The use of Expression Profiling for the Detection of C-Reactive Protein in Human Serum

T. McKenna, C.J. Hughes, M. A. Ritchie , I. Campuzano and J. Langridge
Waters Corporation, MS Technologies Centre, Manchester, United Kingdom

Introduction

A common goal of proteomic experiments is the qualitative identification of proteins from biological samples. However, recently a great deal of interest has been generated on the use of HPLC in combination with mass spectrometry for the quantitative study of gene expression at the proteome level. The study of protein expression levels between samples has become increasingly important in modern biology and medicine.

Several approaches, which utilise stable isotope labelling of samples in combination with separation and subsequent analysis by mass spectrometry, for the study of protein expression levels, have been described [1,2]. In addition a previous study used an oa-Tof mass spectrometer in a parallel manner to provide enhanced protein identification, with relative protein quantification using isotope labelling [3].

We have recently presented a concept for qualitative and quantitative protein profiling using LC-MS, without the use of isotope labelling [4]. This methodology allows the relative change in abundance of individual peptides, and proteins, in highly complex mixtures to be determined. Excellent chromatographic performance, in terms of retention time reproducibility, and elevated resolution and mass measurement accuracy from the mass spectrometer provide additional selectivity and specificity, enabling these studies.

Using replicate injections on a nanoscale LC-MS system allows for the identification and correlation of ions across identically prepared control and experimental samples. We have shown that using this configuration, we can determine the change in relative abundance of a small number of ions, between two conditions, solely by accurate mass and retention time.

In the work presented here we have used this analytical approach to monitor C-reactive protein (CRP) in normal Human serum and in serum containing an elevated level of CRP

C- reactive protein (CRP).

C-reactive protein is part of the pentraxin family and is characterised by discoid assembly and Ca^{2+} dependant ligand binding. It has long been recognised as a non-specific acute phase protein produced by the liver in response to tissue injury, infection, and inflammation. It is found in low levels in normal serum, with the normal range for CRP in serum or plasma typically 0.2 to 5 mg/L. However, in response to an acute event, circulating levels may increase up to 100 fold. [5]. The protein was first discovered over 70 years ago but is currently receiving attention as evidence increases suggesting that elevated levels may be indicative of a number of diseases. Recently a direct role has been proposed for CRP as a biomarker for diagnosing and reflecting the course of metastasis in stage IV melanoma [5]. CRP is also a proposed marker for cardiovascular disease and may be an obesity related inflammatory maker [6].

Methods

Sample depletion
Normal control serum and elevated CRP serum were protein depleted of the six most abundant proteins to extend the dynamic range of the experiment. The samples were reduced, alkylated and digested prior to analysis using the Waters Protein Expression system.

An aliquot of both sera samples was depleted using the multiple affinity removal system, MARS kit (Agilent) designed to remove the six most abundant proteins present in Human serum; Albumin, IgG, anti-trypsin, IgA, transferrin and haptoglobulin. The procedure used is detailed below.

- 40µL of serum was diluted to 200µL prior to injection onto the depletion column using the Waters Alliance Bioseparations system.
- The flow through fraction was collected in 2mL at a flow rate of 500µL/min and concentrated to 300µL using a 5kDa spin cartridge.

The sample was diluted to 2mL with 100mM NH4HCO3 and concentrated again to 300µL.

Sample digestion
A 50µL aliquot of each depleted serum sample was reduced with 10mM dithiothretal (DTT), alkylated with 55mM iodoacetamide (IAA) and digested in the presence of Rapigest SF using sequencing grade trypsin (promega, Madison, WI) to a total volume of 70µL. The resulting tryptic peptide solution was diluted as follows, prior to analysis by LC-MS

- CRP elevated serum- 100ug/mL was diluted 1:5 and 5uL injected.
- Normal serum was diluted 1:1 and 5uL injected
- Tryptic digests of the proteins Enolase and Alcohol Dehydrogenase (ADH) were spiked into both serum samples as internal standards

Nanoscale LC
- Samples were injected onto the Stream Select module of a Modular CapLC system and analysed as previously described [4]

Mass Spectrometry
- Data was acquired using alternating low collision energy (approx 8 - 10eV) and elevated collision energy (23 – 33eV).
- An integration time of 1.5 seconds was used for each function and the reference channel was sampled every 30 seconds.
- A resolution of approximately 17,500 FWHM was used
- Nanolockspray was enabled to continually infuse Glu-Fibrinopeptide B (accurate mass of doubly charged species is 785.8426 amu).

