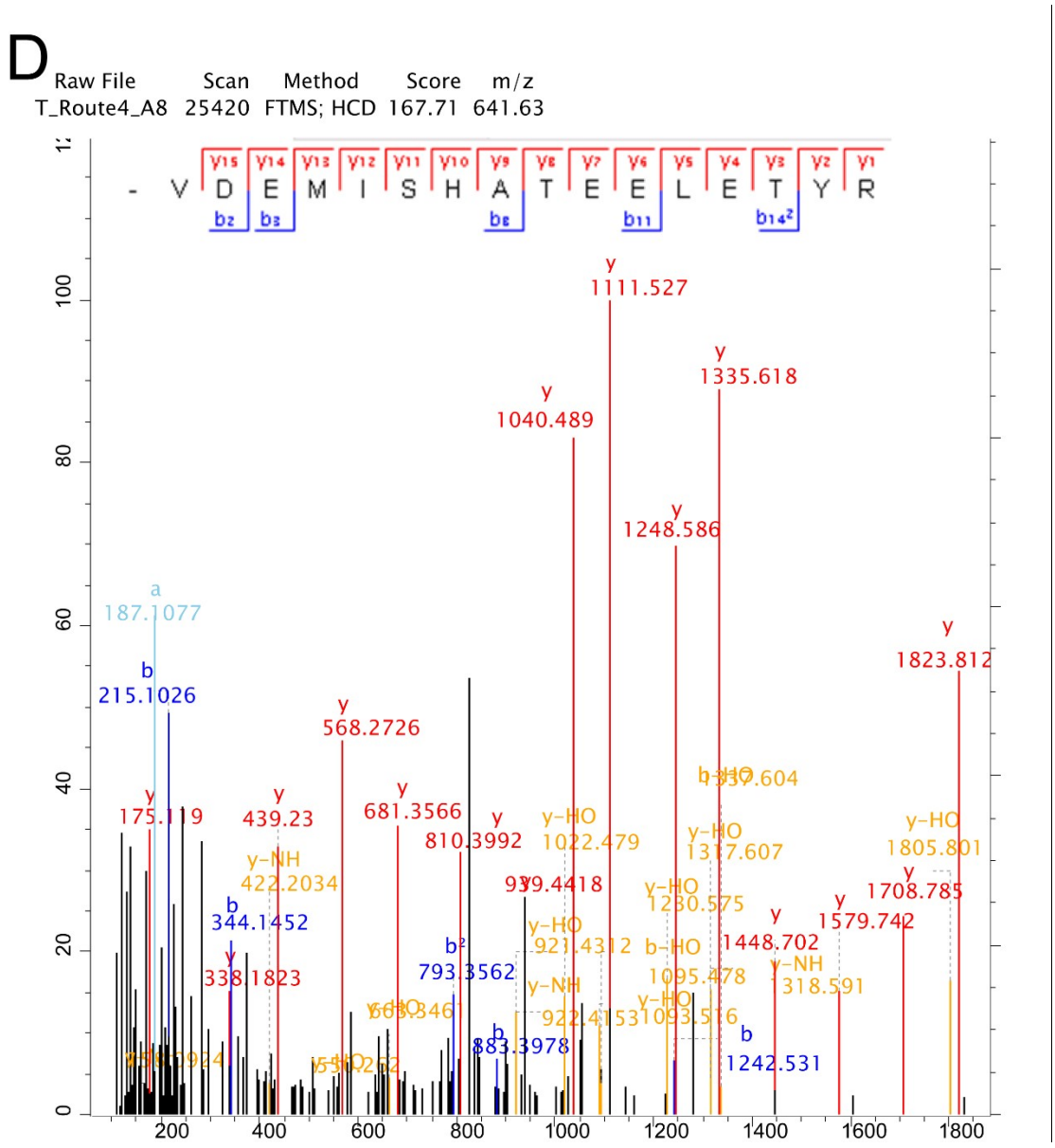


Figure S6. MS/MS spectrum of unique peptides (A) ASLYPPTLLEGPLR, (B) NSELVHEILCLEK, (C) HSEQNELMADISK, and (D) VDEMISHATEELETYR from candidate MP Protein CTAGE1 (Q96RT6) in this study.