# Compound-centric Displacement Proteomics – An advantaged method to survey small molecule-protein interactions

Swapan Roy, Ph.D., Matthew Kuruc,
Biotech Support Group LLC, 1 Deer Park Dr., Suite M, Monmouth Junction, NJ 08852
Presented As A Poster At US HUPO Conference, 04/06/2014-04/09/2014

## Abstract

Functional proteomics relies in part, on the functional or structural features of intact, non-denatured proteins. As such, chemical and affinity-based proteomics can be considered a subset of functional proteomics. Regardless of the evolving terminology, the subject of chemical proteomics is to identify, characterize and quantify the binding interactions of small compounds to proteomes. The consideration of those interactions as functional modulators within the cell is paramount to understanding a potential therapeutic compound's mechanism of action. These same tools and methods also can help survey the promiscuous behavior of compounds towards multiple proteins and posit such behavior as deterministic of either toxicity or efficacy. We describe herein, new tools and methods for this purpose, called Compound-centric Displacement Proteomics (CCDP). Employing a new product - **NuGel™ *PRO*fessor™** which can non-covalently bind proteins, a subset of proteins can be displaced upon introduction of soluble small compounds. Coupled to LC-MS, quantitative metrics of these affinity-eluted sub-proteomes help characterize and identify interacting proteins. These new methods gain efficiencies over prior covalent-based substitution methods and can serve applications in drug target deconvolution, on-target/off-target specificity, and personalized medicine.

## Introduction

The healthcare industry relies on innovations and refinements in Systems Biology technologies, such as Gene Expression and Next Gen Sequencing, to accumulate knowledge surrounding the cellular constituents of different tissue phenotypes. As the vast majority of drugs work by modulating protein function, systems biology interests are vested in characterizing proteins as discretely as possible.

Consequently, the proteomics field continues to advance more efficient methodologies to sequence annotate subsets of proteomes. Known or suspect biomarker proteins can now be surveyed in targeted quantitative multiplexed analysis. Unique to proteomics is the additional capacity to directly assess structure, splice variants, post-translational modifications, and interactions with compounds, substrates and other regulating factors. This much deeper view of functional relevancy and phenotypic characterization cannot come from genomic information alone.

Notably, validated protein targets, definable by their amino acid sequence, are the foundation of modern drug discovery. Yet this foundation rests on an oversimplifying assumption; that is, one gene presupposes one protein sequence, which further presupposes one protein function. While this assumption works for some singular gene disorders, it is particularly flawed for complex multi-factorial diseases like cancer and neurodegenerative disorders. Biology is simply not so simple, as proteins with common sequence can serve many functions, and conversely proteins with dissimilar sequence can perform similar functions. For these reasons, the leap to functional mechanisms remains complicated. Therefore sequence annotation without correlating functional annotation, thus has many of the same limitations in Proteomics as in Genomics. Functional proteomic profiles derived from intact, non-denatured proteins, can thus bolster and deepen correlations between protein sequence, structure and function.

Many structural features participate in defining protein function including: post-translational modification (i.e., phosphate on/off), sub-unit equilibrium, allosteric regulating factors and the unstructured (or "spatially fluid") nature of protein domains. Unstructured domains interact with environmental stimuli, and drive conformational changes in structure – promoting a multitude of alternative functions[1,2]. So the subtleties of protein attributes, when the same or similar polypeptide sequence can have multiple conformations and functions, and when different sequences perform the same or similar functions, are now open to investigation.

Such functional annotation at the protein level is crucial to assessing mechanism of action of drug candidates, particularly since even clinically successful drugs are often promiscuous, modulating multiple proteins at once. The challenge 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. To support these challenges, Biotech Support Group has developed surface chemistries and separations strategies to untangle proteome complexity without compromising protein functional integrity. Here we present a new method for surveying the proteome complement of small compound interactions. One such product reported here is **NuGel™ *PRO*fessor™**, for chemical proteomic applications.

While the terminology can often overlap, chemical proteomics can be considered a subset of functional proteomics and encompasses a range of methods to derive profiles that can identify and classify structurally intact proteins that interact with small molecules of synthetic or natural origin3-9. In this way, functional and chemical proteomics can investigate the effects of a compound on its target and any collateral off-target functions. Such information is critical to safety and efficacy.

The field to date has primarily relied on two methods: activity-based protein profiling (ABPP) and compound-centric chemical proteomics (CCCP).

