Supporting Information

Specific enrichment of peptides with N-terminal serine/threonine by a solid-phase capture-release approach for efficient proteomics analysis

Yating Yao[a,b], Junfeng Huang[a,b], Kai Cheng[a,b], Yanbo Pan[a,b], Hongqiang Qin[a], Mingliang Ye[a,*], Hanfa Zou[a,*]

[a] Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences (CAS), Dalian, 116023, China

E-mail: mingliang@dicp.ac.cn, hanfazou@dicp.ac.cn; Fax: +86-411-84379620; Tel: +86-411-84379620, +86-411-84379610

[b] University of Chinese Academy of Sciences, Beijing 100049, China

SUMMARY

This supporting information file includes additional results and information as described in the text of the main article. Including:

Supplementary figures

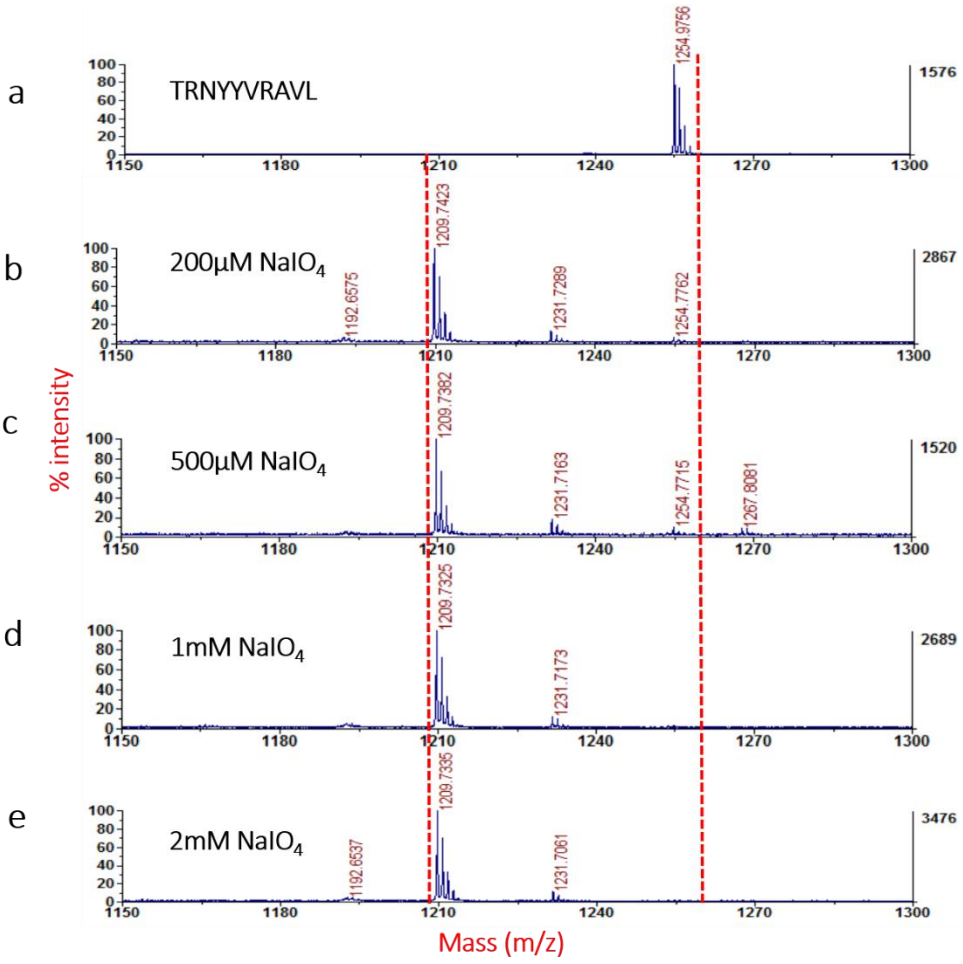

Figure S1. Optimization of sodium periodate concentration for the oxidation of N-terminal threonine. MALDI mass spectra of the standard peptide TRNYYVRAVL containing N-terminal threonine: (a) native peptide and (b-e) peptide oxidized at 200 µM $NaIO_4$, 500 µM $NaIO_4$, 1 mM $NaIO_4$ and 2 mM $NaIO_4$, respectively.