Bioinformatics
- All replicate data files were processed using the Waters Protein Expression System Bioinformatics
- Interrogation of the acquired raw data, via an automated algorithm [4], was used to extract all of the relevant peptide level information, including the de-isotoped monoisotopic exact mass, HPLC retention time and peak intensity/ area for each detectable peptide. This is known as an Exact Mass Retention Time signature (EMRT)
- Subsequent comparison of these EMRT signatures across injections (replicates) and conditions (different samples) was performed automatically, and this produced a list of the EMRT "clusters". The algorithm used to match the EMRT's can take account of variability in retention time reproducibility, and mass measurement accuracy from experiment to experiment.
- The low energy data includes the exact mass of all detectable peptide molecular ions, whilst the parallel fragmentation data from the high energy scans provides information about the amino acid sequence for the peptide.

Results and Discussion

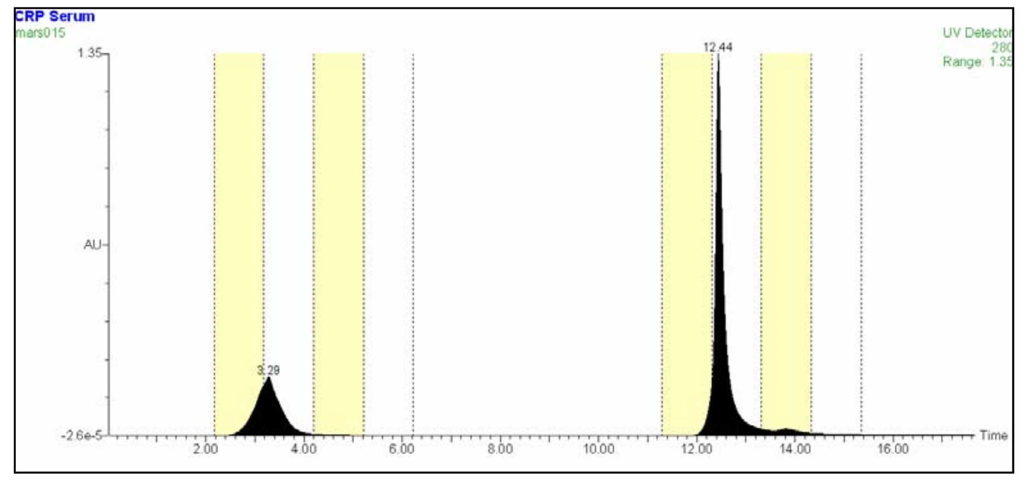

Figure 1. A typical LC-UV chromatogram obtained from the depletion of Human serum is presented (UV=280nm). This clearly shows the un-retained, low abundance proteins eluting at 3.3mins whilst the retained, high abundance proteins, elute at 12.44 mins.

