Dried blood spot proteomics: Automated surface sampling and sample preparation

Nicholas J. Martin [1], Josephine Bunch [2], Helen J. Cooper [1]

1 School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
2 School of Chemistry, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

Overview

- Analysis of dried blood spots (DBS) by mass spectrometry has, to date, focused on small molecules and haemoglobin.
- DBS are a potentially rich source of biomarkers.
- Here, we demonstrate automated, untargeted proteomic analysis of endogenous proteins obtained via surface sampling of DBS.

Introduction

- Liquid extraction surface analysis (LESA) by use of the Triversa Nanomate robotic nanoelectrospray platform (TVNM) (Advion Biosciences) has previously been applied to the analysis of intact haemoglobin proteins from DBS.
- Here we use LESA and the TVNM to perform surface extraction of endogeneous proteins from DBS followed by an automated trypsin digestion for LC MS/MS analysis.
- LC MS/MS analysis of the resulting digests results in the identification of >100 proteins, crossing 4 orders of magnitude of concentration in blood plasma, several of which are biomarker candidates for screening programmes and other clinical assays.

Methods

- Capillary blood was taken from healthy donors via finger prick and applied to NHS dried blood spot cards Ahlstrom grade 226 filter paper (ID Biological systems) and dried overnight prior to surface sampling and digestion.
- Trypsin digestion was performed by a robotic sequence using the Advanced User Interface (AUI) function of TVNM as shown in fig 1.

A. DBS was loaded onto microtitre plate and heated to 40˚C by the temperature control unit of the TVNM. One well of the microtitre plate was filled with 50 mmol NH_4HCO_3 and another was filled with 0.1 µg/µl Trypsin Gold (Promega).
B. 7 µl solvent was aspirated from the solvent well.
C. 6 µl solvent was dispensed onto DBS, forming a liquid microjunction between tip and surface of DBS, allowing intact proteins to diffuse from the DBS into the solvent.
D. Solution was reaspirated and dispensed into a clean well in the microtitre plate.
E. 4.5 µl trypsin solution was aspirated from the trypsin well.
F. Trypsin solution was added to sample.
G. Sample was left to digest for 1 hour
H,I. As solvent begins to evaporate from sample well, additional solvent (7.5 µl) is aspirated from solvent well and added to sample well. (H&I are performed at 30 mins and 1 hr).
J. Proteins are digested into peptides after 1 hour.
K. Plate is transferred to HPLC autosampler and peptides are analysed by LC MS/MS

- Samples were analysed by nanoflow rate LC MS/MS using a Dionex Ultimate 3000 nano LC unit (Thermo Fisher Scientific) and Orbitrap Velos ETD mass spectrometer (Thermo Fisher Scientific).
- Peptides were separated by a 3.2-44% ACN gradient and fragmented by a 'top 7 CID' method, in which a survey scan was followed by CID fragmentation, with a normalised collision energy of 35%, of the seven most abundant precursor ions.
- MS/MS data were searched against the SwissProt human database (downloaded in December 2012), composed of 20233 sequences, using a Mascot and Sequest algorithm in Proteome Discoverer 1.3.
- Parameters were: precursor ion mass accuracy 10 ppm, fragment mass tolerance 0.8 Da, methionine oxidation was allowed as a dynamic modification and up to 2 missed cleavages were permitted in the digestion.
- Peptide false discovery rates were calculated from a decoy database using the percolator function of Proteome Discoverer. Data were filtered to a false discovery rate of 1%.
- The protein grouping algorithm was applied which grouped all non-unique peptides to the highest scoring protein.
- All proteins identified by only one unique peptide were confirmed manually.

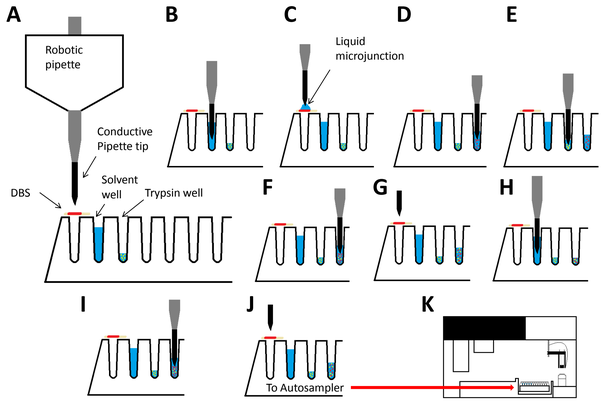

Fig 1 Sequence of robotic aspiration and dispense cycles used in surface sampling and digestion protocol.

