Analytical Methods

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Analysis of the whole serum proteome using an integrated 2D LC-MS/MS system

Mingming Xiao[a], Yajing Chen[b], Huan Yu[a], Su Wei[a], Kang Yu[a], Huan Zhao[c] , and Ruibing Chen[a,*]

[a]Research Center of Basic Medical Sciences, Tianjin Medical University, Tianjin, China

[b]School of Medical Laboratory, Tianjin Medical University, Tianjin, China

[c]Harmonia Biotechnology Ltd. Co., Tianjin, China

*Correspondence to:

Dr. Ruibing Chen, chenruibing@tijmu.edu.cn, Research Center of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China

Key words: whole serum, 2D LC-MS/MS, proteomics, serum proteome

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

An integrated two-dimensional liquid chromatography tandem mass spectrometry (2D LC-MS/MS) system was applied to investigate the proteome from the whole serum sample without abundant protein depletion. An integrated HPLC column containing both a strong-cation-exchange section and a C18 reverse-phase section was used for online 2D LC separation, which was coupled with a linear ion trap mass spectrometer. Over a thousand protein groups were identified from a single normal serum sample. This method enabled two-dimensional separation of tryptic peptides from complex biological samples using a conventional one-dimensional chromatography system. The serum sample from a patient with hepatocellular carcinoma (HCC) was also analyzed, and the gold standard HCC biomarker alpha-fetoprotein (AFP) was detected. Collectively, the integrated 2D LC-MS/MS system was proved to be an efficient tool to analyze the serum proteome.

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Mass spectrometry based proteomic approaches have evolved as powerful tools for the discovery of biomarkers.[1,2] However, the identification of potential protein biomarkers from biofluid samples, such as serum and plasma, remains challenging because of the large range of protein concentrations in those complex samples. Efforts have been made to simplify the serum samples based on affinity chromatography, either to remove the abundant proteins from the serum or to enrich a subproteome with a common chemical structural feature.[3-6] For example, affinity depletion using antibody conjugated materials have been extensively used for such purposes.[3,4] However, those methods risk the possibility of losing proteins with potential interest during the sample preparation process. In addition, the experimental materials required for these types of sample processing are rather expensive for routine clinical analysis. Therefore it is beneficial to develop a straightforward strategy to analyze the proteins from the whole serum.

Multiple-dimensional chromatography is a powerful tool to separate complex biological samples and to improve the dynamic range of mass spectrometric analysis.[7-10] The strong-cation-exchange/ reverse-phase HPLC (SCX-RP HPLC) is one of the most widely used setups.[7-9] The tryptic peptides are first fractionated by a SCX column followed by reverse-phase separation. However, the conventional SCX separation requires a salt gradient, which is not compatible with mass spectrometry. Therefore, offline separation and/or extensive column wash after sample loading is required before mass spectrometric analysis. In addition, for online SCX-RP HPLC separation, a specific HPLC system with two sets of gradient pumps is required. It has

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

been reported that the integrated SCX-RP column could be used on a regular HPLC system to acquire 2D separation. [11] Such method was proved to be efficient for the proteomics analysis of mouse liver tissues. However, it has not been tested for the analysis of biological fluids. In this study, a 2D LC-MS/MS system was built up using an integrated SCX-RP column and a LTQ linear ion trap mass spectrometer, and was employed to study the whole serum proteome.

The serum samples were collected with the patients' consent and the approval from the ethical committee. The acquired serum was spiked with cold acetone to precipitate the serum proteins. The protein samples were digested with trypsin overnight using a protocol described previsously.[12] The 2D LC system was set up as illustrated by Fig. 1. Briefly, a Biphase column was used for online 2D separation (SCX 0.075 mm×5 cm+C18 0.075 mm×10 cm) (Column Technology Inc., Fremont, CA) at a flow rate of 2 µL/min. The pH buffers were delivered using the sample loading mode, and 6 different pH buffers purchased from the Column Technology Inc. were used, including pH2.5, pH 3.5, pH 4.5, pH5.5, pH 6.0, and pH 8.0. The tryptic peptides were delivered onto the integrated column using the pH2.5 buffer, in which most of the peptides were positively charged and bound to the SCX section. When the pH increased above the pI values of certain peptides, the peptides became negatively charged and were eluted to the RP section. The peptides were then separated by the RP HPLC using a 45 min gradient of 5-45% buffer B (Mobile phase A, 0.1% formic acid in H_2O; Mobile phase B, 0.1% formic acid in acetonitrile) followed by detection with mass spectrometer.