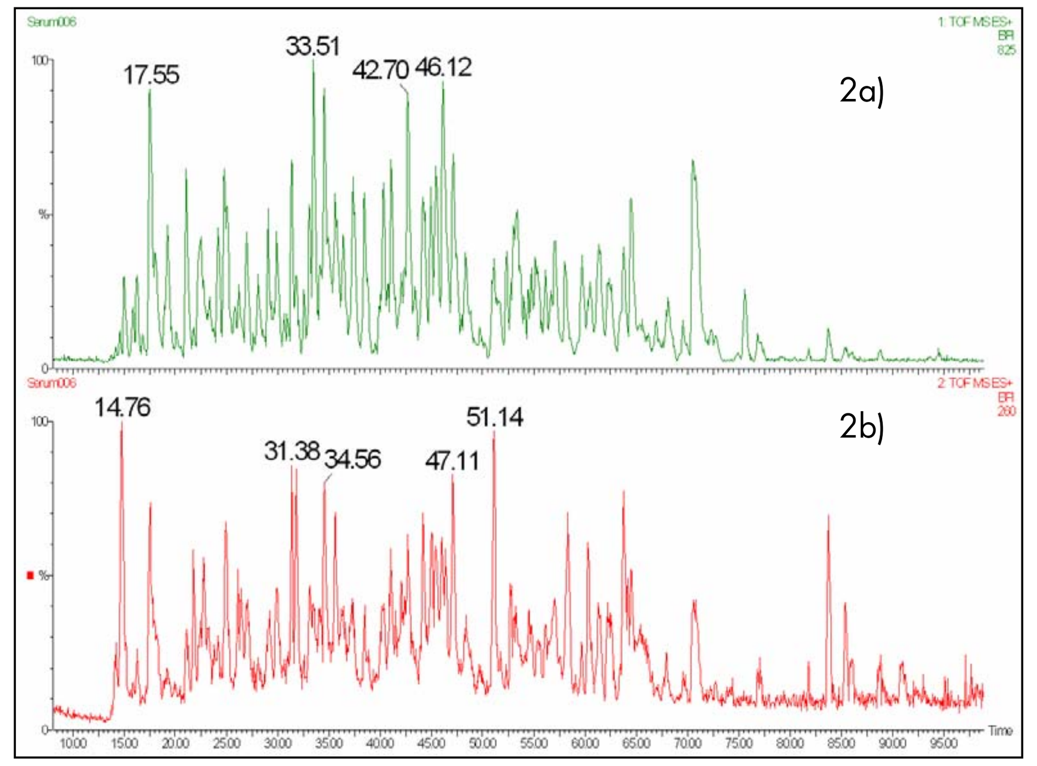

Figure 2. Nanoscale LC-MS analysis of a tryptic digest of depleted CRP elevated Human serum. The chromatograms represent the low energy (2a) and elevated energy (2b) base peak intensity, BPI, plots

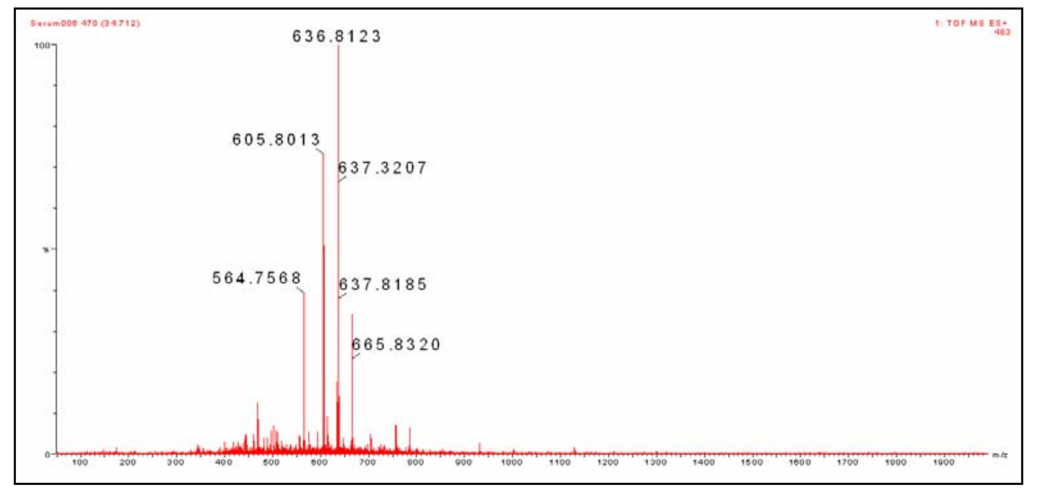

Figure 3. A mass spectrum obtained from the low energy function at a time of 34.7 minutes. It can clearly be seen that a number of ions, many doubly charged are present in the source at one time.

