Evaluation of Different Porcine and Bovine Trypsins Yield Different Cleavage Results

Scott Walmsley*[1], Paul Rudnick[2], Yuxue Liang[2], Qian Dong[2], Stephen E. Stein[2], Alexey Nesvizhskii[1]

[1]*University of Michigan Department of Pathology, Ann Arbor, Michigan*

[2]*National Institute of Standards and Technology, Gaithersburg, Maryland*

Abstract

Trypsin is an endoprotease commonly used for sample preparation for mass spectrometry based proteomics and is typically either porcine (P) or bovine (B) sequence in origin. While the general conditions for optimum trypsin activity are well understood, less is known about its reproducibility and specificity from the various suppliers. Further characterizing trypsin activity would improve the reliability of peptide detection and quantitation for both SRM-based and complex sample biomarker discovery studies. As such, we evaluated six available proteomics-grade trypsins.

Trypsin performance was assayed using a highly purified human serum albumin as the substrate. The human serum albumin sample was aliquoted and then digested in triplicate for 2 or 18 hours for each typsin, and then desalted and analyzed by reversed phase tandem mass spectrometry. Spectra were identified using MSPepSearch,v.0.9 and using a comprehensive human serum albumin spectral library (NIST format,v.071711) as the search space. Peptide intensities (MS1) were calculated using ProMS (NIST,v.May 25,2011).

Percent deviation were generally low (< 5%) for metrics relevant for digestion reproducibility (identifications of spectra, unique ions and unique peptides). Unique peptide and ion identifications were generally higher in the 18 hour digests suggesting an increase toward complete digestion. At the 18hr time point, total unique peptide identifications were 53% higher for the bovine trypsins versus the porcine trypsins. There also were observed differences of unique peptide counts for the fully tryptic (FT), missed cleavage (MC) or semi cleavage (SC) peptides between the porcine and bovine trypsins. Analysis of the MS1 intensities indicated significantly altered abundances between the bovine and porcine trypsins (FT: 9 peptides, MC: 8 peptides, SC: 3 peptides, FC>2, p<0.01). Taken together, these results identify trypsin activity that is reproducible for each manufacturer yet produce significant and reproducible differences between the bovine and porcine trypsins.

A. Materials and Methods (Figure 1)

- Trypsins were selected from the sources listed in table 1.
- A highly purified HSA was aliquoted from the same purification to perform the digests.
- Trypsin digests were repeated in triplicate for each trypsin and for 2 or 18 hours.
- C18-LCMSMS was using LTQ-XL MS of the HPLC separated (60 min gradient) peptides.
- Data analysis was performed using the NIST_MSQC pipeline together with an in-house software pipeline to perform additional statistics.
- Spectra were matched to a comprehensive human serum albumin spectral library.

Figure 1. Analysis Scheme

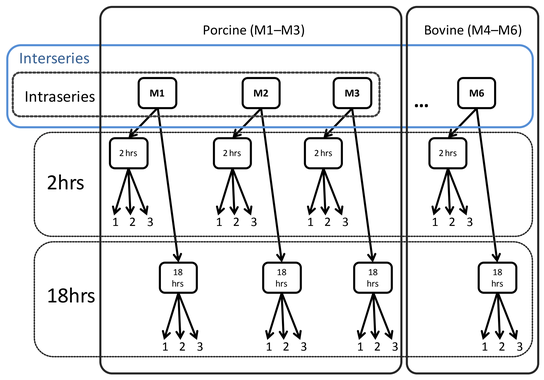
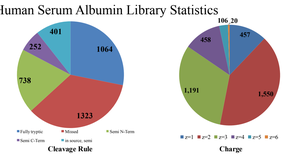


TABLE 1

Supplier	*Cat#*	*Sequence*
G Biosciences	786-245	Porcine
Princeton Separations	EN-151	Porcine
Promega	V511A	Porcine
G Biosciences	786-245B	Bovine
Roche-Diagnostics	11418025001	Bovine
Worthington	LS02120	Bovine

