Improved proteomic enrichment and workflow strategies

Swapan Roy, Ph.D.[1], Haiyan Zheng, Ph.D.[2], Caifeng Zhao[2], Absari Arpa[1], Matthew Kuruc[1],
[1]Biotech Support Group LLC, 1 Deer Park Dr., Suite M, Monmouth Junction, NJ 08852,
[2]Biological Mass Spectrometry facility at RWJ Medical School and Rutgers, Center for Proteomics, Room 010, 174 Frelinghuysen Road, Piscataway, NJ 08854-5638

Presented As A Poster At US HUPO Conference, 04/06/2014-04/09/2014

Abstract

Biotech Support Group has developed a series of enrichment reagents that offer common features and key advantages. After separations, the sub-proteomes retain their structural and functional integrity. All the products are consumable and economical; not being derived from biological sources. We report here applications for these products to specifically deplete albumin by a voidance strategy - called AlbuVoid™, which consequently enriches for the low abundance serum protein content. Such a strategy compares favorably to high abundance immuno-depletion (Agilent column) methods. Also, we continue to advance simplified workflows that utilize on-bead digestion for LC-MS. Together, these methods create efficiencies necessary to support high throughput investigations in all areas of discovery, targeted, functional and chemical proteomics.

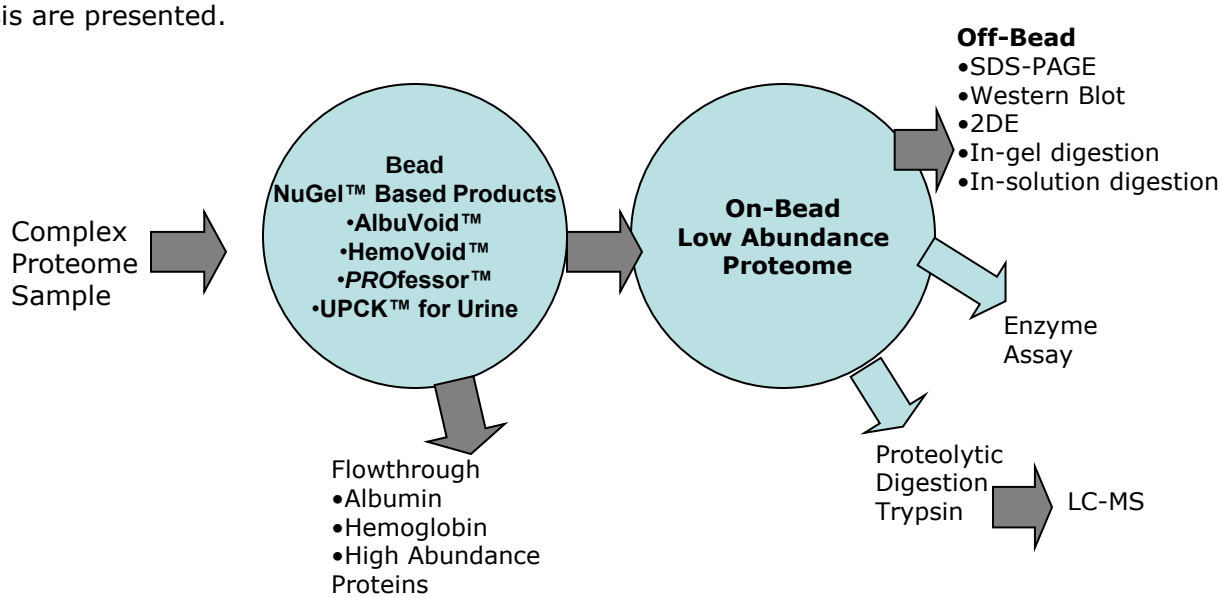
Background & Introduction

A significant challenge in proteomics has been to overcome the analytical bias towards the most abundant proteins, and the complexities of mining the data to a manageable number of biomarker proteins that can be analyzed in more depth. Therefore, three common goals prior to LC-MS are often:

1. To enrich specialized sub-proteomics or alternatively - deplete high abundance protein(s)
2. Get efficient and consistent trypsin digestion
3. Streamline workflows to minimize handling, potential for variance and decreased costs.