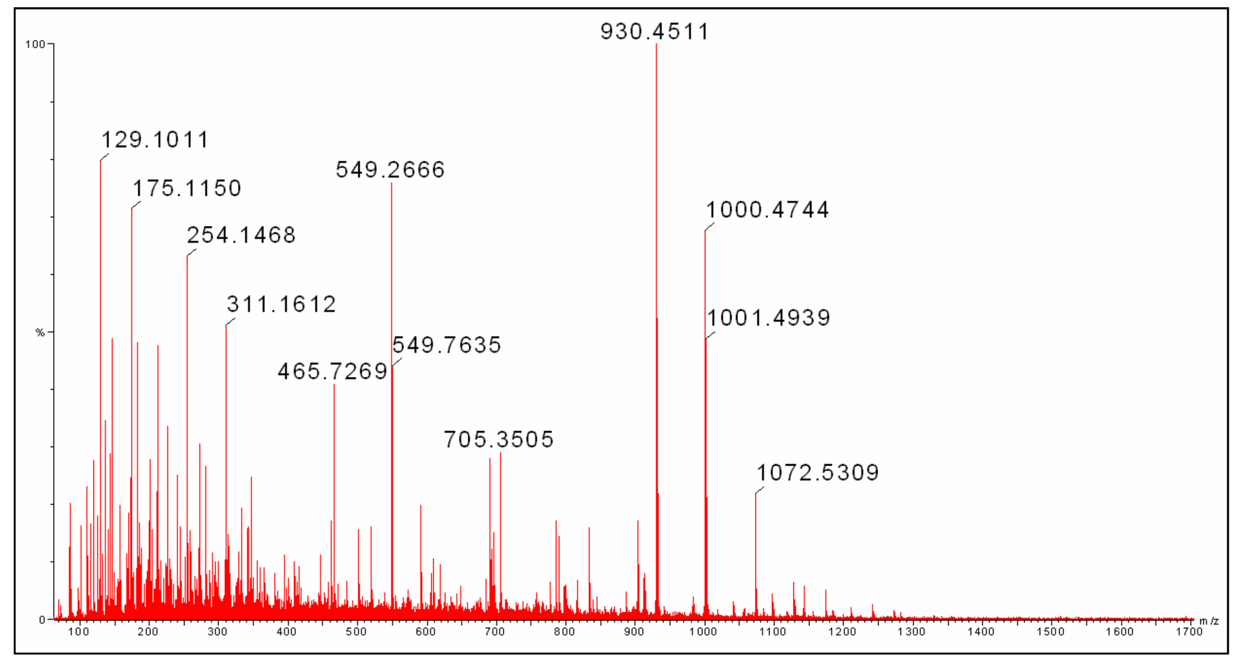

Figure 4 Examination of the corresponding elevated energy spectrum, gives an indication of the complexity of the fragment ion data obtained.

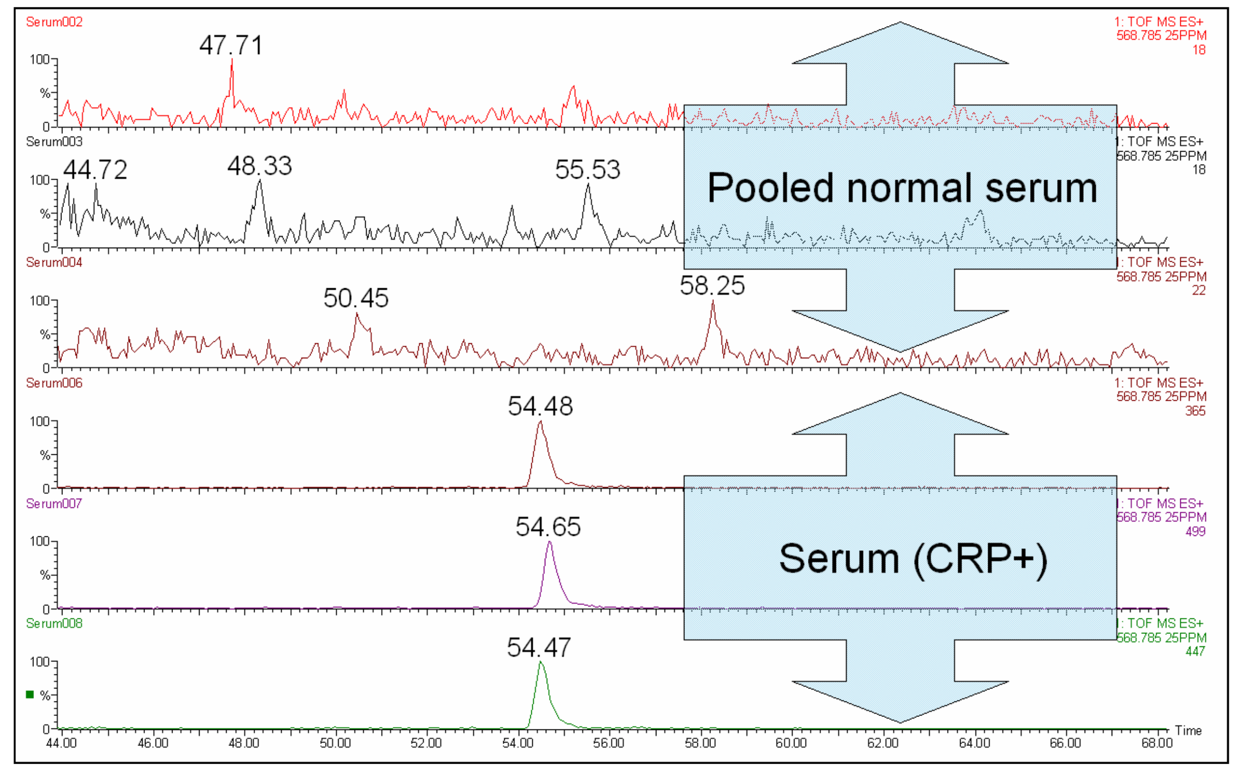

Figure 5 Results from the analysis of the depleted normal serum (top) and the depleted CRP elevated serum (bottom). In the case of the CRP elevated serum, m/z 568.785 (2+), the tryptic peptide GYSIFSYATK from CRP, is apparent in all three experiments with good signal to noise. In this case, the level of CRP in the serum was severely elevated, to a level of 100ug/mL.
Our preliminary experiments show that CRP can be identified in the range 1-100ug/mL in Human serum. Endogenous levels of CRP can be identified in the control serum with several tryptic peptides detected.

