

# Proteome Profiling by Multidimensional Protein Separation and Automated Lab-on-a-Chip Technology

## Application

## Proteomics

### Authors

Peter Mrozinski, Bill Wilson, and Hongbin Liu  
Agilent Technologies, Inc.  
2850 Centerville Road  
Wilmington, DE 19808-1610  
USA

### Abstract

**Technologies and methodologies for proteomic studies often involve a combination of protein fractionation, identification, and quantitation by liquid chromatography tandem mass spectrometry (LC/MS/MS) and in many cases pattern comparison of mass spectra from a complex sample. A common goal of these efforts is the discovery and validation of biomarkers. While current mass spectrometry-based proteomic methods are highly developed and an important component for biomarker discovery, the time required for these methods makes them impractical for screening applications. In addition, current methods of pattern recognition, while capable of the speed required, suffer from a lack of reproducibility, tend to focus on high-abundance proteins, and fail to identify the actual biomarkers.**

**We test here a concept to combine several novel separation approaches for protein fractionation with an automated Lab-on-a-Chip platform and nano LC/MS/MS based technologies to provide a rapid, quantitative proteome profiling method for the serum/plasma proteome. The Lab-on-a-Chip methodology affords high-throughput analysis combined with reproducible sizing and quantitation, making it a potentially powerful and enabling technology to rapidly compare the proteomes of multiple samples in a wide variety of applications.**

### Introduction

A common goal of many laboratories is the discovery and validation of protein biomarkers for use in the study of disease, drug efficacy, off-target drug effects, dose optimization, and toxicity. Multiple techniques, methods, and tools are often applied to this endeavor. Each has its own unique set of advantages and limitations, all of which are constantly debated and discussed within the scientific community. Various tools can be optimally applied at different times within the protein biomarker “lifecycle”. During discovery, mass spectrometry (MS) is the preferred tool and is considered by many to be the “gold standard” for identification and quantitation. During the validation stage, however, the requirements for the measurement approach change to accommodate the large number of samples and the direct comparison of the results. In addition, the validation system must be robust, easy to use, accurate, and linkable to the discovery phase results. Tools involving pattern recognition can be ideally suited for the validation stage as long as the above requirements are met. While the MS-based tools can, in principal, be used in the validation stage, the typical time for analysis of a proteome using these tools can take 1-month or longer, making them impractical for measuring and comparing a large number of proteomes needed for validation efforts.

A common sample for biomarker discovery is human plasma/serum in part because of its ease of availability, but more importantly, plasma/serum is believed to contain the largest representation of the human proteome available from any biological sample. Of particular interest are the subsets of tissue proteomes that may be present in serum



and the changes that are represented within these proteome subsets as a result of physiological state, time, or perturbation. These partial tissue proteomes are comprised of proteins normally secreted by tissues (such as the liver and intestines). These may be comprised of proteins that normally function in cells but are released due to cell damage or death (several of these proteins are already used as markers for myocardial infarction), proteins secreted from tumor cells or other types of diseased tissues, or proteins from xenobiotics.

The analysis of human plasma/serum presents a number of technical challenges from both the dynamic range of protein amounts ( $10^{10}$ – $10^{12}$ ) as well as from the structural complexity of these proteins. Current instrumentation, such as LC/MS, is capable of analyzing proteins present in amounts spanning a dynamic range of  $10^3$ – $10^4$ . Addition of an incremental protein prefractionation step is commonly proposed to enable one to “dig down” an additional decade into the proteome. An open question is, how deep must one “dig down” in order to enable effective biomarker discovery and validation? An additional complication of proteome analysis of human plasma/serum arises from the fact that all of the proteins present are not at equal concentrations. A few are so dominant, such as serum albumin, IgG, IgA, haptoglobin, anti-trypsin and transferrin, that they make up 85% of the total protein mass and are referred to as high-abundance proteins. Selective removal of these proteins is essential because they are so abundant that they effectively “mask” the detection and measurement of most other proteins present in plasma/serum.