Human Serum Albumin Library Statistics



B. Evaluation of Reproducibility of Individual Trypsins

- Commonly used metrics (spectral counts, ion counts, peptide counts,ID's) produced the expected low intra-series variation (<5%), however inter-series variation of these metrics was increased. (Table 2).
- Intra-series deviation decreased when the results were classified by origin of trypsin (porcine versus bovine).
- Increasing digestion time from 2 to 18 hours generally decreased the variability of sample digestion, while the difference between the bovine and porcine groups increased.
- The ratio of those ID's dependent on charge ($z^+/2^+$) together the number of IDs at that charge suggested that there were significant differences in the complement of peptides that were identified using either the porcine or bovine trypsins.

TABLE 2 ***Metrics for analysis of reproducibility***

Metrics from Rudnick et. Al (2010) and described in the text. Mean ± %dev of each metric are reported for each trypsin (M1-M6). Interseries mean±%dev are calculated for all trypsins combined. Porcine or Bovine±%dev are the reported values for the intraseries (between porcine (M1-M3) or bovine(M4-M6) trypsins. Bov/Por % change is the % difference between the mean bovine and porcine reported values. Generally, reported deviations were low, but decreased further from 2 to 18 hrs. Interseries deviations increased from 2 to 18 hrs and indicating greater differences for the reported intraseries values (porcine vs bovine). Bold face type indicates selected metrics whose %dev were higher than 5%. DS: dynamic sampling, P: peptide, IS:ionsource, PL: peptide length. For a further explanation of the DS,P and IS metrics, see Rudnick et. Al (2010).

	Metric Code	*Description*	M1		M2		M3		M4		M5		M6		Inter Series		Porcine		Bovine		*Bov/Por*
			mean	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*%change*
	DS2-B	Spectrum Counts	10081.7	0.7	10085.3	0.5	10073.3	0.6	10187.0	0.8	10099.0	0.5	10121.3	0.5	10107.9	0.4	10080.1	0.1	10135.8	0.5	0.55
	P-2C	Tryptic unique peptides	108.7	3.0	107.7	4.2	96.0	1.8	139.3	2.5	130.3	1.9	146.3	2.1	121.4	17.2	104.1	6.8	138.6	5.8	33.13
	P-2B	Tryptic unique ions	312.0	2.0	296.0	3.3	266.3	1.5	333.3	2.9	320.0	1.7	349.7	0.9	312.9	10.4	291.4	8.0	334.3	4.4	14.72
2Hrs	P-2A	Tryptic total IDs	960.7	3.3	940.0	3.7	875.0	**6.7**	897.0	3.6	877.3	2.8	883.0	3.0	905.5	3.0	925.2	4.8	885.8	1.1	-4.27
	IS-3A	Ratio 1z^+/2z	0.5	**5.2**	0.5	3.2	0.4	3.4	0.3	2.4	0.4	**6.9**	0.4	**6.2**	0.4	11.0	0.4	9.2	0.4	8.3	-14.23
	IS-3B	Ratio 3z^+/2z	1.0	0.5	1.0	**10.7**	0.8	4.1	1.2	2.1	1.1	**6.5**	1.1	**6.0**	1.0	15.1	0.9	13.9	1.1	4.1	21.61
	IS-3C	Ratio 4z^+/2z	0.5	**8.3**	0.6	3.4	0.4	**9.1**	0.7	4.0	0.7	**6.6**	0.7	**10.8**	0.6	22.1	0.5	15.7	0.7	1.9	42.88
	PL-1	AVG length z=1	7.9	0.9	7.7	1.9	7.8	1.7	7.5	1.3	7.6	1.6	7.5	0.8	7.7	1.6	7.8	0.8	7.5	0.6	-3.25
	PL-2	AVG length z=2	10.8	1.2	11.0	2.0	10.7	1.5	10.8	1.2	10.9	0.5	10.9	1.4	10.9	1.0	10.8	1.3	10.9	0.5	0.46
	PL-3	AVG length z=3	16.1	0.7	16.1	2.0	15.8	3.0	16.1	1.2	16.1	3.4	16.6	1.2	16.2	1.8	16.0	1.0	16.3	1.8	1.71
	PL-4	AVG length z=4	22.0	0.6	21.7	0.5	22.1	0.3	22.3	1.7	22.3	2.1	22.1	1.5	22.1	1.1	21.9	0.9	22.2	0.4	1.52