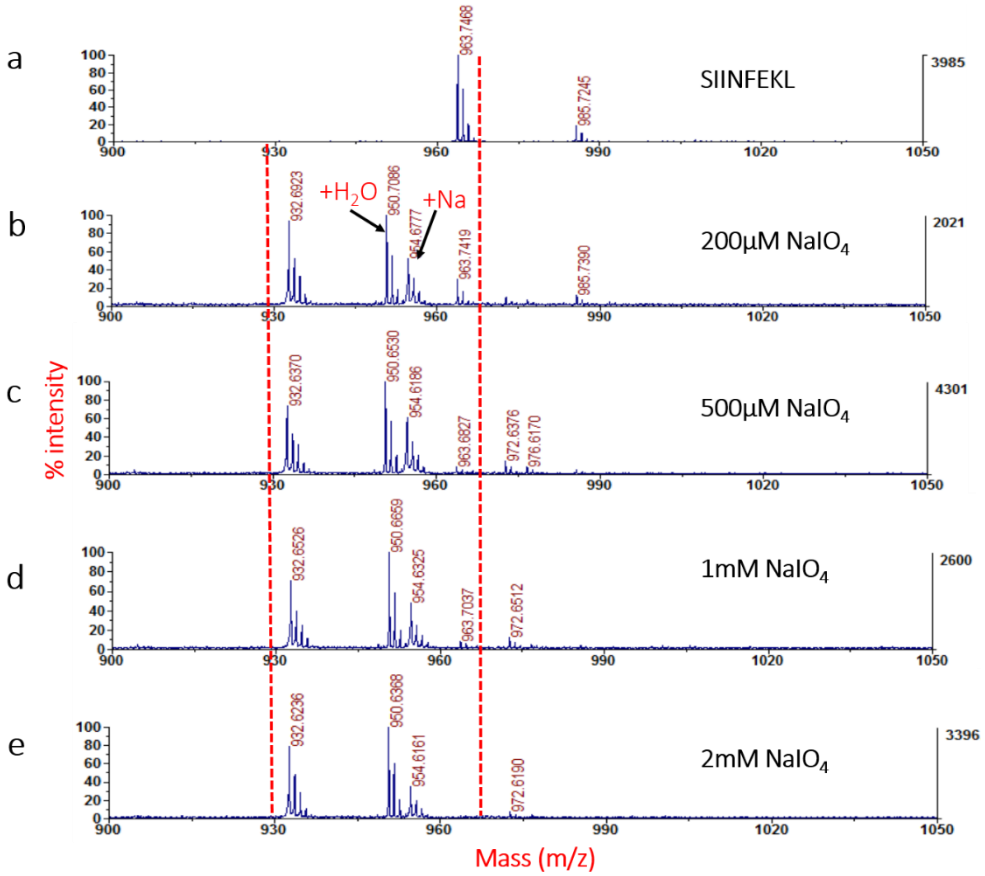

Figure S2. Optimization of sodium periodate concentration for the oxidation of N-terminal serine. MALDI mass spectra of the standard peptide SIINFEKL containing N-terminal serine: (a) native peptide and (b-e) peptide oxidized at 200 μM $NaIO_4$, 500 μM $NaIO_4$, 1 mM $NaIO_4$ and 2 mM $NaIO_4$, respectively.