To assist in all 3 aspects at once, we have developed a platform surface chemistry with derivative products to support different applications. The surface chemistry platform called **NuGel™**, is based on porous silica, a chromatography industry standard and an advantageous matrix suitable for high performance liquid chromatography. With **NuGel™**, non-specific sites have been virtually eliminated making it an ideal support for protein separations. Through a proprietary polymer coating, silica is crosslinked forming a reactive Poly-Epoxy functionality stable across a wide pH range (pH 2 to 10). From this foundational chemistry, all of the **NuGel™** products are derived.

The **NuGel™** consumable product line offers a bead-based platform to either deplete high abundance proteins through a selective binding strategy, or through a selective voidance strategy, references for which can found on the Biotech Support Group website[1]. In the latter strategy, the high abundance proteins selectively flow through the beads, capturing and enriching the low abundance proteome on the bead. Four such products feature this strategy and are described throughout this poster report. All feature simple bind/wash microfuge protocols requiring no specialized instruments or HPLC. By binding the low abundance proteome to the surface, efficient new workflows for digestion, assay and analysis are presented.

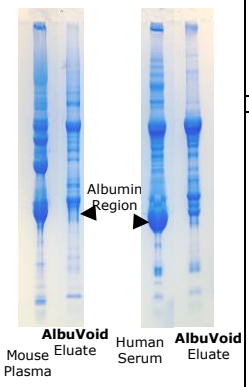


New Strategies for Albumin Depletion

Serum Albumin has the singular highest protein concentration in plasma and sera from all animals. In human serum, albumin represents more than half of the protein mass. To detect low abundance proteins, depletion of one or more high abundance proteins is common and desirable. To address this need, several commercial products incorporating either immobilized monoclonal antibodies or combinatorial peptide libraries have been reported to either deplete high abundance proteins or enrich for low abundance proteins.

Derived from the **NuGel™** silica-based platform, we have devised two strategies for albumin depletion**, AlbuSorb™** and **AlbuVoid™**.

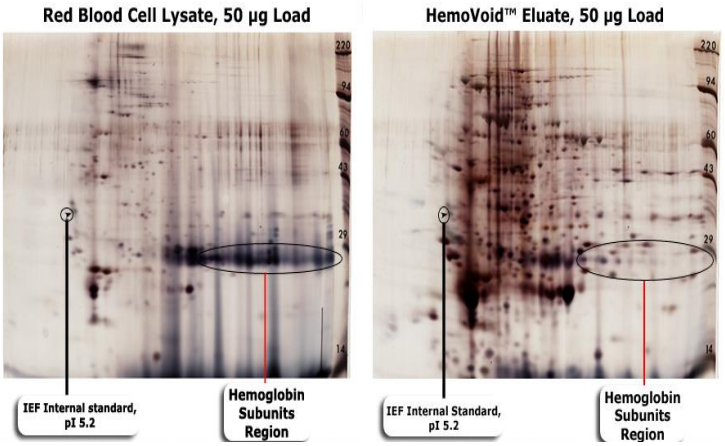
These non-reduced SDS-Page profiles show that albumin removal is specific, and suitable for both mouse plasma and human serum with equal efficiency.



AlbuSorb™ has been used extensively to selectively bind albumin from a wide range of mammalian sera. **AlbuVoid™** is designed to deplete albumin and enrich for low abundance proteins. Both are derived from the **NuGel™** platform. Upon empirically evaluating many chemical mixed-mode ligand/buffer/load binding combinations, a singular surface chemistry was selected that inhibits binding, achieving significant voidance of albumin while adsorbing the vast amount of the remaining protein mass. As it is not immunological, it works across all mammalian species.

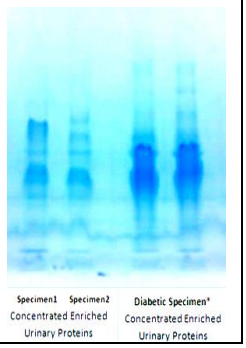
New Strategies for Hemoglobin Depletion

In a similar voidance strategy and derived from the **NuGel™** platform is **HemoVoid™**, for the selective depletion of hemoglobin from erythrocyte lysate. As hemoglobin accounts for 97% of the protein mass of erythrocytes, Lange et al., applied **HemoVoid™** to annotate the human erythrocyte proteome. The article reports 778 proteins from the hemoglobin-depleted soluble fraction, 171 of which were not represented in either the soluble non-depleted fraction or the membrane fraction[2]. In the 2DE comparison below, enrichment of the low abundance proteome is not biased to either molecular weight or pI.