Results

	Protein Description	Total Coverage	Unique peptides	Concentration in plasma / µmol/L [1]
1	Hemoglobin subunit beta	99.32%	16	
2	Hemoglobin subunit delta	75.51%	5	
3	Hemoglobin subunit gamma-1	74.15%	7	
4	Hemoglobin subunit alpha	71.83%	15	
5	Protein S100-A9	55.26%	4	0.05-0.14
6	Apolipoprotein A-I	54.68%	18	30-70
7	Glyceraldehyde-3-phosphate dehydrogenase	50.15%	11	
8	Ig kappa chain C region	49.06%	3	68-150
9	Carbonic anhydrase 1	44.06%	8	
10	Carbonic anhydrase 2	43.46%	8	
11	Peroxiredoxin-6	40.18%	6	
12	Protein S100-A6	38.89%	3	
13	Keratin, type I cytoskeletal 10	37.67%	17	
14	Complement C3	35.18%	38	5 – 10
15	Serum paraoxonase/arylesterase 1	34.93%	6	
16	Keratin, type II cytoskeletal 1	34.78%	18	
17	Apolipoprotein A-II	33.00%	3	30 – 60
18	Serum albumin	31.86%	29	500 – 800
19	Ig gamma-1 chain C region	31.82%	5	68-150
20	Flavin reductase	31.07%	4	
21	Keratin, type I cytoskeletal 9	30.50%	11	
22	Keratin, type II cytoskeletal 2 epidermal	28.79%	9	
23	Alpha-1-antitrypsin	28.71%	11	18-40
24	Keratin, type I cytoskeletal 13	28.17%	9	
25	Apolipoprotein C-III	27.27%	2	6-20
26	Dermcidin	25.45%	3	
27	Protein S100-A8	24.73%	2	0.05-0.14
28	Inter-alpha-trypsin inhibitor heavy chain H4	23.66%	11	2-4
29	Bisphosphoglycerate mutase	22.78%	3	
30	Peroxiredoxin-2	22.73%	5	
31	Actin, cytoplasmic 1	22.67%	5	
32	Histone H2A type 1-H	22.66%	1	
33	Inter-alpha-trypsin inhibitor heavy chain H2	21.25%	10	2-5
34	Ceruloplasmin	20.94%	13	2-5
35	Apolipoprotein C-II	20.79%	1	2-7
36	Vitamin D-binding protein	20.04%	7	
37	Nucleoside diphosphate kinase A	19.08%	2	
38	Serotransferrin	17.62%	10	25 – 45
39	Complement factor B	16.88%	10	
40	Ig kappa chain V-III region SIE	16.51%	1	68-150
41	Alpha-1-acid glycoprotein 1	16.42%	3	9-20
42	Keratin, type II cytoskeletal 2 oral	16.30%	6	
43	GTP-binding nuclear protein Ran	16.20%	3	
44	Hemopexin	15.58%	6	9-20
45	Catalase	15.37%	6	
46	Apolipoprotein A-IV	15.15%	4	3-6
47	Clusterin N	14.70%	5	1-2
48	Alpha-2-macroglobulin	14.65%	14	7-17
49	Fibrinogen beta chain	14.26%	6	10-27
50	Lysozyme C	14.19%	2	0.01-1
51	Ig lambda-7 chain C region	14.15%	1	68-150
52	Peroxiredoxin-1	14.07%	3	
53	Fibrinogen gamma chain	13.69%	3	9-24
54	Semenogelin-1	13.64%	4	
55	Peptidyl-prolyl cis-trans isomerase A	13.33%	2	
56	Apolipoprotein C-I	13.25%	1	6-12
57	Purine nucleoside phosphorylase	13.15%	3	
58	Keratin, type I cytoskeletal 14	12.92%	2	
59	Inter-alpha-trypsin inhibitor heavy chain H1	12.84%	7	2-4
60	Fructose-bisphosphate aldolase A	12.64%	3	