	Metric Code	*Description*	M1		M2		M3		M4		M5		M6		Inter Series		Porcine		Bovine		*Bov/Por*
			mean	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*mean*	*%dev*	*%change*
	DS2-B	Spectrum Counts	10026.0	0.9	10115.0	1.0	10077.7	1.2	10085.7	1.2	10143.0	0.7	10113.3	1.4	10093.5	0.3	10072.9	0.4	10114.0	0.3	0.41
	P-2C	Tryptic unique peptides	98.3	2.6	90.7	4.2	87.7	1.7	142.7	1.6	132.0	0.8	148.7	0.4	116.7	24.2	92.2	5.9	141.1	6.0	53.02
	P-2B	Tryptic unique ions	370.3	1.6	357.3	1.3	320.0	1.9	432.7	1.8	406.7	3.1	427.7	3.2	385.8	12.5	349.2	7.5	422.4	3.3	20.95
18 Hrs	P-2A	Tryptic total IDs	1432.0	**5.5**	1419.3	3.0	1253.0	5.0	1337.0	2.3	1349.7	3.9	1332.0	5.0	1353.8	4.4	1368.1	7.3	1339.6	0.7	-2.09
	IS-3A	Ratio 1z^+/2z	0.4	4.9	0.4	0.5	0.4	3.9	0.3	2.9	0.4	1.5	0.3	**7.1**	0.4	6.5	0.4	5.7	0.3	4.5	-8.32
	IS-3B	Ratio 3z^+/2z	0.8	0.7	0.8	4.9	0.7	3.9	1.2	2.8	1.0	3.7	1.1	3.3	0.9	22.7	0.8	12.5	1.1	6.7	43.73
	IS-3C	Ratio 4z^+/2z	0.3	4.9	0.4	3.4	0.3	4.3	0.7	**6.4**	0.5	3.2	0.7	**6.5**	0.5	35.9	0.3	13.0	0.6	11.7	91.68
	PL-1	AVG length z=1	7.7	0.8	7.8	1.2	8.0	0.5	7.5	1.4	7.6	1.5	7.6	0.5	7.7	2.6	7.8	2.0	7.6	0.7	-3.36
	PL-2	AVG length z=2	10.8	0.8	10.8	1.5	10.6	1.2	11.1	0.3	11.1	0.6	11.2	1.8	10.9	2.2	10.7	0.9	11.1	0.4	3.60
	PL-3	AVG length z=3	15.9	1.0	15.7	0.9	15.8	0.9	16.4	0.8	16.5	1.1	16.7	1.0	16.2	2.7	15.8	0.6	16.5	0.9	4.54
	PL-4	AVG length z=4	21.1	1.2	20.4	0.9	20.4	0.3	23.0	1.3	23.2	0.9	23.3	0.8	21.9	6.8	20.6	1.8	23.2	0.6	12.25

C. Profiling Differences in Cleavage Results

- The means of the relative intensities classified by digestion for either porcine or bovine trypsins were different (figure 2). More fully tryptic (FT) and semi tryptic(SC) serum albumin peptides were produced by the porcine trypsins whereas more peptides with a missed cleavage (MC) were produced when using the bovine trypsins.
- Albumin sequence position indicated that the complement of peptides produced were reproducible. The fraction of total abundance contributed by the FT, MC, or SC peptides is shown (Figure 3A).
- Analysis of select peptides indicated higher propensity for missed cleavage products by the bovine trypsins (Figure 3-B and C).

Figure 2.