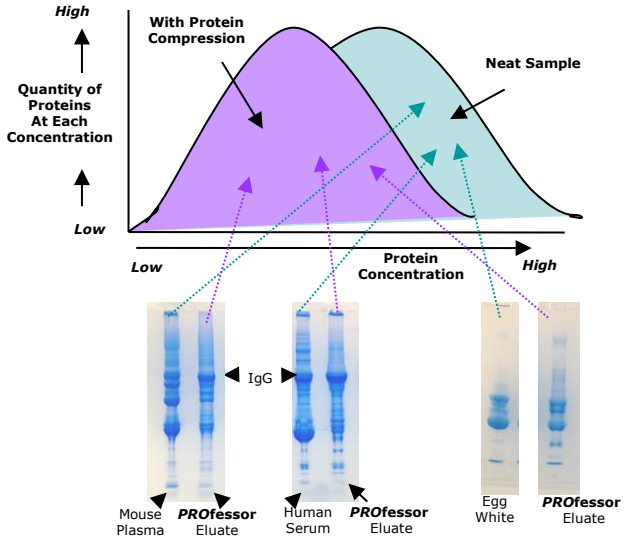
UPCK™ Kit for Urine Proteome Enrichment

As an alternative to solvent precipitation, the **UPCK™** kit adapts the **NuGel™** surface chemistries to enrich and concentrate urine proteins in <60 minute protocol. The proteins retain their enzymatic activity and on-bead digestion can be applied.



NuGel™ *PRO*fessor™ - High Abundance Depletion & Protein Compression

Many methods have been commercialized to reduce the dynamic range of protein concentrations within complex mixtures. **NuGel™ *PRO*fessor™** is the only one that is universal as it can be applied to any tissue from any biological source, and is suitable for small sample and protein loads; only about 1 mg total protein being required. It also preserves the functional and structural integrity of the enriched proteomes as no harsh denaturants are used for elution. The **NuGel™ *PRO*fessor™** material is a composite mixture of 6 mixed mode interaction surface characteristics, each with its own complement bias. The highest abundance proteins bias towards some but not all surface chemistries. Under saturation overload conditions, the highest protein concentrations are then reduced due to competitive displacement effects – a consequence of weak affinity binding. As each of the 6 surface architectures has sufficiently different bias in its complement proteome, together they work to compress the overall protein concentrations relative to their starting concentrations, the low abundance protein content being enriched relative to the high abundance protein content. This is illustrated by the graph and analyses below.



NuGel™ *PRO*spector™ Protein Compression. These non-reduced SDS-Page profiles show that major high abundance bands are greatly reduced. The low abundance bands are enriched relative to their high abundance bands, visually estimated to be 5-10X for most bands. Because of its heterogeneity, the IgG band remains in high abundance. The protein compression is both tissue and species agnostic, with no bias towards molecular weight or pI.

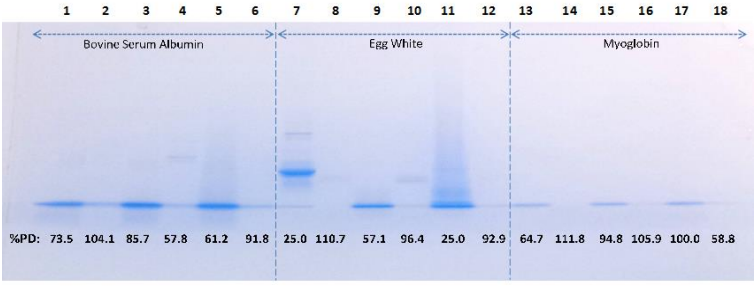
On-Bead Digestion

Common methods to digest proteins with trypsin are: in-solution whereby the steps take place with the sample proteins in solution; in-gel whereby the sample proteins are contained within a gel slice after SDS-PAGE; and immobilized trypsin whereby the sample proteins in solution are contacted with trypsin covalently immobilized to a solid support. Nevertheless, with greater interest in the proteomics community for better workflows and performance for LC-MS analyses, bead-based separations have been coupled to direct proteolytic digestion of the bound protein content. This is commonly called "on-bead digestion" in the literature[3-5]. In a study reported previously using **AlbuVoid™**, we reported several performance metrics whereby the quality and number of tryptic peptides after on-bead digestion, was comparable to those obtained after elution from the beads and in-solution digestion[6].