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Peptides eluted from the integrated column were analyzed by a LTQ XL linear ion trap mass spectrometer (Thermo Electron Corp, Waltham, MA). The instrument was operated with a spray voltage of 3.5 kV. The information-dependent acquisition (IDA) mode of operation was employed in which a survey scan from m/z 400 to 1800 was acquired followed by collision-induced dissociation (CID). For MS/MS, normalized collision energy of 35% was used with an activation q of 0.25 for 30 ms. Ion selection thresholds for MS and MS/MS were 1,000 and 500 counts, respectively. Raw LTQ data was searched against the IPI human protein database using the SEQUEST algorithm embedded in the Protein Discoverer 1.2 Software (Thermo Electron Corp , Waltham, MA). Iodoacetamide derivative of Cys, deamidation of Asn and Gln, oxidation of Met were specified as variable modifications. Peptide identifications were accepted with > 95.0% probability as specified by the result filter (Xcorr> 1.9 if Charge=1, Xcorr> 2.2 if Charge=2, Xcorr> 3.75 if Charge=3). Protein identifications were accepted with at least two observed unique peptides and proteins identified with the same set of peptides were grouped.

A control serum was analyzed to evaluate the integrated 2D LC-MS/MS system. The total ion chromatograms of the pH buffer fractions were shown in Fig. 2A. Six different pH buffers were used for SCX separation in this study, while more pH gradient buffers could be used to further separate the peptides to achieve higher dynamic range. The abundant serum proteins, such as albumin, apolipoprotein and sub-types of immunoglobulins, were detected with the largest number of unique peptides and PSM values as expected, since the serum sample was not depleted as the

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

conventional protocol. Meanwhile, we were able to detect a large number of proteins with lower concentration using this method. Over 1000 protein groups were detected in each single run and 1793 total protein groups were detected combing two technical replicates (Supplemental Figure S1). As shown in Fig. 2B, majority of these proteins were identified with 2 unique peptides (65%). 4% of them were identified with more than 10 peptides. These were the most abundant proteins in the serum, such as albumin, apolipoprotein, haptoglobin, immunoglobulins, complement C, etc. As shown in Fig. 2C, 751 protein groups were observed in the two replicates of control serum, which accounts for 42% of the total proteins. The consistency of detection between runs was not ideal especially for proteins with lower abundance (PSM<4). It may be due to the type of mass spectrometer used in this study. Higher detection reliability could be achieved by coupling the integrated 2D LC system to a mass spectrometer with higher resolution and mass accuracy, such as LTQ-Orbitrap, FT-ICR, etc. And also, more replicates should be analyzed when dealing with real samples in order to improve identification confidence. To evaluate the dynamic range of the integrated 2DLC-MS/MS system, reported concentrations of 39 detected proteins were plotted in Fig. 3. The protein concentration information was acquired from the literature.[13-15] As shown in Fig. 3, the lowest concentration of identified proteins was around 10^3 pg/mL, and the dynamic range of detection is eight orders of magnitude.

A serum sample from a hepatocellular carcinoma (HCC) patient was also analyzed using the same method (Supplemental Fig. S2). This patient has been

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

recently diagnosed with HCC and hasn't received any treatment yet. Multiple proteins of interest that may be related with HCC were detected from this sample, such as the heat shock protein 70 (HSP70), transforming growth factor beta-1 (TGF-β1) and alpha-fetoprotein (AFP).[16,17] AFP is the only FDA approved HCC serum biomarker for clinical applications.[18] It was identified in this study with two tryptic peptides as shown in Fig. 4.