The aim of this application note is to examine the value of using the 5100 automated Lab-on-a-Chip platform (ALP) for protein fraction analysis. A control human serum sample (normal) and a human serum sample from a colon cancer patient (colon cancer) are prepared by a combination of novel separation technologies for protein fractionation (Multiple Affinity Removal System and macroporous reversed phase (mRP-C18) chromatography) and analyzed by the 5100 ALP and nano LC/MS/MS. Comparisons of the resulting data indicate several advantages to using the 5100 ALP system for biomarker validation and other, proteome-comparative investigations.

## Experimental

### Samples

Colon cancer and normal serum samples were purchased from Genomics Collaborative (Cambridge, MA). The colon cancer serum donor and the normal serum donor had similar demography, ethnicity and medical history. Thirty microliters of each of the two serum samples was prepared for analysis.

### Affinity Separation

Sample processing and fractionation were performed according to the protocol provided with the Agilent Multiple Affinity Removal System. Crude human serum was diluted five times with Buffer A (30- $\mu$ L crude serum + 120- $\mu$ L Buffer A). To remove particulates, the diluted serum samples were spun through a 0.22- $\mu$ m spin tube at  $16,000 \times g$  at room temperature for 2 minutes. Seventy-five microliters of this diluted sample was injected onto a  $4.6 \times 50$ -mm affinity column (Agilent part number 5185-5984).

Affinity chromatographic separation was performed using the customized analytical system comprising G1312A Binary Pump, G1329A ALS with G1330B Thermostatted, G1316A Thermostatted Column Compartment, G1315B DAD, and the G1364 Fraction Collector with G1330B Thermostatted. The conditions were: Buffer A at a flow rate of 0.25 mL/min for 9 min, Buffer B at a flow rate of 1.0 mL/min for 3.5 min, Buffer A at a flow rate of 1.0 mL/min for 7.5 min (Table 1).

**Table 1. Affinity Separation Conditions**

Solvent A = Buffer A, Solvent B = Buffer B

LC Timetable				
	Time (min)	%B	Flow rate (mL/min)	Max. press (bar)
1	0.00	0	0.25	120
2	9.00	0	0.25	120
3	9.01	100	1.00	120
4	12.50	100	1.00	120
5	12.60	0	1.00	120
6	20.00	0	1.00	120

The flow-through fraction was collected at 2.0–4.5 min. Each of the two samples (colon cancer and normal) was injected twice and the flow-through fractions pooled. Each pooled flow-through consisted of approximately 300 µg of protein.

### mRP-C18 Separation

Each of the two depleted samples (colon cancer and normal) was fractionated further on an Agilent 4.6-mm id × 50-mm mRP-C18 column (part number 5188-5231). The analytical system consisted of: G1312A Binary Pump, G1329A ALS with G1330B Thermostatted, G1316A Thermostatted Column Compartment, G1315B DAD, G1364 Fraction Collector with G1330B Thermostatted. The total protein (~300 µg) from each flow-through fraction (in approximately 1 mL of Buffer A) was prepared for direct-loading onto the mRP-C18 column by adding 0.48 g of urea pellets and 13 µL of neat glacial acetic acid. This step ensures that the proteins will not be carbamylated during analysis at the elevated temperatures (80 °C) required for the separation. Approximate final concentrations were 6M urea and 1.0% acetic acid. After sample preparation, the immunodepleted serum was separated under RP conditions using a linear multi-segment gradient (Table 2).

**Table 2. mRP-C18 Separations Conditions**

Solvent A: water/0.1% TFA	Pressure limits: 200 bar
Solvent B: acetonitrile/0.08% TFA	Column Temp: 80 °C

#### LC Timetable

	Time (min)	%B	Flow (mL/min)	Max. press (bar)
1	0	3	0.75	200
2	1	3	0.75	200
3	6	30	0.75	200
4	39	55	0.75	200
5	49	100	0.75	200
6	53	100	0.75	200
7	58	3	0.75	200

Fraction collection was performed by time, collecting 1.5-min time slices starting at 6.5 min and continuing to 53 min. A total of 32 fractions were collected for each of the colon cancer and normal samples.