ABPP employs specially designed small molecule probes that covalently attach to active sites of their targeted protein family. Such probes require conserved active sites and detailed structural information and are thus restricted to pre-determined targets.

More broadly applied is compound-centric chemical proteomics (CCCP); an affinity chromatography method whereby the small molecule compound is covalently immobilized to a solid support, and incubated with a protein lysate to pull down the interacting proteins.

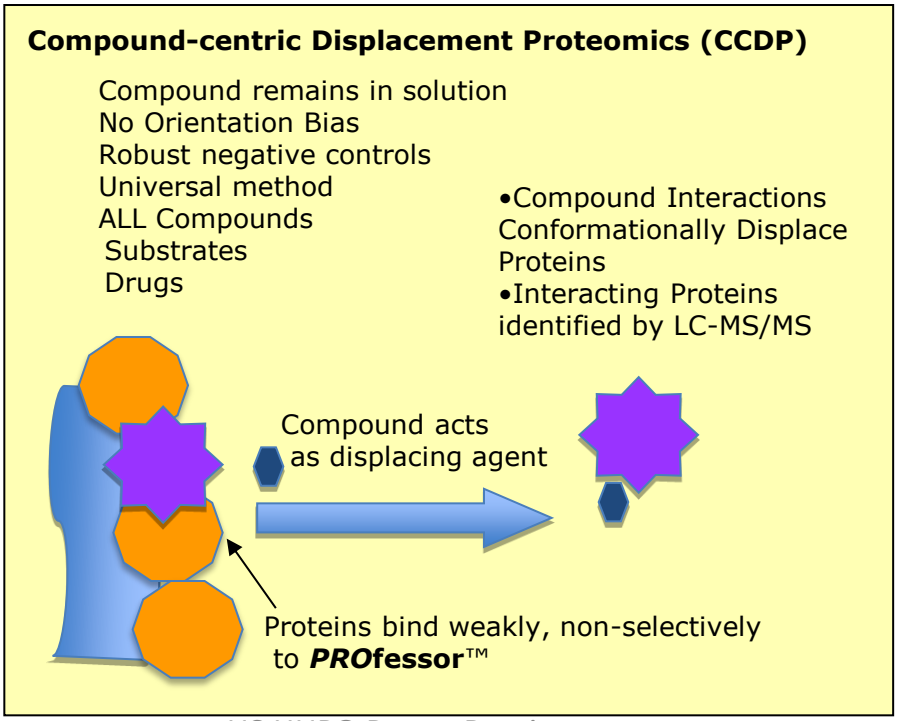
However, as therapeutic compounds exhibit their specificity due to their unique spatial orientation within the protein complex, covalent immobilization can compromise efficiency, leading to false specificity, high background binding and insufficient survey. Finally, compounds vary greatly in their immobilization efficiencies and this can distort the negative controls when comparing multiple compounds. The products and methods describe here circumvent many of these limitations.

**Technology Description**

As an alternative to conventional affinity chromatography, whereby the compound is first covalently immobilized, our method makes use of affinity elution, also known as 'biospecific elution', or 'substrate elution'. This process makes use of a small molecule's ability to conformationally alter the binding characteristics of proteins to a non-covalent general protein adsorbent.

We call our method, Compound-centric Displacement Proteomics or "CCDP". CCDP employs the **NuGel™ *PRO*fessor™** product – Figure below. Using this product, whole proteomes from cell lysates for example, can be first captured, then compound interactions can displace (or 'affinity elute') select proteins from the solid support. Such affinity-selected proteins can then be identified by LC-MS.

**NuGel™ *PRO*fessor™** is a composite mixture of mixed mode architectures designed to both compress protein concentrations from highly complex proteomes, and to non-covalently immobilize protein content with weak binding energy. It is supplied in a prep kit format, using simple bind/wash/elute microfuge protocols. One prep processes about 1 mg total protein, with final eluate volumes of 50 µls. As a one-time use consumable product, no specialized instruments or HPLC is required.



**Materials and Method**

The **NuGel™ *PRO*spector™** protocol is extremely simple and does not require any specialized instruments or HPLC. Simply, weigh the **NuGel™** powder, place in spin-filters, mix and centrifuge.