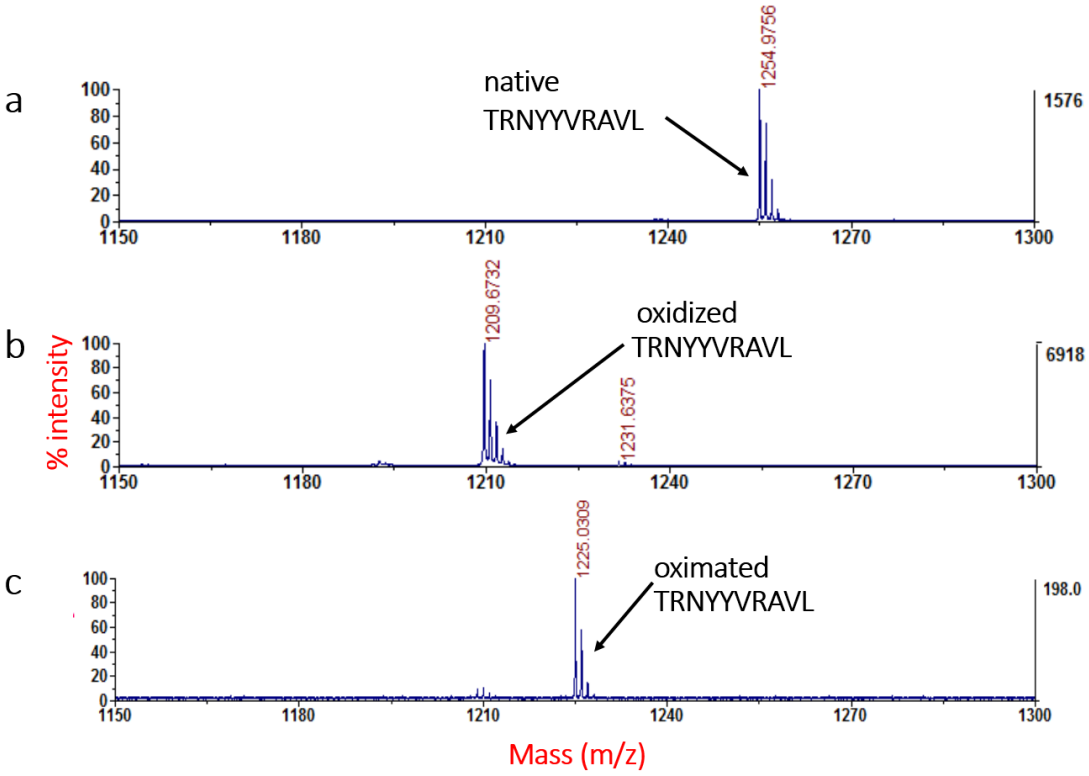

Figure S3. Enrichment of the standard peptide TRNYYVRAVL with N-terminal threonine: (a) original peptide, (b) peptide oxidized at 2 mM $NaIO_4$, and (c) peptide released from hydrazide beads with hydroxylamine chloride.

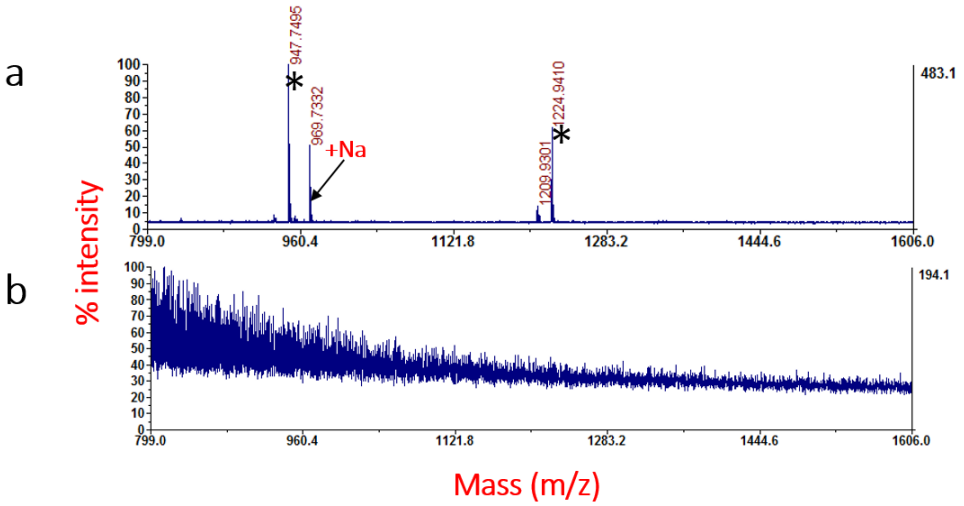

Figure S4. Investigating the efficiency of the release step with standard peptide mixture. The peptide mixture containing two standard peptides, SIINFEKL with N-terminal serine, TRNYYVRAVL with N-terminal threonine were coupled onto the hydrazide beads and released with two successive elutions with hydroxylamine chloride. MALDI mass spectra of: (a) the first elute, and (b) the second elute (peaks with N-terminal serine/threonine peptides are marked with asterisks).