### Fraction Analysis

#### 5100 ALP

The gel electrophoretic separation with the Agilent 5100 ALP is not compatible with acetonitrile. Therefore, each of the mRP-C18 fractions was evaporated from 1.125 mL to complete dryness with a SpeedVac. The samples were then reconstituted in 10 µL of PBS (phosphate buffered saline solution).

The Protein 200 HT-2 assay and microfluidic chip (G3011-87300) were used according to the manufacturer's recommendations in conjunction with the appropriate reagent kit (G3011-68750). Unlike the Agilent 2100 Bioanalyzer chip, the 5100 ALP chip is reusable for 2000 samples and has external capillaries that enable it to withdraw samples from a sealed titer plate. The reagent kit comes complete with sample buffer including internal standards, protein calibration ladder, chip storage buffer, and ready to use reagent plates. Fraction samples were prepared in Eppendorf PCR twin plates. Four microliters of sample was dispensed to the well, followed by 2 µL of sample buffer. The plate was sealed with a Remp plate sealer and was heated to 95 °C for 5 minutes on a PCR (polymerase chain reaction) thermocycler. The plate is then cooled, the foil seal was removed, and 24 µL of deionized water was added to each well and mixed thoroughly. The plate was then resealed with foil and was ready to run. Approximately 40% of each reconstituted fraction was used for this analysis; the remaining 6 µL was retained for MS analysis.

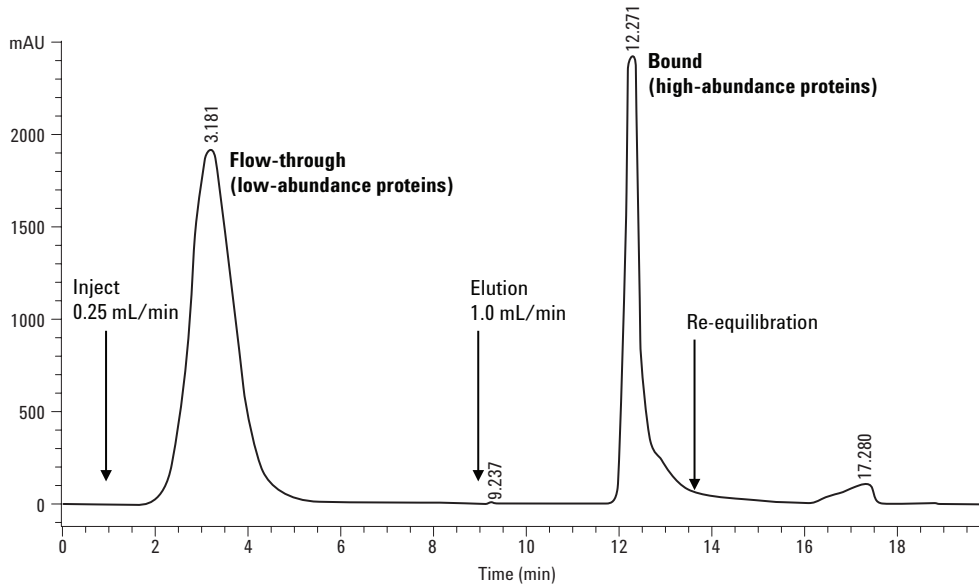
### Nano LC/MS/MS Analysis

The fractions from the mRP-C18 column were prepared with a tryptic digestion using a 2,2,2-trifluoroethanol (TFE) protocol [1]. The digested samples were analyzed by (LC/MS/MS) on an Agilent LC/MSD TRAP XCT. LC/MSD conditions were: scan range of 400–1500 *m/z*, AutoMS settings of Prefer Double Charged Ions, Active Exclusion ON, SmartFrag ON, Peptide Mode ON, and Averages set to 2. Nano LC separations were performed on a ZORBAX 300SB-C18, 3.5 µm, 150 mm × 75-µm column, 20 µL injected (enrichment column was a ZORBAX 300SB-C18, 5 mm × 300 µm). Solvent A: 0.1% formic acid in water, Solvent B: 0.1% formic acid in acetonitrile (90-minute gradient from 5% B to 90% B), flow rate: 300 nL/min. Data analysis of all LC/MS/MS experiments was performed with Agilent Spectrum Mill software.