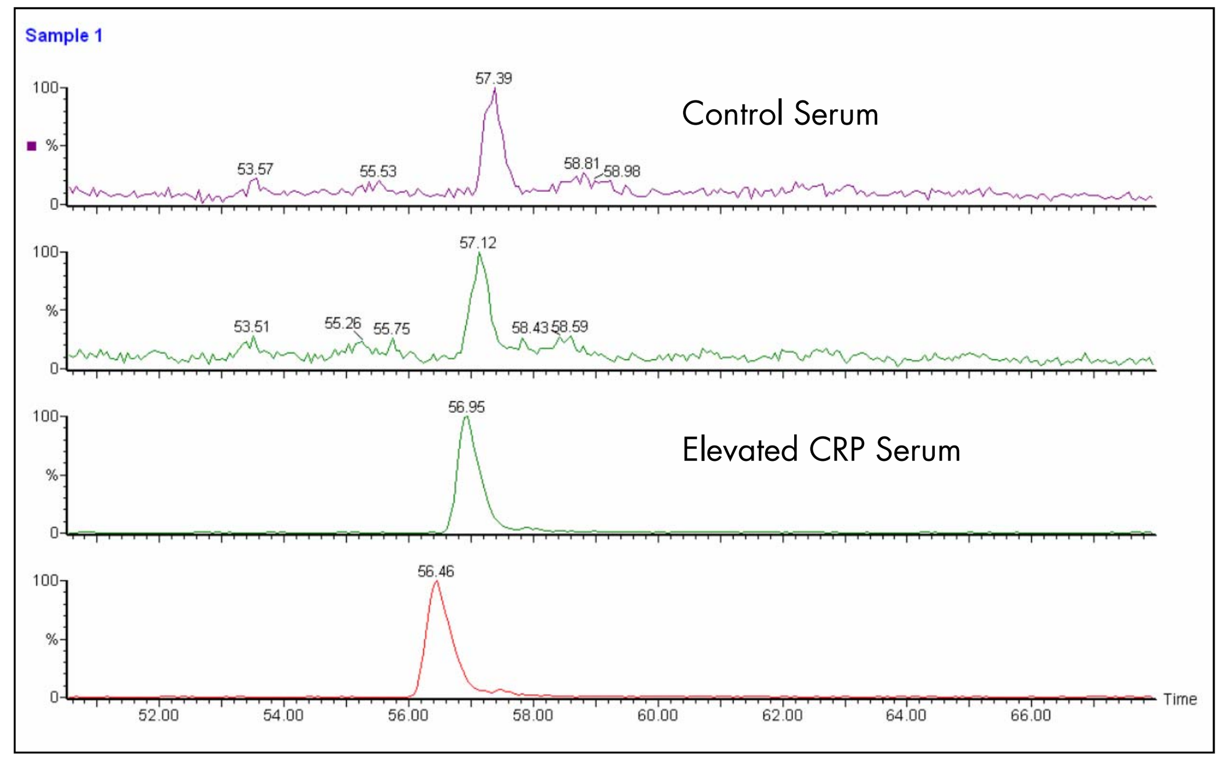

Figure 6 Mass chromatograms from a tryptic peptide from CRP at a retention time of 55 minutes, from duplicate injections of the control serum digest, and from the CRP elevated serum.