The protocol in detail can be found on the product insert, or company website (www.biotechsupportgroup.com). Briefly, 15 mg of dry powder **NuGel™ *PRO*spector™** material is added to a spin-filter. All steps continue with the spin-filter format using a microfuge for separations. 100 µls of clarified rat brain homogenate sample, 10 mg/ml, was conditioned 1:1 v/v by a binding buffer (0.1 M HEPES, pH 6), and applied to buffer treated **NuGel™ *PRO*spector™** material. The protein bound ***PRO*fessor™** material , was washed twice with 250 µl Wash buffer (0.001 M HEPES pH 7). For Compound-centric displacement experiments (on right), 'Ctrl' elution was the same as Wash buffer, displacement challenge compounds Imatinib or Caffeine-Benzoate were at 10 mM, in the same Wash buffer.

**Protein identification by LC-MSMS:**

LC-MSMS analysis was performed at the Biological mass spectrometry facility at RWJ Medical School and Rutgers (http://cabm-ms.cabm.rutgers.edu/) which supports client-sponsored research. Samples were run as a gel plug using a Novex gel Bis-Tris 10% gel. The entire band was excised and proteins in the gel were reduced, carboxymethylated, and digested with trysin using standard protocols. Peptides were extracted, solubilized in 0.1% trifluoroacetic acid, and analyzed by nanoLC-MS/MS using a RSLC system (Dionex, Sunnyvale CA) interfaced with a Velos-LTQ-Orbitrap (ThermoFisher, San Jose, CA). Samples were loaded onto a self-packed 100um x 2cm trap packed with Magic C18AQ, 5um 200 A (Michrom Bioresources Inc, Aubum, CA) and washed with Buffer A( 0.2% formic acid) for 5 min with flowrate of 10ul/min. The trap was brought in-line with the homemade analytical column (Magic C18AQ, 3um 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 15min, 15-25%B in 45min, and 25-55%B in 30 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 (acquired in linear ion trap) of 20 most intense ions with a repeat count of two and the dynamic exclusion duration of 60 sec.

The LC-MSMS data was searched against the *rat* complete protein database Rattus_norvegicus.Rnor_5.0.72. (ensembl.org) using a local version of the Global Proteome Machine (GPM cyclone, Beavis Informatics Ltd, Winnipeg, Canada) with carbamidoethyl on cysteine as fixed modification and oxidation of methionine and tryptophan as variable modifications using a 10 ppm precursor ion tolerance and a 0.4 Da fragment ion tolerance.

For protein confidence, we include only proteins that have more than 2 unique peptides and also loge<=-5. For peptide, FPR is automatically generated, valid log(e)<-0.4, ρ = 93, FPR=0.61%. To clarify, all cutoffs, filters or thresholds, are using all files. For example, each sample we ran 3 times, a total 6 runs. So any protein having at least 2 unique peptides identified (counting all 6 files), passed threshold.

Label-free peptides with corresponding gene/protein identifications were counted and tabulated. The most abundant 200 proteins are shown on right. A total of approximately 1200 proteins were annotated to some level between the three samples, indicating a broad level of proteome binding coverage by the ***PRO*fessor™** surface chemistries. These MS2 spectral counts serve as a relative quantification metric between samples. Some interpretation is required however, in evaluating the relative abundances in the different pools. We used a threshold under 25 as being not very quantitative, while anything close to 50 or more was considered quantitative at least with respect to comparisons of abundances between pools.

Table 1 shows a partial list of some significant abundance differences between the samples

**Table 1**
**MS2 Spectral Counts of Compound Displaced Subproteomes**

| Protein Description | Caffeine | Imatinib | Neg. Cont. |
|---|---|---|---|
| Hemoglobin subunit beta-1 | 87 | 550 | 53 |
| Glucose-6-phosphate isomerase | 192 | 459 | 76 |
| Malate dehydrogenase | 117 | 356 | 35 |
| transketolase | 72 | 160 | 24 |
| Cytochrome c, somatic | 47 | 123 | 3 |
| Succinyl-CoA:3-ketoacid transferase | 69 | 122 | 19 |
| Transgelin | 0 | 84 | 0 |
| Annexin A2 | 26 | 66 | 0 |
| fumarate hydratase | 17 | 42 | 2 |
| annexin A3 | 5 | 36 | 0 |
| glutathione reductase | 9 | 38 | 0 |

A partial list of LC-MS/MS identification and spectral counts demonstrate Imatinib interaction proteins from a common tissue homogenate, using CCDP. Caffeine was employed as a non-specific control compound, negative control was the final wash buffer.