## Results and Discussion

### Protein Fractionation

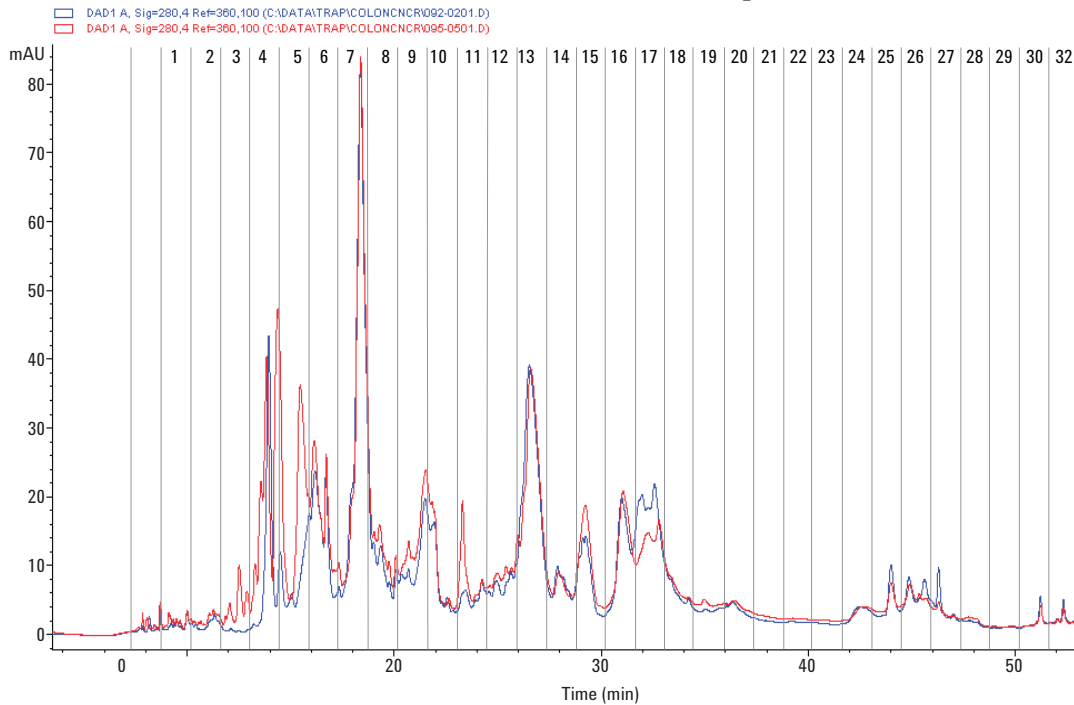
Both serum samples were first immunodepleted of their six most abundant proteins - albumin, transferrin, IgG, IgA, haptoglobin, and antitrypsin in a single step with very high specificity. Figure 1 shows the typical separation achieved. Selective removal is essential because these proteins are so abundant that they effectively “mask” the detection and measurement of most other proteins present in plasma/serum [2].



**Figure 1. Chromatogram from Multiple Affinity Removal column. The flow-through was collected from 2 to 4.5 minutes.**

Reproducibility of the depletion column was tested by the analysis of 300 injections on a single column. The two buffers provided with the system are optimized for promoting maximum binding capacity (sample loading and selectivity), column lifetime (elution and regeneration), and reproducible sample fractionation. The depleted fraction is amenable to the mRP-C18 separation step with minimal preparation.

The immunodepleted samples were then fractionated via the mRP-C18 column into 32 fractions each (Figure 2). A recent development, the mRP-C18 column affords near quantitative recovery of proteins [3]. Previous work showed that highly reproducible protein separations with high recovery (>98%) and excellent resolution was achieved under these operating conditions. This column is also able to tolerate high protein loading (up to 400 µg) and still provide excellent recovery and separation.