	Protein Description	Total Coverage	Unique peptides	Concentration in plasma / µmol/L [1]
61	Ras-related protein Rab-14	12.56%	1	
62	Glutathione S-transferase A1	12.16%	1	
63	Ig alpha-1 chain C region	11.90%	3	8-50
64	Polyubiquitin-C	11.82%	1	
65	Serum amyloid A-4 protein	11.54%	1	4
66	Alpha-synuclein	11.43%	1	
67	Thioredoxin	11.43%	1	
68	Cofilin-1	10.24%	1	
69	Protein DJ-1	10.05%	1	
70	Mucin-like protein 1	10.00%	1	
71	Complement C4-B	9.81%	9	0.5-2
72	Ig gamma-3 chain C region	9.55%	1	2-16
73	Alpha-1B-glycoprotein	9.49%	4	3-5
74	Vitronectin	9.41%	3	1-5
75	Gelsolin	9.21%	4	3-5
76	Haptoglobin	9.11%	3	0-40
77	Keratin, type II cytoskeletal 5	8.98%	2	
78	Lactotransferrin	8.59%	3	
79	Semenogelin-2	8.42%	4	
80	Ig kappa chain V-I region Lay	8.33%	1	68-150
81	Trypsin-1	8.10%	1	
82	Angiotensinogen	8.04%	3	1
83	Keratin, type II cytoskeletal 6A	7.98%	1	
84	Heat shock protein beta-1	7.80%	1	
85	Prothrombin	7.56%	3	1-5
86	Histidine-rich glycoprotein	7.43%	3	1-3
87	Apolipoprotein E	7.26%	1	0.6-2
88	Ig lambda chain V-III region LOI	7.21%	1	36-48
89	Alpha-enolase	7.14%	2	
90	Zinc-alpha-2-glycoprotein	7.05%	1	0.8-1.6
91	Alpha-1-antichymotrypsin	6.86%	2	4-9
92	Apolipoprotein-L1	6.78%	1	0.2
93	Kininogen-1	6.68%	4	3
94	Adenylate kinase isoenzyme 1	6.19%	1	
95	Fibrinogen alpha chain	6.00%	3	10-27
96	Delta-aminolevulinic acid dehydratase	5.76%	1	
97	Alpha-2-HS-glycoprotein	5.45%	1	9-30
98	Plasma protease C1 inhibitor	5.40%	2	
99	Ig gamma-2 chain C region	4.91%	1	20 – 90
100	Arginase-1	4.66%	1	
101	Zymogen granule protein 16 homolog B	4.33%	1	
102	Complement factor H	4.31%	3	2-5
103	Rab GDP dissociation inhibitor beta	4.27%	1	
104	Heat shock-related 70 kDa protein 2	4.23%	1	
105	14-3-3 protein theta	4.08%	1	
106	Plasminogen	3.95%	2	2-6
107	Fibroblast growth factor 2	3.82%	2	
108	Liver carboxylesterase 1	3.70%	1	
109	Heparin cofactor 2	3.61%	2	1.5
110	Serpin B3	3.59%	1	
111	Protein disulfide-isomerase A2	3.43%	1	
112	Zinc finger protein 611	3.40%	1	
113	Apolipoprotein B-100	3.31%	11	1-3
114	Phosphoglycerate kinase 1	2.64%	1	
115	Retinal dehydrogenase 1	2.59%	1	
116	Lumican	2.37%	1	
117	Complement C1s subcomponent	2.33%	1	1
118	Plasma kallikrein	2.19%	1	
119	Ig mu heavy chain disease protein	2.05%	1	4-25
120	Complement component C9	1.97%	1	0.4-1

Table 1 Proteins identified from LC MS/MS run of trypsin digestion of endogenous proteins sampled from surface of DBS

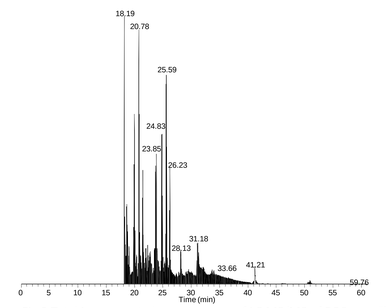

Fig 2 Total ion current chromatogram of LC MS/MS of DBS digest.

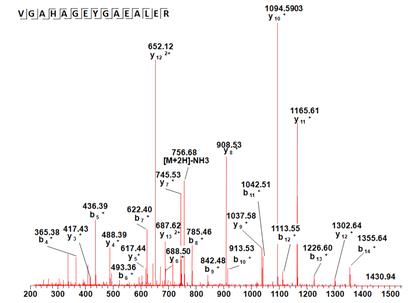

Fig 3 CID spectrum of haemoglobin alpha peptide [VGAHAGEYGAEALER] $^{2+}$.

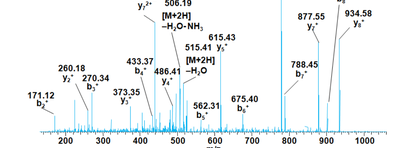

Fig 4 CID spectrum of inter alpha trypsin inhibitor H4 peptide [LGVYELLLK+2H]$^{2+}$.

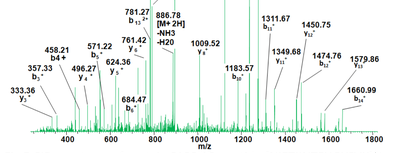

Fig 5 CID spectrum of protein S100 A9 peptide [NIETIINTFHQYSVK]$^{2+}$.

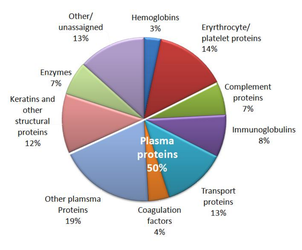

Fig 6 Categories of proteins identified from the DBS.

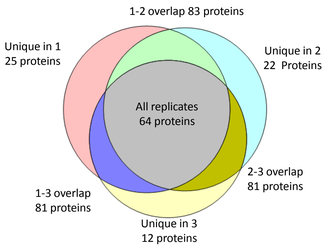

Fig 7 Venn diagram showing overlap of proteins across 3 replicates.

Conclusions and future work

- Over 100 proteins can be identified per DBS. Proteins identified cross a concentration range of 4 orders of magnitude in human plasma [1].
- A number of the proteins identified are biomarkers of diseases e.g. haemoglobin, cerruloplasmin and alpha-1-antitrypsin.
- The digestion procedure is fully automated and can be used to digest multiple spots in parallel.
- Minimal amounts of trypsin and other reagents are required. Only 0.45µg of trypsin is required per digestion.

References

[1] Glen L. Hortin, Denis Sviridov and N. Leigh Anderson High abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. *Clinical Chemistry* **54** 1608-1616 (2008)

Acknowledgments: the EPSRC, Advion Biosciences, the Biochemical Society and the British Mass Spectrometry Society are greatly appreciated for financial support