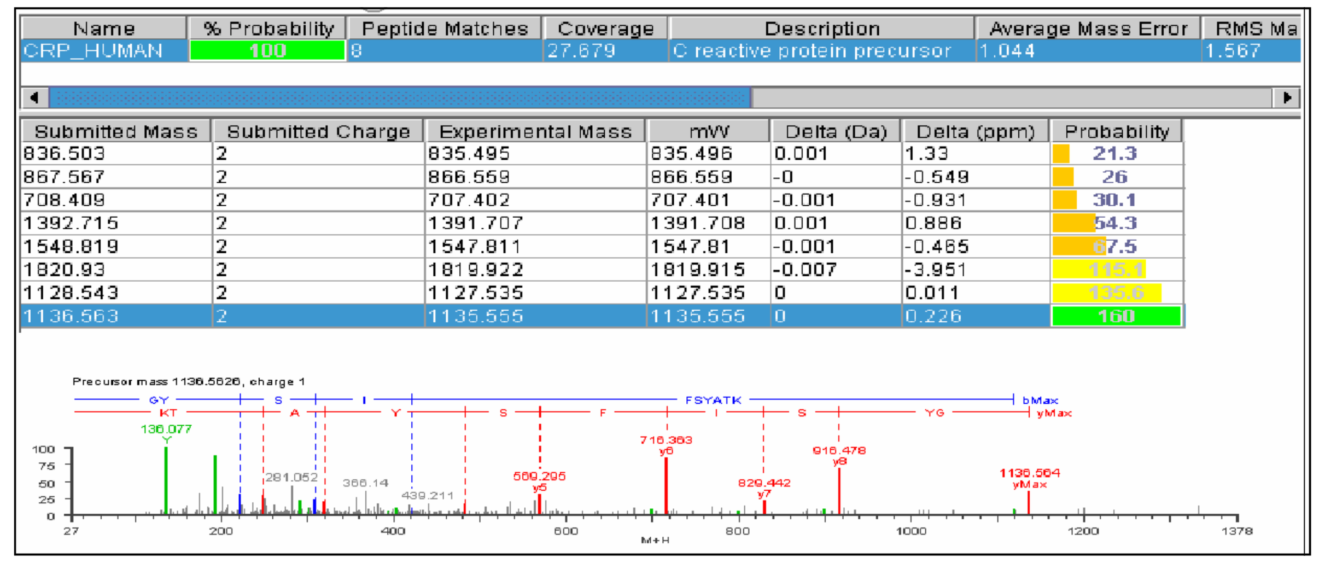

Figure 7 Identification of CRP from the LC-MS expression data acquired from the elevated CRP sample. Eight peptides matched to Human C-reactive protein precursor, corresponding to 27,7% coverage of the amino acid sequence. The average mass error for these peptides was 1.04ppm with an RMS error of 1.57ppm.
The matched fragment ions are highlighted in the display for peptide GYSIFSYATK

Observed protein differences in CRP elevated serum
In addition to the elevation of CRP, a number of proteins were found to be unique to either the control or CRP elevated serum and therefore expression ratios could not be quantified.

Peptide level quantification
The peptide ions (EMRT's) from replicate injections of the same sample and across different sample conditions are clustered by their accurate mass and retention time. These clustered EMRT pairs can subsequently be plotted, to display the average mass of the clustered peptide vs log Expression ratio. This allows up and down regulation of peptides between samples to be visualized and subsequently the parent protein can be identified.

ExpressionAnalysis 14 - Result 1

Protein ID	Description	Sample 2:Sample 1
4	Antithrombin III precursor ATIII PRO0309 Homo sapiens Human	Sample 2
24	C reactive protein precursor Homo sapiens Human	Sample 2
30	Antithrombin III precursor ATIII PRO0309	Sample 2