We continue these investigations here by demonstrating the speed and efficiency of on-bead digestion for hard-to-digest proteins, using the composite **NuGel™ *PRO*fessor™** product. While trypsin digestion is common and very efficient for most proteins in support of LC-MS analysis, improvements in workflow, time, reproducibility and effectiveness for some hard-to-digest proteins is still desirable. Better methods to monitor the consistency of digestion, understanding the mechanisms which impede digestion, and alternative workflows to work around such inefficiencies, will contribute to the higher throughput and reliability required as the field matures.

Material and method:

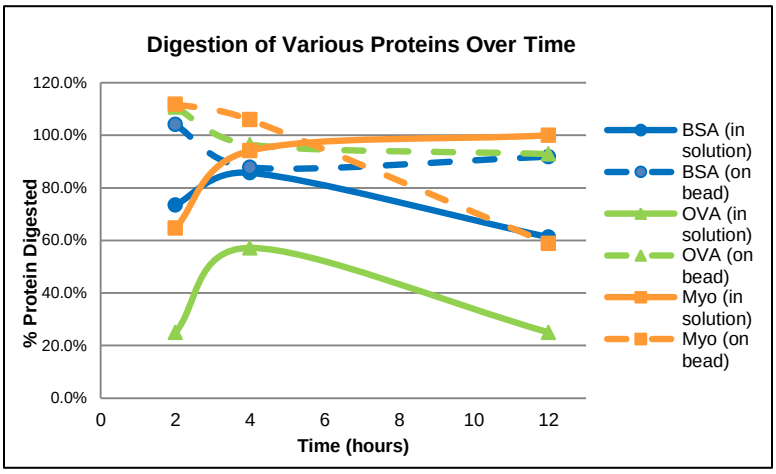
Myoglobin (MYO) and Ovalbumin (OVA) from egg whites were selected as representative hard-to-digest proteins. Bovine serum albumin (BSA) represented an easy-to-digest protein. Myoglobin and BSA were purchased (Sigma), while egg white protein was prepared from a fresh egg. All protein samples were prepared to 2 mg/ml. In solution digestions were carried out on soluble protein. On bead digestions were carried out on 15 mg ***PRO*fessor™** surface according to our established protocol; *see product sheet on website for full details*[1]. Briefly, 100 µls of each sample were applied to the beads and washed. The bead-bound proteins were reduced with DTT, alkylated, then digested with trypsin. The extent of digestion was quantified using a commercial kit (ProDM™, ITSI-Biosciences). Using the kit, digestion is monitored by a colorimetric assay, measured by Nanodrop spectrometry. The results were further analyzed qualitatively using SDS-PAGE, which was run at 150 volts for 50 minutes in a 4-15% Tris-HCl gel.



Lane	Loaded Samples
1	BSA - 2 hr digest (in soln)
2	BSA - 2 hr digest (on bead)
3	BSA - 4 hr digest (in soln)
4	BSA - 4 hr digest (on bead)
5	BSA - ov. digest (in soln)
6	BSA - ov. digest (on bead)

Lane	Loaded Samples
7	EW - 2 hr digest (in soln)
8	EW - 2 hr digest (on bead)
9	EW - 4 hr digest (in soln)
10	EW - 4 hr digest (on bead)
11	EW - ov. digest (in soln)
12	EW - ov. digest (on bead)

Lane	Loaded Samples
13	Myo - 2 hr digest (in soln)
14	Myo - 2 hr digest (on bead)
15	Myo - 4 hr digest (in soln)
16	Myo - 4 hr digest (on bead)
17	Myo - ov. digest (in soln)
18	Myo - ov. digest (on bead)



Results and Discussion

The digestion efficiency in discovery proteomics is as important as the digestion repeatability in targeted proteomics. Having control over the digestion parameters during sample preparation is essential to correctly identify and quantify proteins, and produce reliable results.

We found that the in-solution and on-bead digestions follow the same trend for both Myoglobin and egg white ovalbumin, whereby BSA serves as the positive control. The most significant finding is that the on-bead digestions are much more efficient than the in-solution digestions for all times tested. This was quantified with the monitoring kit, and is further supported by the SDS-PAGE which indicates a close alignment of the bands with the percentage of protein digested (%PD). One possible explanation for these results, is that after disulfide bonds are broken, the protein interiors are better exposed, which allows a more efficient digestion by trypsin. This needs further investigation.