Either label or label-free quantification methods will support the CCDP method. In addition to spectral counting, other label-free methods that can computationally monitor aligned MS1 spectral intensities (i.e., XIC) can be used. Label methods such as iTRAQ should be adaptable with consideration given to removal of any amine reactive species.

**Summary of Chemical Proteomics Methods**

**Activity-based Protein Profiling (ABPP)**
**small molecule probes**
**covalent attachment to active sites**
**requires conserved active sites & detailed structural information**

**Compound-centric chemical proteomics (CCCP)**
**covalent surface immobilization**
**affinity chromatography separations**
**Inefficiencies dominate when analyzing multiple compounds**

**Compound-centric displacement proteomics (CCDP)**
**NuGel *PRO*fessor™ - based method**
**No covalent substitutions**
**Consistent negative controls when analyzing multiple compounds**

## Results and Discussion

Normalization of background protein binding and robust negative controls are intrinsically challenging with prior compound immobilization methods. With CCDP, the background proteins are always the same, regardless of the challenge compound, so many different compounds can be run in parallel, with good quantitative metrics of comparison. Such is not the case when comparing different immobilized compounds, all derived from different covalent reactions and efficiencies.

By taking a composite of proprietary surface chemistries, we are able to weakly bind complex proteomes to the **NuGel™ *PRO*spector™** beads, and derive compound interaction sub-proteomes. The advantage of the CCDP method is that solution phase compounds are not subject to orientation biases. Negative controls are robust. Furthermore, it can be applied universally towards any compound or substrate, as it does not require covalent substitution or tag modifications. As compound concentrations can be varied, some measure of relative binding affinities between compounds can be characterized. Finally, when sufficient protein quantity is available, the compression of protein concentrations is possible, helping to minimize bias towards the high abundance protein content and uncover low abundance interactions that otherwise would have been hidden. No other chemical proteomics method collectively provides these benefits.

The interaction sub-proteomes are functionally and structurally intact and can be further analyzed in any suitable manner to support the goals of the investigation. LC-MS/MS sequence characterization of these sub-proteomes (the interactomes) can identify prospective biomarker proteins that bind with the challenging compound. Alternative splice variants or post-translational modifications can be identified. With these efficiencies gained, we envision that compound-centric displacement proteomics will support the annotation and correlation of drugs to tissue specific isoforms, the deconvolution of targets from high-throughput phenotypic screens, and the identification of functional biomarkers for personalized medicine.

## Conclusion & Future Directions

The proteomics field continues to advance more efficient methodologies to sequence annotate subsets of proteomes. Known or suspect biomarker proteins can now be surveyed in targeted quantitative multiplexed analysis. With chemical proteomics, there is now the additional capacity to annotate structure, splice variants, post-translational modifications, and correlate interactions with compounds, substrates and other regulating factors. This provides a much deeper view of functional relevancy and phenotypic characterization.

The collective advantages provided by the CCDP method are:

1.NuGel™ *PRO*fessor™ non-covalently binds proteins, a subset of which can be displaced upon introduction of soluble small compounds

2.Integrated with label or label-free quantitative LC-MS, quantitative metrics are applied to characterize and identify interacting proteomes

3.Efficiencies are gained over prior covalent-based substitution methods:

- more complete survey,
- no orientation bias,
- robust background controls for comparing multiple compounds at once

4.Applications in drug target deconvolution, on-target/off-target specificity, and personalized medicine

5.Proteins can be conformationally categorized to be:

- Frequently displaceable, those highly promiscuous to many compounds
- Assignable to specific challenge compounds
- Identifiable but with no other conformational categorization