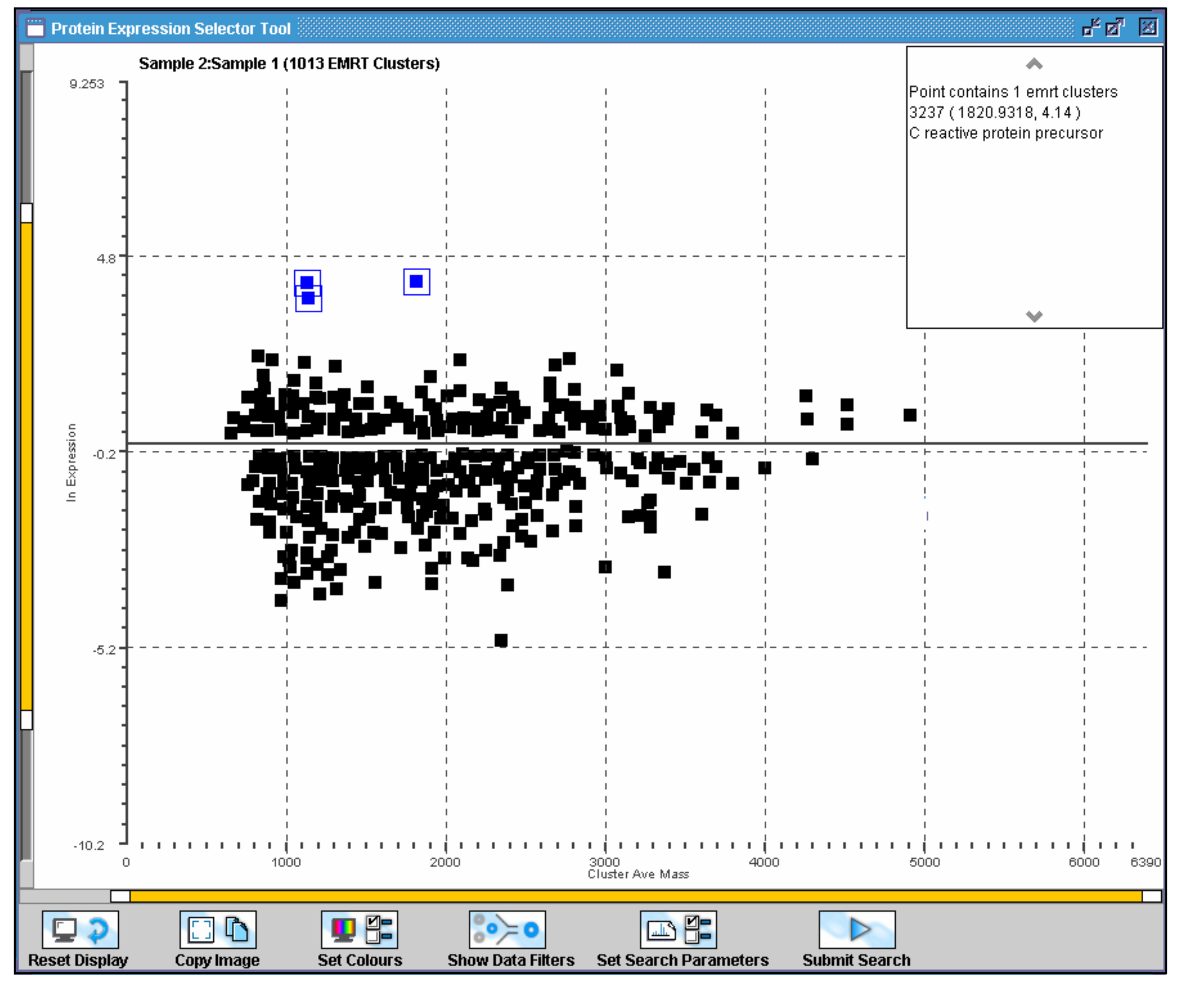

Figure 8. EMRT cluster results plotting the mass of the peptide cluster (X axis) against the log10 expression ratio (Y axis). In this example we are comparing the normal control serum (sample 1) against the elevated CRP serum (sample 2). A total of 1,013 filtered EMRT's are displayed in this plot. From this plot groups of EMRT's can be selected for databank searching to identify the parent protein. The expression ratio, or fold change, can be used to drive the selection of EMRT's to search against the databank. For example, in this case the EMRT's highlighted in blue all show a similar fold change and were selected for databank searching, resulting in the identification of CRP. The matched peptides and their calculated fold change are shown below:

ExpressionAnalysis 12 - Result 1

Cluster ID	Average Mass (MH+)	Average RT (Mins)	Description	Peptide Sequence	Probability	Sample 2:Sample 1
1183	1136.5602	37.912	C reactive protein precursor	GYSIFSYATK	0.515	41.2644 (3.72 +/- 0.58) [1
1154	1128.5414	17.57	C reactive protein precursor	ESDTSYVSLK	0.515	60.9467 (4.11 +/- 0.9) [1.
3237	1820.9318	50.132	C reactive protein precursor	YEVQGEVFTKPQLWP	0.515	62.8028 (4.14 +/- 1.05) [1

In addition to the peptides (EMRT's) that show an up-regulation in the CRP elevated sample, several groups of EMRT's exhibited a significant down regulation, shown in Figure 9a (highlighted in blue). Databank searching of these EMRT's (Fig 9b) identified the parent proteins Complement C4 and apo-lipoprotein H. Interestingly Complement C4 is part of the acute-phase plasma protein response and its down regulation in this sample is unexpected.