Another significant finding is that all of our model on-bead proteins appear to fully digested in only 2 hours after onset, compared to the in-solution samples that took at least 6 to 8 hours for optimal digestion; and most commonly, are incubated overnight. An inconclusive result is the downward monitor trend of the overnight digestions. Having purchased a commercial kit, the chemistry of the monitoring reagents was unknown to us and its results remain open to interpretation. However, this may be evidence that the proteins are in part over-digested when incubated overnight, after most of the substrate has been hydrolyzed, and trypsin auto-digests. If that is the case, the excess smaller, non-tryptic specific peptide sequences would not be correctly inferred from LC-MS computational protein data-mining. This as well warrants further investigation. Although the optimal time of digestion was not the main focus of this experiment, it does appear that on-bead digestion helps to equalize the digestion of samples that may contain mixtures of easy and hard to digest proteins.

Numerous reports describe methods to accelerate or improve digestion efficiency including denaturants, solvents, heat & pressure cycling and microwaves[7]. Our work suggests that on-bead digestion equalizes the digestion time for easy to hard to digest proteins. Proteases with alternative peptide production, such as Lys-C, Asp-N, Glu-C, or other hydrolyzing enzymes, such as glycosidases, offer still more options. All of these refinements can potentially be adapted to the basic on-bead procedure described here.

In summary, on-bead digestion provides advantaged speed, simplicity and efficiency

- Improves workflow
- Less sample handling & manipulations, lowers variance
- No in-gel digests
- No solution digests
- No C18 desalting
- More consistency, better reproducibility

Serum Proteome Annotation Using AlbuVoid™ On-Bead Protocol

The Biological mass spectrometry facility at RWJ Medical School and Rutgers (http://cabm-ms.cabm.rutgers.edu/) supports client-sponsored research. In the investigation reported here, the facility optimized the method for enriching glyco-peptides from serum. We compared both AlbuVoid™ protocols, on-bead and in-solution, to other enrichment methods including a immuno-depletion column and a lectin affinity gel. The starting sample was pooled rat serum. The investigation considered both the merits of performance and workflows with respect to annotation of total protein and glycoprotein identifications. Below is an overview of the workflow.

Material and method:

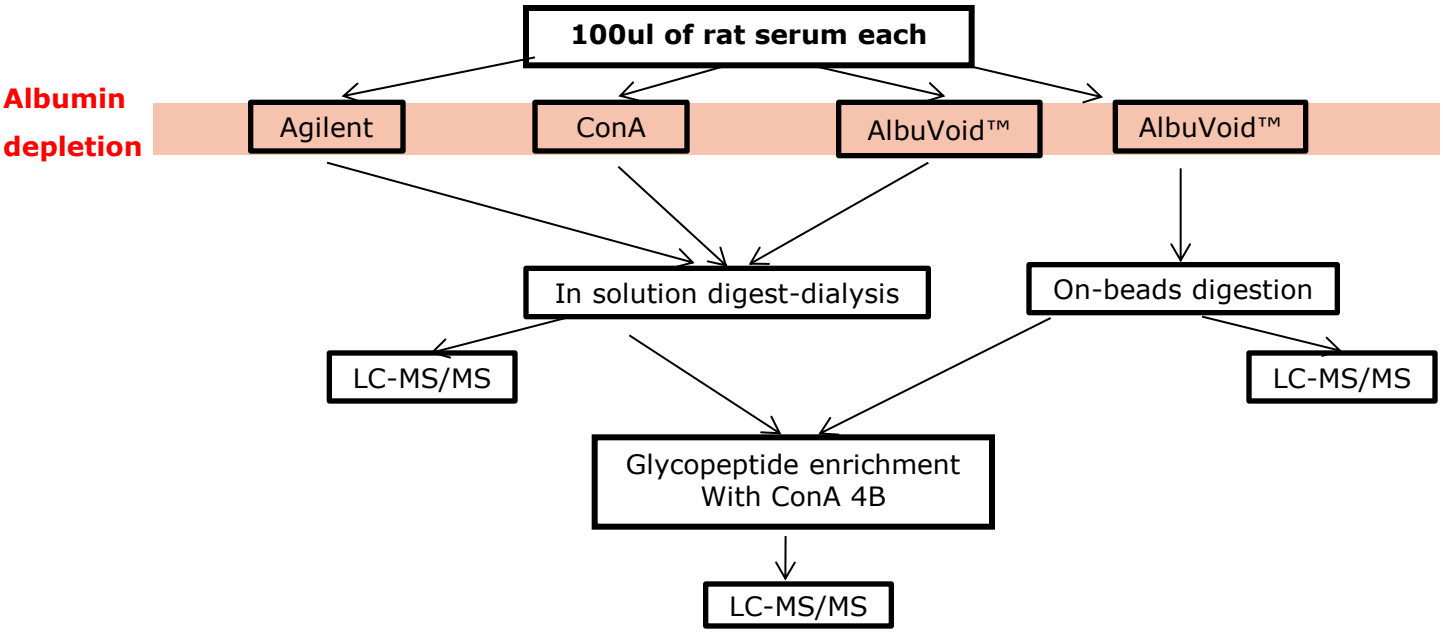
Material: Rat serum: kindly provided by Dr. Kunjan Dave (University of Miami)