**Figure 2. RP chromatographic separation of depleted samples from colon cancer and normal serum. Overlay of 280-nm absorbance signals for cancer (red chromatogram) and normal (blue chromatogram). Fractions collected are represented by the vertical lines along the time scale (32 fractions collected).**

As apparent in the overlaid chromatograms shown in Figure 2, several fractions show obvious differences while others are very similar. For example, fraction 5 shows a quartet of peaks for the normal sample and essentially a flat baseline for the colon cancer sample. Fraction 6 shows similar peaks but at different amounts. Conversely, fraction 14 shows almost identical peak profiles.

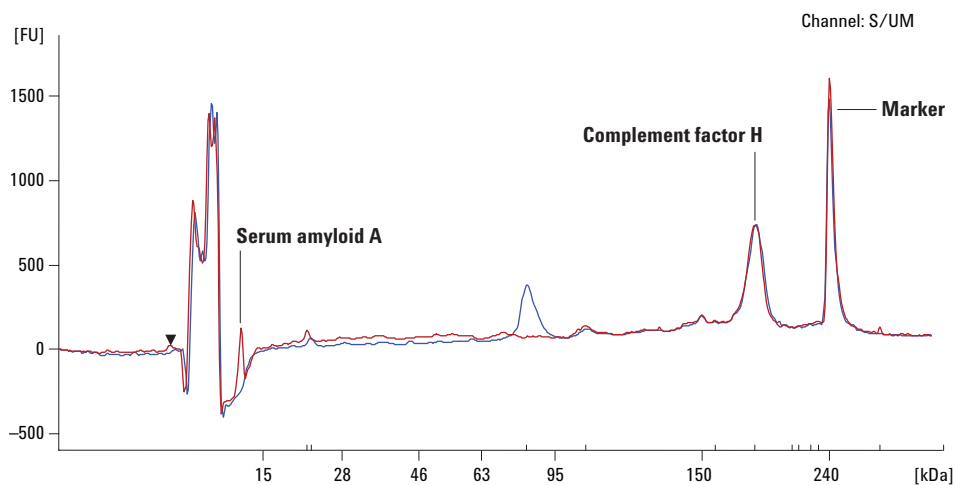
### Protein Analysis

The 32 fractions from each sample were analyzed by on-chip electrophoresis on the Agilent 5100 ALP. Similar to the 2100 Bioanalyzer, the 5100 ALP uses chip-based capillary gel electrophoresis to rapidly separate proteins [4]. External capillaries on the chip allow the instrument to draw from up to 12-sealed titer plates for high-throughput sizing and quantitation of proteins. Calibration ladders are run at regular intervals to ensure accurate sizing and can provide semiquantitative results for unknown proteins. Detection is achieved by a fluorescent dye that is incorporated into the solvent delivery system (SDS) layer surrounding the protein. Consequently, the 5100 ALP has essentially universal detection for proteins in the

10–2000 ng/ $\mu$ L range. Because of this unique detection mechanism, the instrument is insensitive to interference from small molecule additives. After performing its fully automated instrument preparation steps, the 5100 ALP was able to analyze all 64 fractions in a little over 1 hour.

The overlaid electropherograms for fraction 5 are given in Figure 3. There is excellent agreement between the ALP and MS data for the peak at 187 kDa on the electropherogram. The MS identifies this as Complement factor H and shows equal abundance for the cancer and normal serum. Similar results are observed in the electropherogram. The peak at  $\sim$ 13 kDa is present in the cancer fraction and absent from the normal fraction. The peak at 82 kDa in the normal serum is clearly a distinguishing feature between the two samples; unfortunately, a protein at this mass was not detected by the LC/MS/MS.