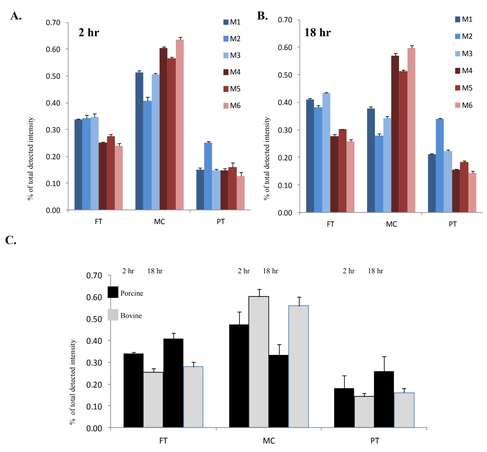
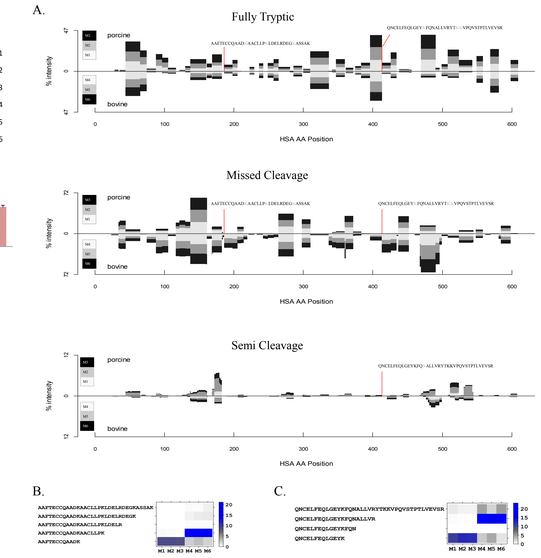


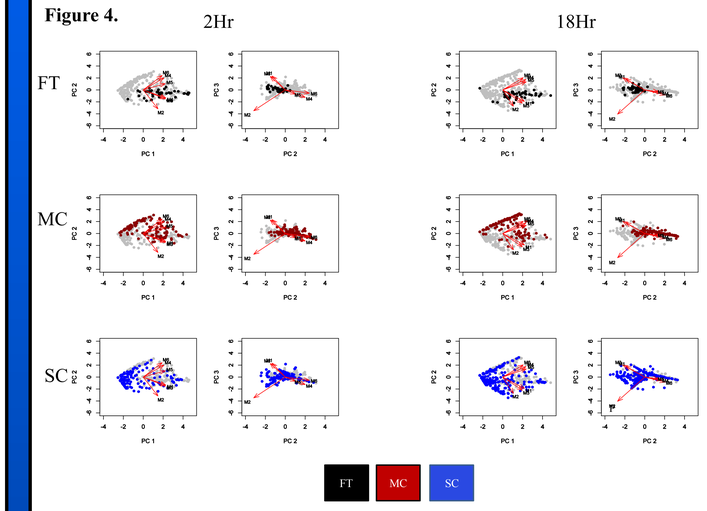
Figure 3.



D. Classification Of Trypsins by PCA

- PCA was completed using the detected intensities and visualized by cleavage rule: fully tryptic (FT) (black), missed cleavages (MC) (red) or partial tryptic-semi cleavage (SC) (blue) (**Figure 4**).
- Each grey dot represents the total intensity of an individual albumin peptide
- Arrows indicate the loading contributed to each trypsin.
- PC1 explains contributing peptide abundance toward each trypsin, PC2 explains the classification of the porcine versus bovine trypsins, and PC3 explains the differences between the individual trypsins.
- M1-M3: Porcine trypsins, M4-M6: bovine trypsins.

Figure 4.



E. Significant differences in peptide abundances

- Signature albumin peptides significantly altered between bovine and porcine trypsins (Table 3).
- Several parent albumin peptide sequences were identified with nested peptides that produced different abundances by the porcine and bovine trypsins.