As shown above, our results demonstrate that the integrated 2D LC-MS/MS system is a useful tool to analyze a particular serum protein or to investigate the serum proteome. The sensitivity of this method is comparable to other widely used serum proteomics analysis methods. It can be used for analyzing whole serum without any affinity depletion, which reduced the cost of serum proteomics and also eliminate the possibility of losing certain proteins of interest during sample processing. Such method can be easily adapted to any type of commercial liquid chromatography system. For sample preparation, the whole serum sample sometimes is sticky and contains lipids and other impurities that may block the column. Filtration with 0.22 μm filters and/or deplipidation before trysin digestion is usually helpful to prolong the lifetime of the column. In addition, we used a LTQ mass spectrometer for detection, which is not suitable for isotopic labeling based quantitative proteomics analysis. For biomarker discovery, label free quantification can be used to investigate differential proteins for potential biomarker identification. Furthermore, it can be coupled with mass spectrometers with higher mass accuracy and resolution to achieve better performance.

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Conclusions

In this work, we applied an integrated two dimensional liquid chromatography tandem mass spectrometry (2D LC-MS/MS) system to investigate the proteome from the whole serum sample without affinity depletion. Over a thousand protein groups were identified in a control serum sample using this method. We were also able to detect the gold standard hepatocellular carcinoma (HCC) biomarker alpha-fetoprotein (AFP) from a HCC serum sample. Such method enables multiple-dimensional separation of tryptic peptides from complex biological samples using a typical one dimensional chromatography system and provides an efficient and straightforward way to analyze the serum proteome.

Analytical Methods Accepted Manuscript

Acknowledgements

This study was supported by grants from National Natural Science Foundation of China (21205088), and Doctoral Research Fund from the Ministry of Education of China (20121202120001). This work was also sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

1. J. M. Jacobs, J. N. Adkins, W. Qian, T. Liu, Y. Shen, D. G. Camp II and R. D. Smith. *J. Proteom. Res.* 2005, **4**, 1073-1085.
2. C. E. Parker and C. H. Borchers. *Mol. Oncol.* 2014 pii: S1574-7891(14)00054-4.
3. T. Plavina, E. Wakshull, W. S. Hancock and M. Hincapie. *J. Proteom. Res.* 2007, **6**, 662-671.
4. A. Kovàcs and A. Guttman. *Curr. Med. Chem.* 2013, **20**, 483-490.
5. X. Wei, A. Herbst, D. Ma, J. Aiken and L. Li. *J. Proteome. Res.* 2011, **3**, 2687-2702.
6. Z. Dai, J. Zhou, S. Qiu, Y. Liu and J. Fan, *Electrophoresis.* 2009, **30**, 2957–2966.
7. M. P. Washburn, D. Wolters and J. R. Yates 3rd. *Nat. Biotechnol.* 2001, **19**, 242-247.
8. N. Wang, C. Xie, J. B. Young and L. Li. *Anal. Chem.* 2009, **81**, 1049-1060.
9. J. Zhang, K. A. Lanham, R. E. Peterson, W. Heideman and L. Li. *J. Sep. Sci.* 2010, **33**, 1462-1471.
10. J. A. Dowell, W. V. Heyden and L. Li. *J. Proteome. Res.* 2006, **5**, 3368-3375.
11. J. Dai, C. H. Shieh, Q. Sheng, H. Zhou and R. Zeng. *Anal. Chem.* 2005, **77**, 5793-5799.
12. X. Zhang, F. Zhang, L. Guo, Y. Wang, P. Zhang, R. Wang, N. Zhang and R. Chen. *Mol. Cell Proteomics.* 2013, **12**, 3199-3209.
13. L. Anderson. *J. Physiol*. 2004, **563**, 23-60.
14. P. Mitchell. *Nature Biotech.* 2010, **28**, 665-670.