In fraction 6, the ultraviolet (UV) trace showed similar peak retention times (RTs) between the cancer and normal serum but differences in amount. The electropherograms (Figure 4) confirms this by showing a common doublet in both

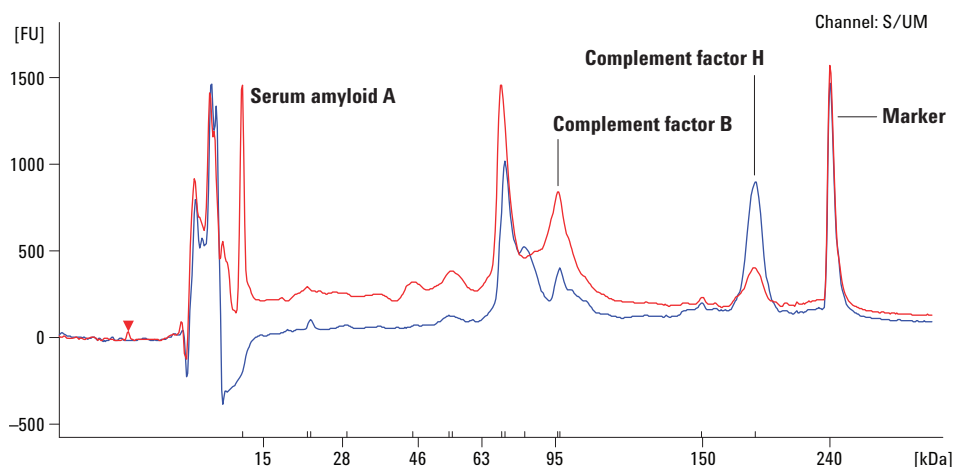


**Figure 3. Overlaid electropherograms for fraction 5 (blue = normal/red = cancer).**

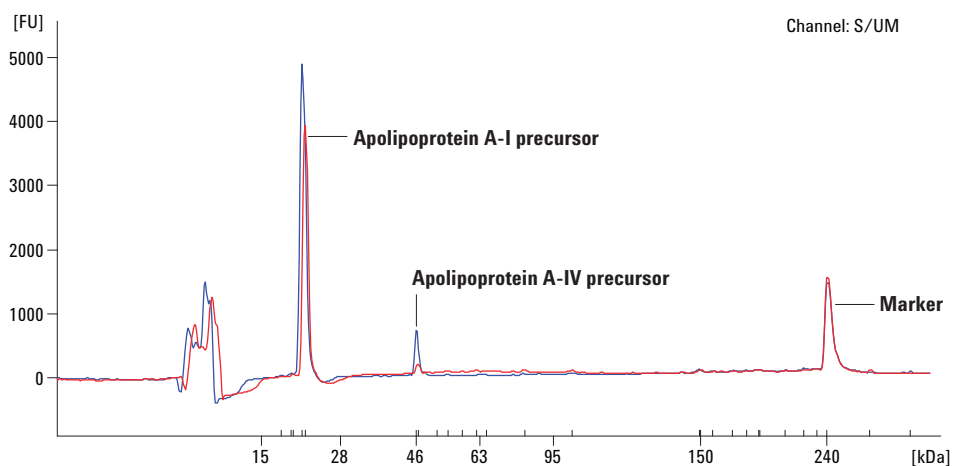
traces at 71 and 95 kDa. The large peak at ~13 kDa likely corresponds to the first peak in fraction 6 as it is absent from the normal serum and would partially elute in fraction 5 as noted above. The LC/MS/MS analysis identifies this protein as serum amyloid A (MW 13,532 Da) and corroborates the over 100-fold increase observed in the electropherogram. Complement factor H precursor (MW 139,126 Da) is seen at a four-fold higher concentration in the normal serum compared to the cancer serum on both the MS and ALP data. In addition, several differences can be seen in the electropherograms and offer a potential way to discriminate between the two samples. For example, the cancer serum shows a unique peak at 44 kDa while the normal serum shows a unique peak at 81 kDa.

Fraction 14 was very similar based on the UV trace. The electropherograms are shown in Figure 5 and confirm that several similar proteins are present in this fraction. One of the advantages of electrophoretic analysis, illustrated here, is the comparative quantitation that is possible. The ratio of the two apolipoprotein precursors between normal and cancer serum is different and this subtlety is readily seen on the electropherograms. The relative amount seen corresponds well with the MS data.