TABLE 3 ***Signature peptides for each specific cleavage rule and significantly altered abundances by porcine and bovine trypsins.***

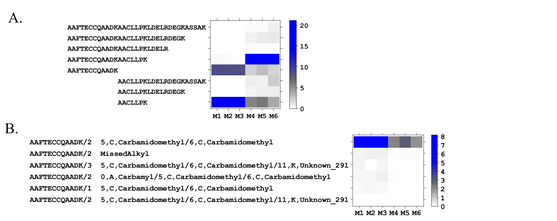
Peptide sequences were identified with a significant difference for the mean of the peptide abundances between the intraseries Porcine and Bovine trypsins. Relative intensities reported for each trypsin (M1-M6) and the mean ± %dev (CV) of the Porcine (M1-M3) or Bovine (M4-M6) trypsins. AA position is the relative amino acid sequence position of HSA. Significance reported required p < 0.01 and a fold change > 2. †,‡: reported peptides that are nested.

Full Tryptic Sequence	*AA position*	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*	*Porcine±CV*	*Bovine±CV*	*p**	*log FC (Bov/Por)*
SLHTLFGDK	88	20.39	25.28	22.26	1.31	2.11	0.25	22.65±0.11	1.23±0.76	0.0032	0.1
LCTVATLR	97	15.36	20.28	17.91	1.68	3.28	0.87	17.85±0.14	1.94±0.63	0.0049	0.1
ETYGEMADCCAK	105	4.15	5.03	4.07	2.88	3.47	2.37	4.42±0.12	2.91±0.19	0.0069	0.7
AAFTECCQAADK†	186	8.79	8.77	8.59	2.23	3.04	2.02	8.72±0.01	2.43±0.22	0.0020	0.3
AACLLPK	198	14.63	15.67	15.04	5.66	6.21	4.00	15.12±0.03	5.29±0.22	0.0040	0.4
NYAEAK†	341	1.62	1.99	1.86	0.78	1.07	0.78	1.82±0.10	0.88±0.19	0.0054	0.5
CCAAADPHECYAK	383	10.57	11.70	10.37	8.04	8.67	7.48	10.88±0.07	8.06±0.07	0.0028	0.7
QNCLLFEQLGEYK	413	10.75	12.36	12.18	2.67	4.35	1.44	11.77±0.08	2.82±0.52	0.0100	0.2
FQNALLVR	426	23.30	24.89	24.05	7.94	11.14	4.94	24.08±0.03	8.01±0.39	0.0096	0.3

Missed Tryptic Sequence	*AA position*	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*	*Porcine±CV*	*Bovine±CV*	*p**	*log FC (Bov/Por)*
DAHKSEVAHR	24	1.20	0.46	0.85	3.13	2.89	3.19	0.84±0.44	3.069±0.05	0.0072	6.5
LVRPEVDVMCTAFHDNEETFLKK	138	75.36	76.96	75.79	59.32	56.58	59.16	76.03±0.01	58.352±0.03	0.0082	25.0
AAFTECCQAADK/AACLLPK†	186	0.28	0.18	0.15	18.94	21.11	17.51	0.20±0.33	19.186±0.09	0.0032	4.0
NYAEAKDVFLGMFLYEYARRHPDYSVVLLLR‡	341	1.82	0.87	2.09	3.64	2.86	4.03	1.59±0.40	3.512±0.17	0.0029	0.3
RHPDYSVVLLLR‡	360	28.15	28.26	27.62	18.86	21.44	20.43	28.01±0.01	20.241±0.06	0.0015	4.1
KTPQVSTPTLVEVSR	437	31.97	33.60	33.00	29.32	30.46	29.66	32.86±0.03	29.813±0.02	0.0029	0.6

- The abundances of each nested sets of HSA peptides were altered dependent on trypsin sequence (Figure 5A) and by detection of multiple peptide ions for each peptide (Figure 5B).

Figure 5.



Conclusion

A comprehensive characterization of trypsin performance was completed using multiple trypsins and a human serum albumin spectra library. Sample and intra series replicates were highly reproducible, whereas differences emerged dependent on origin of trypsin and the cleavage rule. Several parent albumin peptide sequences were identified with sibling/daughter peptides that produced different abundances by the porcine or bovine trypsins. Further characterizing these differences in trypsin activity and selectivity may improve the reliability of peptide detection and quantitation for both SRM-based and complex sample biomarker discovery studies.