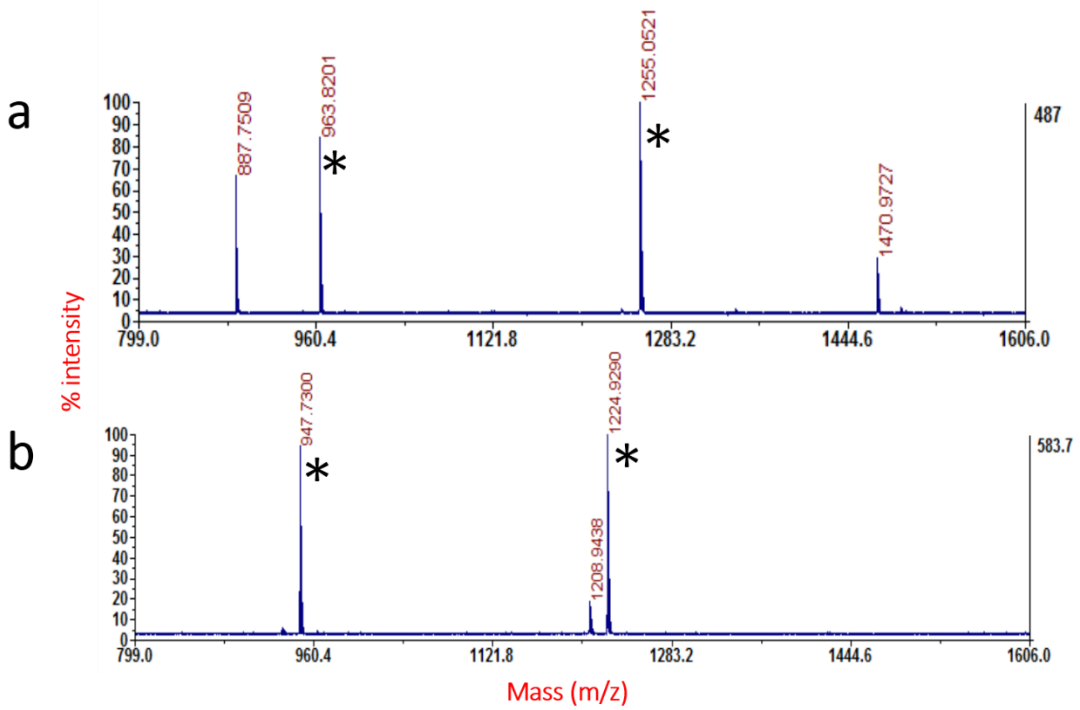

Figure S5. Specificity investigation of the enrichment strategy with standard peptide mixture. MALDI mass spectra of the peptide mixture containing four standard peptides, SIINFEKL with N-terminal serine, TRNYYVRAVL with N-terminal threonine, GRRNSIGK and EESLEDSDVDADF with neither N-terminal serine nor N-terminal threonine: (a) peptide mixture before enrichment and (b) peptide mixture after enrichment (peaks with N-terminal serine/threonine peptides are marked with asterisks).

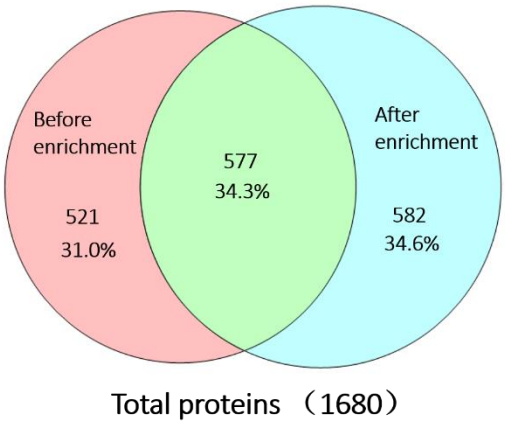

Figure S6. Complementary protein identifications by selective enrichment of N-terminal serine/threonine peptides from mouse liver samples digested by Lys-C.