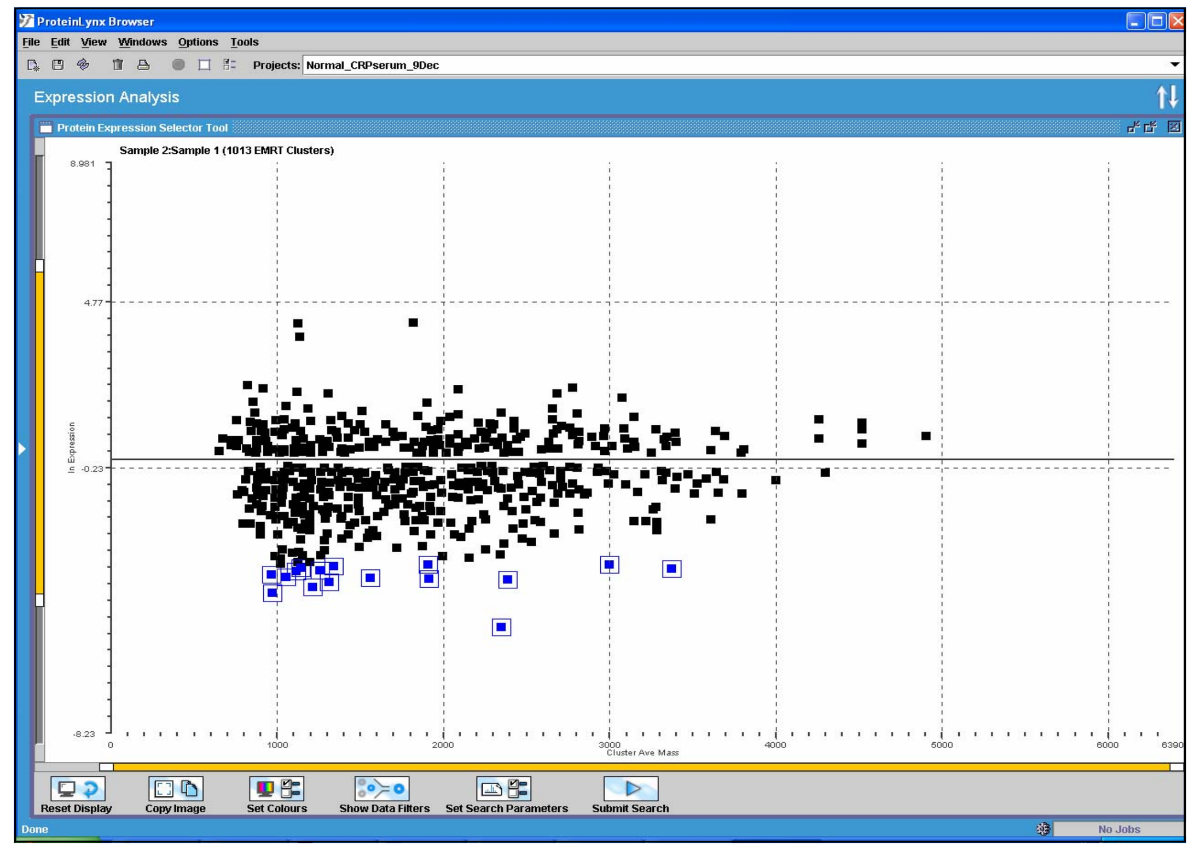

Figure 9a.

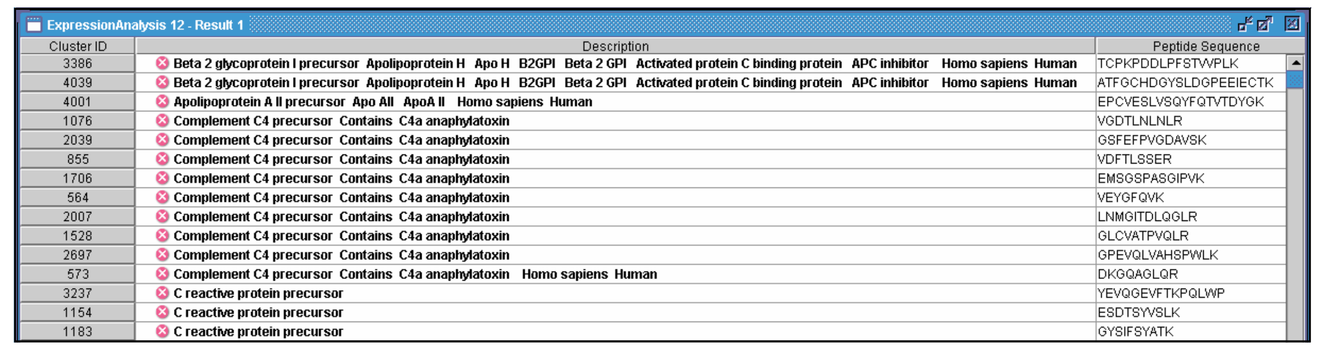

Figure 9b.

Summary

- Expression profiling of C-reactive protein (CRP) in depleted Human serum has been shown using an LC-MS based method
- The use of an affinity based depletion to specifically remove the six most abundant proteins was required to enable identification of CRP in both control and CRP elevated samples.
- Identification of CRP was made from the LC-MS datasets following data processing and databank searching
- Relative quantification of CRP could be made at the peptide level with the mean value of 55, calculated from 3 peptides
- In addition to the up regulation of CRP, a number of further proteins were down regulated in the patient serum.
- This approach provides a comprehensive method of relative protein profiling suitable for time course or multi-patient studies

References
1. Gygi SP, *et al.*. *Nat Biotechnol.* 1999 Oct;17(10):994-9
2. T.P. Conrads, K*et al. Anal. Chem. 73* (2001), p. 2132.
3. S. Purvine, *et al.* Proteomics 2003, 3, 847
4. T. Mckenna et al. *Proceedings of the 52nd ASMS conference*, Nahville, TN
5. Claus, D., A. Osmand, and H. Gewurz. 1976. *J. Lab. Clin. Med.* 87:120-128
6. Deichmann M, *et al.* .*British Journal of Cancer advance online publication*, 27 July 2004.