Method: 100ul of rat Serum samples were subjected to protein reduction by either Multi Affinity Removal Spin Cartridge Mar3 Mouse-3 (Agilent), AlbuVoid™ kit (Biotech support Group) or HiTrap ConA 4B 1 ml cartridge (GE Healthcare) based on manufacturer's instructions. For in solution digests, after proteins were collected, 6M Guanadine HCl, 10mM Tris HCl, pH8, 20mM DTT were added to the samples and incubated at 60°C for 30 min. The samples were cooled to room temperature prior to addition of 40 mM of iodoacetamide and incubated in the dark at room temperature for 1 hour. The samples were then dialyzed against 2M Urea, 50 mM NH4HCO3 for 2 hours and then 50mM NH4HCO3 for 4 hours. Trypsin was added to the sample at 1:50 (w:w, trypsin:sample) and incubated at 37ºC overnight. The digestion protocol was modified for AlbuVoid on-bead digestion. After protein binding and wash steps are complete, we added 100µls of 10 mM DTT solution to the beads, mixed and incubate at 60°C for ½ hour. After cooling, we added 100 µls of 50 mM iodoacetamide to the DTT/bead suspension, mixed and incubated in the dark for 1 hour, then centrifuged and discarded supernatant. On-bead digestion is done by adding 200 µls at 1:50 (w:w, trypsin:sample) to the beads. After overnight digestion at 37°C, the tubes were centrifuged and peptide filtrate was retained. To further extract remaining peptides, we added 200 µls of 10% solution of formic acid to the beads and incubated for 15 minutes at 37°C. This was then centrifuged and this volume was added to the first volume. The tryptic digests were vacuum-dried.

For glycol-peptide enrichment, peptides equivalent of 25ul of serum were solubilized in binding buffer (20mM Tris-HCl, 0.5M NaCl, 1mM MnCl2, 1mM CaCl2, pH7.4), incubated with Con-A 4B (40ul of 50% slurry in binding buffer) for 1 hour, washed 5 times with Washing buffer (10% acetonitrile, 50mM NH4HCO3) then eluted 2 times with 0.5M mannoside. The glyco-peptides enriched samples were then treated with PNGase F at 37 degree for overnight.

LC-MSMS was done using nanoLC-MS/MS using a RSLC system (Dionex, Sunnyvale CA) interfaced with a LTQ Orbitrap Velos (ThermoFisher, San Jose, CA). Samples were loaded onto a self-packed 100µm x 2cm trap packed with Magic C18AQ, 5µm 200 A (Michrom Bioresources Inc, Aubum, CA) and washed with Buffer A(0.2% formic acid) for 5 min with flow rate of 10ul/min. The trap was brought in-line with the homemade analytical column (Magic C18AQ, 3µm 200 A, 75 µm x 50cm) and peptides fractionated at 300 nL/min with a multi-stepped gradient (4 to 15% Buffer B (0.16% formic acid 80% acetonitrile) in 35 min and 15-25%B in 65 min and 25-50%B in 50 min). Mass spectrometry data was acquired using a data-dependent acquisition procedure with a cyclic series of a full scan acquired in Orbitrap with resolution of 60,000 followed by MSMS scans (CID 35% of collision energy) of 20 most intense ions with a repeat count of two and the dynamic exclusion duration of 60 sec.