The comparative analysis done on the 5100 ALP provided an additional benefit for the LC/MS/MS analysis. The sizing information of the differentially expressed proteins from the 5100 ALP analysis can be used with the MS search results for a



**Figure 4. Overlaid electropherograms for fraction 6. (blue = normal/red = cancer).**



**Figure 5. Overlaid electropherograms for fraction 14. (blue = normal, red = cancer).**

more definitive identification of the possible biomarker proteins. There was excellent agreement between the 5100 ALP and XCT TRAP data for both the proteins found in each fraction and their relative amounts. The MS-based quantitation values were determined by both Spectrum Mill's Mean Peptide Spectral Intensity and by a Spectral Sampling method [5]. Table 3 shows the proteins found in fractions 5, 6, and 14 and their relative amounts in the normal and cancer serum sample.

## Conclusion

In this proof-of-principle application, the combination of protein fractionation (immunodepletion and mRP-C18) with the 5100 ALP demonstrates some of the requirements needed for comparing multiple proteomes during biomarker validation and other applications. High-throughput, automation, and reproducibility are critically important for analyzing the large number of validation samples. The accuracy, ease of use, and the ability to link protein size back to the MS results used during discovery increase the utility of this approach. While MS techniques form the foundation of biomarker discovery investigations, other techniques like Lab-on-a-Chip may complement and be able to accelerate the confirmation of biomarkers and other proteome-comparative studies.

## References

1. Meza, J. et. al., "Improved Tryptic Digestion of Protein Using 2,2,2-Trifluoroethanol (TFE)", Poster presentation, ABRF 2004 (5989–1781).
2. Mrozinski, P. et. al., "Enhanced High Capacity Affinity Depletion of the Top Six Proteins from Human Serum", Agilent Technologies, publication 5989–2881.
3. Barrett, W. et. al., "Comparative Analysis of Serum Samples with the Agilent mRP-C18 Column - A New Technology for Protein Fractionation", Agilent Technologies, publication 5989–2505.
4. Hawtin, P. et. al., "High-Throughput Protein Expression and Purification Analysis Using the 5100 Automated Lab-on-a-Chip Platform", Agilent Technologies, publication 5989–3507.
5. Liu, H. et. al., "A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics", (2004) *Anal. Chem.*, **76**, 4193–4201.

**Table 3. MS/MS and ALP Search Results**

Protein	MW (MS)	mRP-C18 fraction	Normal	Colon cancer
Serum amyloid A	13532	5 & 6	1X	100X
Complement factor H precursor	139126	5	1X	1X
Complement factor B precursor	85533	6	1X	2X
Complement factor H precursor	139126	6	4X	1X
Apolipoprotein A-IV precursor	45371	14	3X	1X
Apolipoprotein A-1 precursor	30778	14	1.2X	1X

Proteins listed showed comparable relative quantitation values on both the XCT TRAP and 5100 ALP.

## For More Information

For more information on our products and services, visit our Web sites at [www.agilent.com/chem/bioreagents](http://www.agilent.com/chem/bioreagents)

[www.agilent.com/chem/5100alp](http://www.agilent.com/chem/5100alp)

Email technical inquiries to [affinity\\_removal@agilent.com](mailto:affinity_removal@agilent.com)

Peter Mrozinski and Bill Wilson are Proteomics Application Scientists for Proteomics BioReagents in the Integrated Biology Solutions Group with Agilent Technologies.

Hongbin Liu is a Proteomics R&D Scientist for Proteomics BioReagents in the Integrated Biology Solutions Group with Agilent Technologies.

Correspondence: Peter Mrozinski  
Email: [peter\\_mrozinski@agilent.com](mailto:peter_mrozinski@agilent.com)

Agilent shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance, or use of this material.

Information, descriptions, and specifications in this publication are subject to change without notice.

© Agilent Technologies, Inc. 2005

Printed in the USA  
October 3, 2005  
5989-3919EN