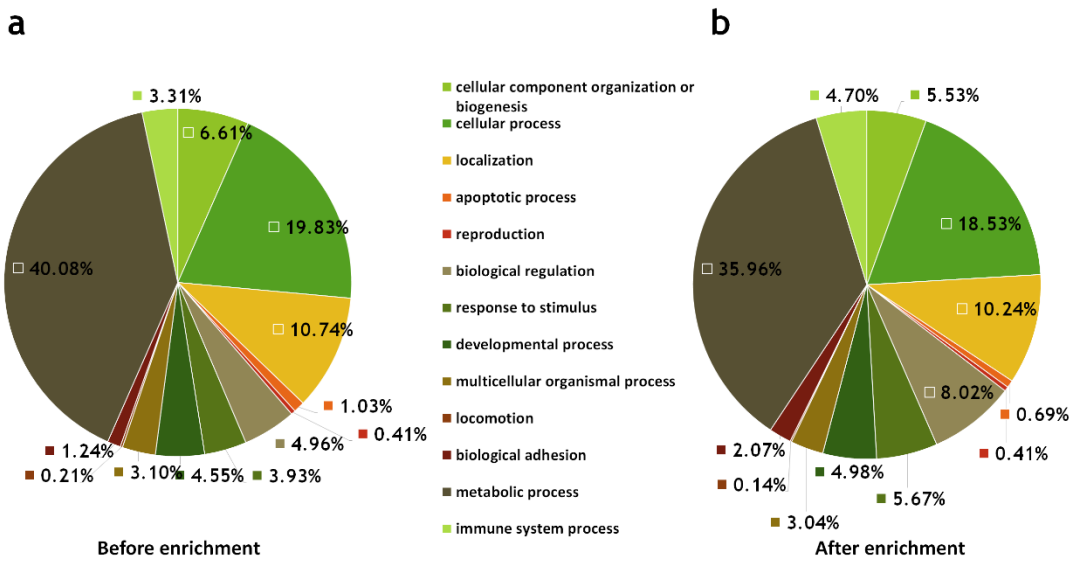

Figure S7. Pie chart representations of the distribution of uniquely identified proteins from the trypsin digested mouse liver sample with and without enrichment according to their biological processes.

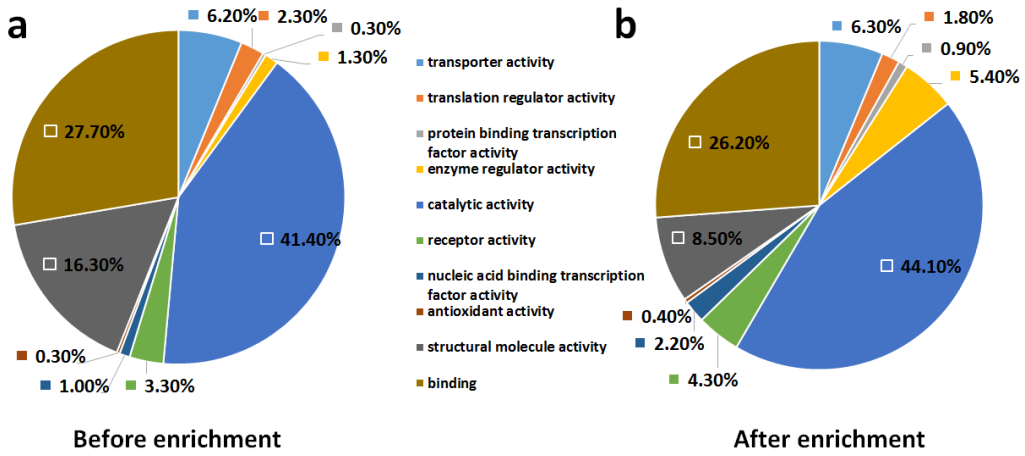

Figure S8. Pie chart representations of the distribution of uniquely identified proteins from the trypsin digested mouse liver sample with and without enrichment according to their molecular functions.