The LC-MS/MS data was searched against the Rat Ensembl database using X!tandem (thegpm.org) with carbamidomethylation on cysteine as fixed modification and oxidation of methionine and deamidation on Asparagine as variable modifications using a 10 ppm precursor ion tolerance and a 0.4 Da fragment ion tolerance. Glyco-peptides were determined as deamidation at the asparagine where NXS/T motif.



Albumin depletion impacts glycopeptide enrichment The two methods of AlbuVoid™ digestion and Agilent are of comparable performance based on the # of glycopeptides or glycoprotein IDs. ConA as a protein-level enrichment step is not suitable.

Albumin depletion	AlbuVoid-on beads digestion	AlbuVoid-Elution	ConA	Agilent
# of proteins	265	272	226	211
# of unique peptides	3195	2870	3093	3409

Comparison between 4 serum depletion method based on number of proteins and peptides identified
The two methods of AlbuVoid™, and Agilent and ConA enrichment of glyco-protein yielded similar results based on number of proteins and number of unique peptides identified. Equal amount of peptides (equivalent serum volume) were loaded on the nano-LC-MSMS using a 3 hour gradient.

Glycopeptide enrichment with ConA	AlbuVoid-on beads digestion		AlbuVoid-elution in solution digestion		ConA		Agilent	
# of glycopeptides	496	477	516	514	379	367	575	551
# of glycoproteins	145	137	143	139	108	107	137	141

Comparison of depletion methods by glycopeptides
The number of glycopeptides and glycoproteins are similar between Agilent and both AlbuVoid™ digestion methods. AlbuVoid™ on-bead digestion and elution with solution digest produced 70% common glyco-peptides. However, only about 30% of the glycopeptides are common between Agilent and AlbuVoid™ methods. Which means the two methods are complementary. Each method has around 150 glycoproteins and 600 glycopeptides, added together, there will be around 200 (35% more) glycoproteins and 1050 glycopeptides (75% more).

Glycopeptide enrichment with ConA	Agilent	AlbuVoid-on beads digestion	Common	Combined	Agilent	AlbuVoid-Elution in solution digestion	Common	Combined
# of glycopeptides	662	544	174	1032	662	568	182	1048
# of glycoproteins	153	151	95	209	153	150	97	236

Reproducibility
As the AlbuVoid™ on-bead method produced similar results to its in-solution method, on-bead was chosen for additional tests due to its much simpler protocol. Multiple different samples were then evaluated using both the Agilent and the AlbuVoid™ methods.

In the table below, the AlbuVoid™ method consistently produced 20-30% more protein IDs.

	Agilent method - shown in # of samples				AlbuVoid on-bead method - shown in # of samples			
	One	Two	Three	Four	One	Two	Three	Four
# of proteins	240	198	184	163	330	254	235	217
# of peptides	4103	3008	2338	1734	3935	2896	2481	2162

Albumin Depletion, Low Abundance Enrichment and On-Bead Digestion

We have demonstrated that for annotating the serum proteome, **AlbuVoid™** selectively voids albumin and enriches the low abundance proteome providing a suitable alternative to other depletion strategies. Furthermore, the on-bead digestion protocol provides comparable annotation and better workflow.

•Unlike immuno-affinity, it is a consumable, one-time use product with no potential for cross-contamination or reduced performance upon regeneration.
•20-30% more protein/peptide identifications than immuno-affinity depletion.
•On-Bead digestion improves workflow efficiency
•No in-gel digests, no solution digests, no C18 desalting, more consistent, reproducible results.

In this assessment, it would appear that strategies to annotate serum glycoproteins is better suited at the peptide level than at the protein level. Others have reported that a significant percentage of tryptic N-glycopeptides is not in the preferred detection mass range of shotgun proteomics, that is, from 800 to 3500 Da. Potentially, the large size of glycan groups may block trypsin access resulting in glycopeptides not suitable to computational data-mining. Thus many glycoproteins may not be annotated effectively, if only trypsin is used to digest proteins[8]. Our results suggest the same as only a very few (about 10-14) glycoproteins were annotated prior to ConA peptide level enrichment in all cases. From our preliminary LC-MS operating parameters, approximately 150 additional proteins were annotated after glycopeptide enrichment/de-glycosylation, supporting a total annotation in the range of 350-580 protein identifications from the **AlbuVoid™** depletion method. Finally, we conclude that a series strategy of albumin depletion/low-abundance enrichment, followed by glycopeptide enrichment/de-glycosylation is an effective and efficient workflow strategy to annotate glycoproteins from serum.