| Gene ID | description | Imatinib | ctrl | caffeine |
|---|---|---|---|---|
| ENSRNOG00000049424 | Hemoglobin subunit beta-1 | 550 | 53 | 87 |
| ENSRNOG00000033190 | Gpi p, Glucose-6-phosphate isomer | 459 | 76 | 192 |
| ENSRNOG00000004291 | Cltc p, clathrin heavy chain 1 | 81 | 251 | 203 |
| ENSRNOG00000034254 | Actb p, Actin, cytoplasmic 1 Actin, | 108 | 224 | 87 |
| ENSRNOG00000004559 | RGD1560000 p, Tubulin alpha-1A c | 94 | 224 | 204 |
| ENSRNOG00000017895 | RGD1562394 p, Alpha-enolase | 172 | 152 | 174 |
| ENSRNOG00000011841 | microtubule-associated protein 2 | 71 | 195 | 154 |
| ENSRNOG00000029886 | Hba-a2 p, Hemoglobin subunit alph | 417 | 37 | 176 |
| ENSRNOG00000016396 | Got1 p, Aspartate aminotransferase | 376 | 81 | 131 |
| ENSRNOG00000009625 | Dpys l2 p, Dihydropyrimidinase-rela | 117 | 190 | 57 |
| ENSRNOG00000013009 | Ldha p, L-lactate dehydrogenase A | 323 | 70 | 137 |
| ENSRNOG00000011329 | Pyruvate kinase isozymes M1/M2 | 158 | 109 | 142 |
| ENSRNOG00000023647 | Aldoc p, Fructose-bisphosphate ald | 188 | 96 | 107 |
| ENSRNOG00000030625 | Ttr p, Serotransferrin | 190 | 95 | 219 |
| ENSRNOG00000018630 | RGD1560581 p, Glyceraldehyde-3- | 77 | 169 | 61 |
| ENSRNOG00000017558 | Tubb2c p, Tubulin beta-2A chain | 30 | 43 | 96 |
| ENSRNOG00000001440 | Mdh2 p, Malate dehydrogenase, mi | 356 | 35 | 117 |
| ENSRNOG00000033835 | Dynamin-1 | 46 | 124 | 42 |
| ENSRNOG00000010170 | Tubb3 c, tubulin beta-4B chain | 137 | 209 | 41 |
| ENSRNOG00000029211 | Alb p, Serum albumin | 43 | 95 | 113 |
| ENSRNOG00000007219 | Hsp90aa1 p, Heat shock protein HS | 29 | 77 | 134 |
| ENSRNOG00000019834 | Hsp90ab1 p, Heat shock protein HS | 61 | 130 | 71 |
| ENSRNOG00000013000 | Ldhb p, L-lactate dehydrogenase B | 93 | 89 | 94 |
| ENSRNOG00000011782 | Got2 p, Aspartate aminotransferase | 152 | 42 | |
| ENSRNOG00000034093 | no protein information available | | | 89 |
| | no protein text annotation available | 22 | 105 | 74 |
| ENSRNOG00000010872 | Ckb p, Creatine kinase B-type | 37 | 76 | 78 |
| ENSRNOG00000002840 | Atp5b p, ATP synthase subunit beta | 43 | 78 | 77 |
| ENSRNOG00000017163 | Pfkp p, 6-phosphofructokinase type | 29 | 90 | 77 |
| ENSRNOG00000017428 | microtubule-associated protein 1B | 24 | 85 | 41 |
| ENSRNOG00000001992 | Atp6v1a p, V-type proton ATPase c | 24 | 74 | 80 |
| ENSRNOG00000009628 | Pfkm p, 6-phosphofructokinase, m | 26 | 76 | 63 |
| ENSRNOG00000010222 | Glud1 p, Glutamate dehydrogenase | 34 | 79 | 81 |
| ENSRNOG00000007583 | Pygb p, glycogen phosphorylase, b | 57 | 88 | 68 |
| ENSRNOG00000016064 | Tkt p, transketolase | 160 | 24 | 72 |
| ENSRNOG00000024128 | Aco2 p, Aconitate hydratase, mitoc | 44 | 75 | 76 |
| ENSRNOG00000034242 | Vcp p, Transitional endoplasmic ret | 17 | 82 | 57 |
| ENSRNOG00000011452 | Aldoc p, Fructose-bisphosphate ald | 48 | 64 | |
| ENSRNOG00000014525 | Hspd1 p, | | | 66 |
| | Hspd1 p, 60 kDa heat shock protei | 34 | 76 | 77 |
| ENSRNOG00000044438 | Cnn1 p, Contactin-1 | 29 | 63 | 39 |
| ENSRNOG00000049937 | Pdxk p, pyridoxal kinase | 46 | 66 | 52 |
| | no protein text annotation available | 38 | 65 | 67 |
| ENSRNOG00000019889 | Pgm1 p, phosphoglucomutase-1 | 54 | 62 | |