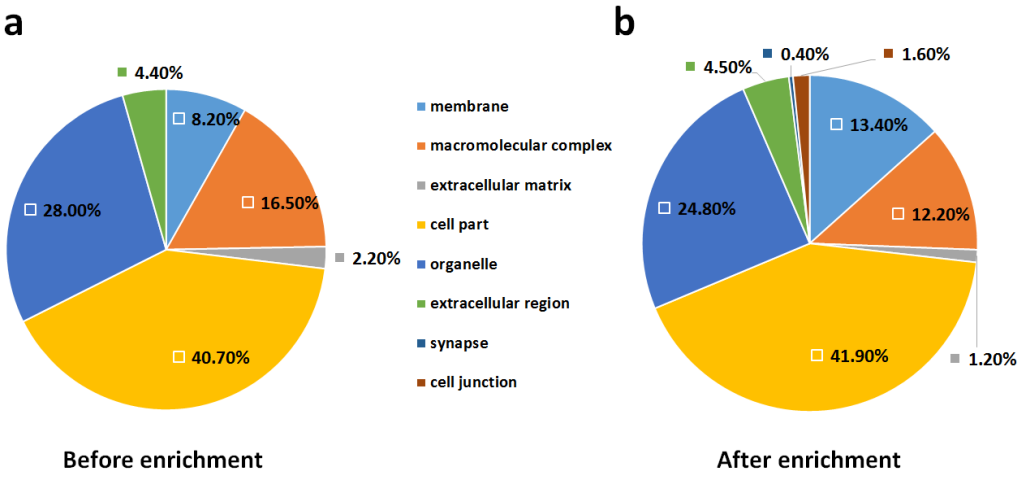

Figure S9. Pie chart representations of the distribution of uniquely identified proteins from the trypsin digested mouse liver sample with and without enrichment according to their cellular components.

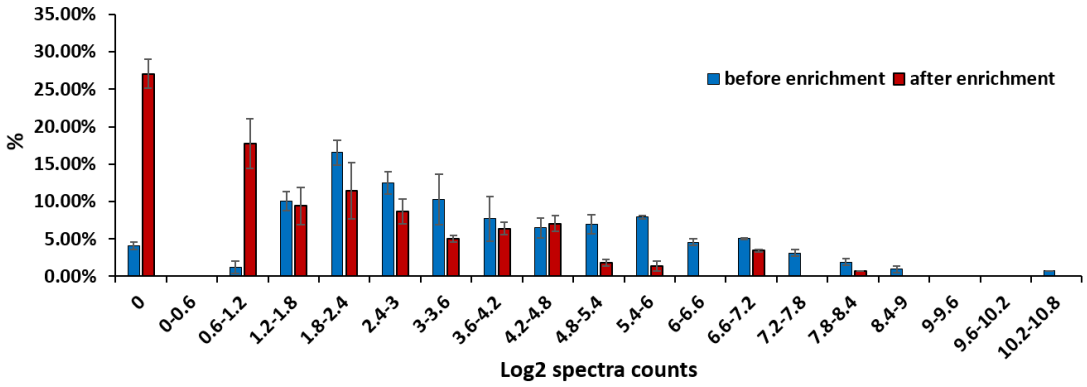

Figure S10. The distribution of the identified serum proteins from tryptic digests before and after enrichment across the spectra counts. The above data were averaged from 3 replicates (error bars represent the standard deviation).

Supplementary table

Table S1. The number of peptides containing the target sequences and their representative coverage of the mouse proteome based on in silico digestion.

	Total peptide number	#.S/T containing peptide number	% of peptide	% of coverage
Trypsin digestion	3107991	440429	14.17%	88.17%
Lys-C digestion	1452193	187402	12.90%	72.75%
Arg-C digestion	1506814	184495	12.24%	70.69%
Glu-C digestion	1957465	225212	11.50%	73.14%

Numbers were obtained through in silico digestion of the protein sequences in the mouse proteome from the UniProtKB/Swiss-Prot database with 51221 total protein entries. Peptides with 6-30 residues and with 0-1 missed cleavage site were considered.