Conclusions & Future Directions

The **NuGel™** surface chemistry has provided a platform for products with efficient proteomic workflows:

•Low abundance enrichment
•On-bead utility, enzyme digestion and assay
•No loss of functional or structural integrity, suitable for functional proteomics
•No requirement for specialized instruments or HPLC
•Simple bind/wash/elute protocols, adaptability to multiwell formats,
•Consumable, no potential for cross contamination, no regeneration.
•Substantially reduced cost, no bio-engineered ligands
•Tissue and species agnostic
•No bias to sample size or molecular weight
•Compatibility with quantitative LC-MS and iTRAQ methods

In the Rutgers Proteomics Center workflow, only N-linked glycoproteins were considered. Another **NuGel™** product, **NuGel™ PBA Kit**, features immobilized Phenyl-Boronic Acid. This chemical ligand has lectin-like character as it binds to *cis*-diols of carbohydrates. In principle, Phenyl-Boronic Acid will bind heterogeneous sets of both N-linked and O-linked glycoproteins and glycopeptides. We soon plan to test the coverage of the **NuGel™ PBA Kit** and to develop streamlined workflows of the **NuGel™** products to help annotate the serum proteome. Alternatively, enzymatic de-glysosylation prior to trypsin digestion may turn out to be a viable approach for basic sequence annotation and quantification of the serum proteome, and merits further investigation. Likewise, we anticipate the diversity of **NuGel™** proteome fractionation products, along with their unique features, will find utility serving all facets of proteomics.

Acknowledgment – We thank Dr. Kunjan Dave from Department of Neurology, University of Miami Miller School of Medicine for supplying the Rat serum.

References

1)www.biotechsupportgroup.com.
2)Philipp F Lange, Pitter F Huesgen, Karen Nguyen, and Christopher M Overall. *Annotating N termini for the Human Proteome Project: N termini and Nα-acetylation status differentiate stable cleaved protein species from degradation remnants in the human erythrocyte proteome*, J. Proteome Research., Just Accepted Manuscript • DOI: 10.1021/pr401191w • Publication Date (Web): 21 Feb 2014.
3)Fukuyama H, et al. *On-bead tryptic proteolysis: An attractive procedure for LC-MS/MS analysis of the Drosophilia caspase 8 protein complex during immune response against bacteria*. J of Proteomics. (2012)75:4610-4619
4)Labugger R, et al. *Strategy for analysis of cardiac troponins in biological samples with a combination of affinity chromatography and mass spectrometry*. Clin Chem (2003) 49(6 Pt1):873-9.
5)DeMiguel N, et al. *Proteome analysis of the surface of Trichomonas vagilis reveals novel protein and strain-dependent differential expression*. Mol Cell Proteomics (2010) 9(7):1554-66.
6)Swapan Roy, Ph.D., Matthew Kuruc, Krisha Patel, Suzanne Ackloo, Sven Nahnsen, Ph.D. *New On-Bead Digestion Protocols Improve LC-MS Workflows Of Albumin Depleted Samples.* Presented as a poster, US HUPO Conference March 11,12, 2013.
http://biotechsupportgroup.com/sites/default/files/New%20On-Bead%20Digestion%20Protocols%20Improve%20LC-MS%20Workflows%20Of%20Albumin%20Depleted%20Samples_0.pdf.
7)Hanne Kolsrud Hustoft, Helle Malerod, Steven Ray Wilson, Leon Reubsaet, Elsa Lundanes and Tyge Greibrokk (2012). *A Critical Review of Trypsin Digestion for LC-MS Based Proteomics, Integrative Proteomics*, Dr. Hon-Chiu Leung (Ed.), ISBN: 978-953-51-0070-6, InTech, Available from: http://www.intechopen.com/books/integrative-proteomics/a-critical-review-of-trypsin-digestion-for-lc-ms-based- proteomics.
8)Chen R. et al. *Glycoproteomics Analysis of Human Liver Tissue by Combination of Multiple Enzyme Digestion and Hydrazide Chemistry.*J Proteome Res., 2009, 8 (2), pp 651–661.