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

15. L. N. Anderson and N. G. Anderson. *Mol Cell Proteomics*, 2002, **1**, 845-867

16. L. Zhou, J. Liu and F. Luo. *World J. Gastroenterol.* 2006, **12**, 1175-1181.

17. Y. Zhao, Q. Ju and G. Li. *Mol. Clin. Oncol.* 2013, **1**, 593–598.

18. P. J. Johnson. *Clin. Liver Dis*. 2001, **5**, 145-159.

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure Legends

Figure 1. Illustration of the integrated 2D LC-MS/MS system. A Biphase column was used for online 2D LC separation. pH buffers were delivered using the sample loading mode. Peptides eluted from the SCX dimension were separated with the RP section using acetonitrile gradient.

Figure 2. The proteomics analysis of the whole serum sample. (A) Total ion chromatograms of the 2D LC-MS/MS analysis of the human serum using 6 pH buffers. (B) Number of the unique peptides identified for each protein group. (C) The comparison of detected protein groups between two replicate analyses of a whole control serum samples.

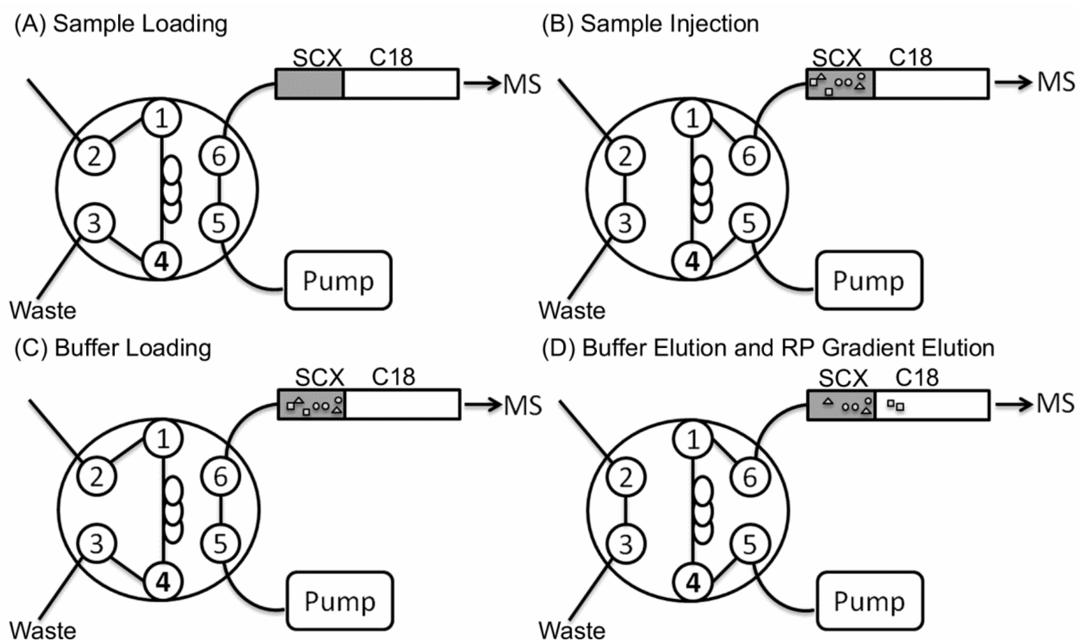
Figure 3. The dynamic range of detected serum proteins. Values were taken from the literature.

Figure 4. MS/MS spectra of alpha-fetoprotein (AFP) identified from the HCC serum using the integrated 2D LC-MS/MS system. (A) Mass spectrum from tryptic peptide GDVLDCLQDGEK. (B) Mass spectrum from tryptic peptide CFQTENPLECQDKGEEELQK. The iodoacetamide derivative of Cys was indicated with small capital c.

Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

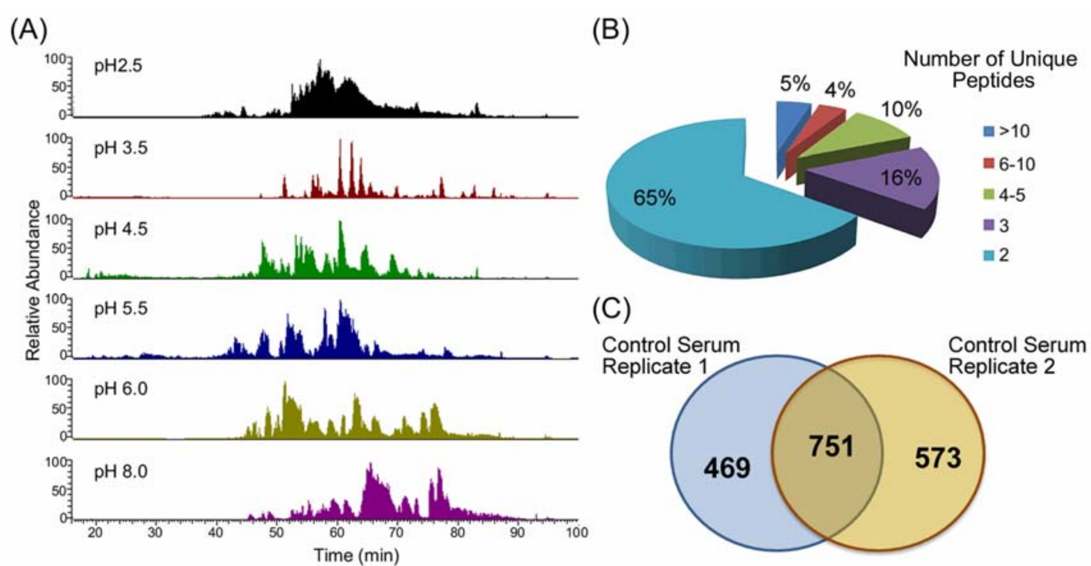
Figure 1.



Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

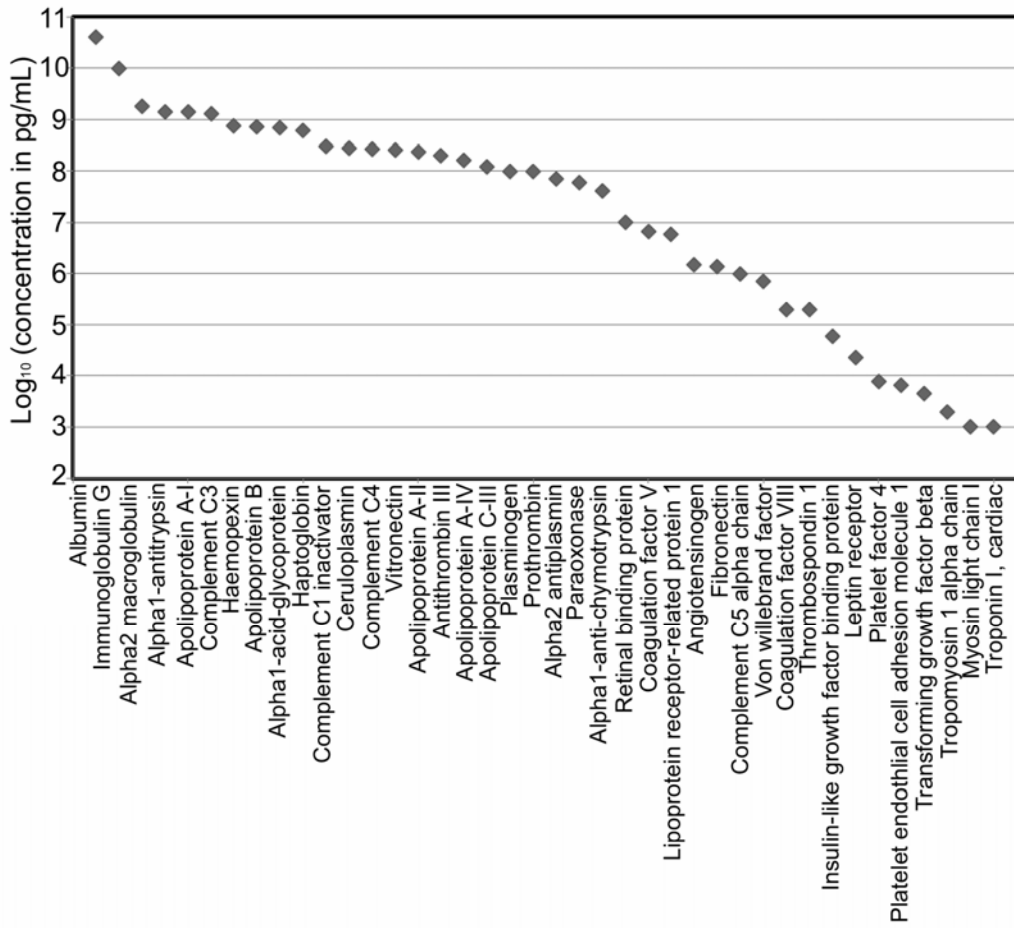
Figure 2.



Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

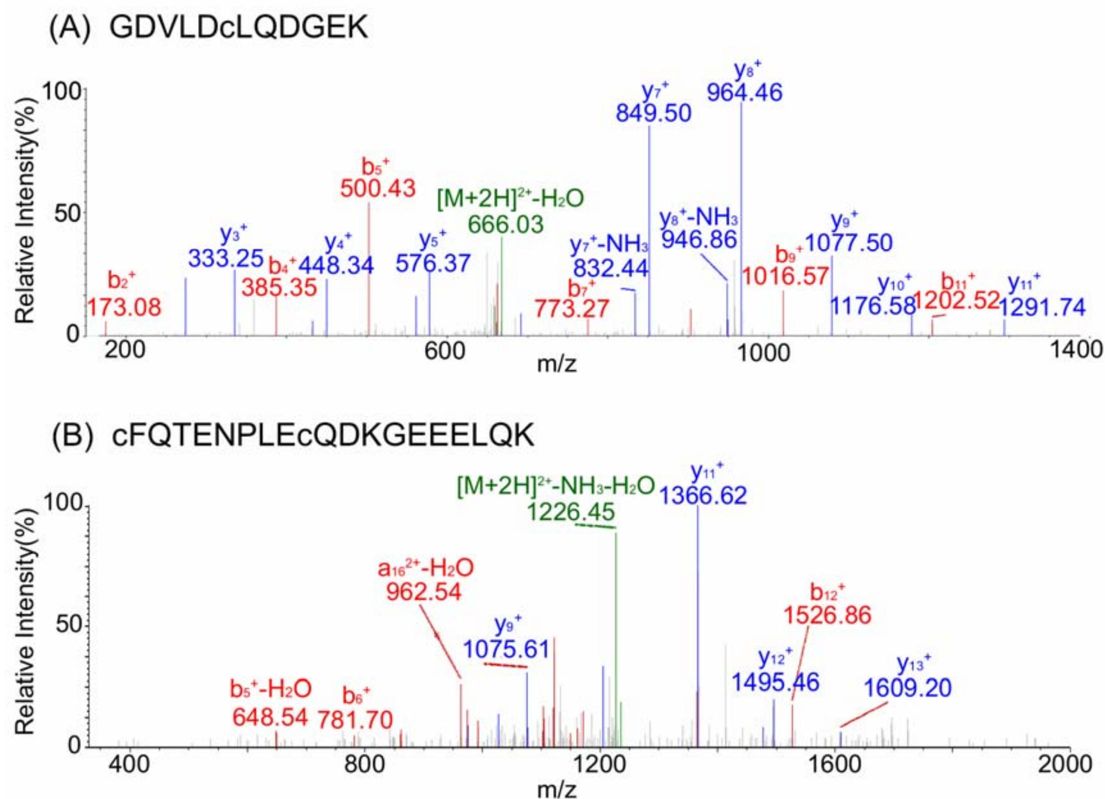
Figure 3.



Analytical Methods Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 4.



Analytical Methods Accepted Manuscript