| ENSRNOG00000064423 | Ppia p, Peptidyl-prolyl cis-trans isor | 117 | 12 | 64 |
| ENSRNOG00000048365 | no protein information available | | | 66 |
| | no protein text annotation available | 9 | 67 | 51 |
| ENSRNOG00000043094 | Oxct1 p, Succinyl-CoA:3-ketoacid c | 122 | 19 | 69 |
| ENSRNOG00000011891 | Atp6v1b2 p, V-type proton ATPase | 15 | 52 | 82 |
| ENSRNOG00000024467 | Pgk1 p, Phosphoglycerate kinase 1 | 41 | 59 | |
| ENSRNOG00000059585 | no protein information available | | | 52 |
| | Phosphoglycerate mutase 1 | 41 | 31 | 48 |
| ENSRNOG00000043205 | Ig gamma-2A chain C region | 148 | 11 | 55 |
| ENSRNOG00000005290 | Ywhae p, 14-3-3 protein epsilon | 25 | 60 | 53 |
| ENSRNOG00000018282 | Gda p, guanine deaminase | 14 | 64 | 55 |
| ENSRNOG00000009988 | Ap2b1 p, AP-2 complex subunit bet | 31 | 60 | 28 |
| ENSRNOG00000024468 | Tnr p, tenascin R | 15 | 52 | 58 |
| ENSRNOG00000010365 | Syn1 p, Synapsin-1 | 50 | 31 | 24 |
| ENSRNOG00000013905 | Nsf p, vesicle-fusing ATPase | 14 | 64 | 59 |
| ENSRNOG00000014230 | Map1a p, | | | 50 |
| | Map1a p, microtubule-associated p | 9 | 65 | 48 |
| ENSRNOG00000011282 | Ppp2r1a p, serine/threonine-protei | 16 | 56 | 48 |
| ENSRNOG00000010277 | Idh3a p, isocitrate dehydrogenase | 24 | 50 | 52 |
| ENSRNOG00000015396 | spectrin alpha chain, non-erythroc | 4 | 76 | |
| ENSRNOG00000037245 | Gdi1 p, Rab GDP dissociation inhibi | 8 | 63 | 50 |
| ENSRNOG00000008195 | 14-3-3 protein zeta/delta | 25 | 45 | |
| ENSRNOG00000019534 | Ap2a2 p, | | | 39 |
| | Ap2a2 p, AP-2 complex subunit alp | 10 | 62 | 37 |
| | Crmp1 p, Dihydropyrimidinase-rela | 38 | 49 | 37 |
| ENSRNOG00000008103 | Mdh1 p, Malate dehydrogenase, cy | 26 | 44 | 30 |
| ENSRNOG00000006709 | Rpp p, Alpha-1-macroglobulin Alpha | 10 | 40 | 13 |
| ENSRNOG00000035133 | Mapt p, microtubule-associated pr | 2 | 2 | 49 |
| | microtubule-associated protein tau | 40 | 43 | |
| ENSRNOG00000018798 | brevican core protein isoform 1 pre | 6 | 47 | 46 |
| ENSRNOG00000018629 | Car2 p, Carbonic anhydrase 2 | 44 | 5 | 39 |
| ENSRNOG00000022343 | Uchl1 p, Ubiquitin carboxyl-termin | 22 | 45 | 2 |
| ENSRNOG00000022636 | Abat p, 4-aminobutyrate aminotran | 19 | 51 | 35 |
| ENSRNOG00000049560 | Glutamine synthetase | 19 | 50 | 35 |
| | no protein information available | | | 51 |
| ENSRNOG00000031230 | LOC689064 p, Hemoglobin subunit | 103 | 12 | 57 |
| ENSRNOG00000007862 | Acat1 p, Acetyl-CoA acetyltransfer | 37 | 40 | 44 |
| ENSRNOG00000009882 | no protein information available | | | 49 |
| | Serine/threonine-protein phosphat | 17 | 40 | 42 |
| ENSRNOG00000017032 | Atp5a1 p, ATP synthase subunit alp | 22 | 40 | 37 |
| ENSRNOG00000017209 | Tubb3 p, Tubulin beta-3 chain | 27 | 39 | |
| ENSRNOG00000007895 | Pdhb p, Pyruvate dehydrogenase E | 14 | 44 | 36 |
| ENSRNOG00000045636 | Fatty acid synthase S-acetyltransfe | 3 | 53 | 45 |
| | no protein information available | | | 17 |
| ENSRNOG00000018289 | RGD1560648 p, Protein ED-1 | 48 | 23 | 36 |
| ENSRNOG00000020266 | Eef2 p, Elongation factor 2 | 12 | 54 | 39 |
| ENSRNOG00000024698 | Wdr1 p, WD repeat-containing prot | 11 | 43 | 35 |
| ENSRNOG00000018937 | Gstm7 p, Glutathione S-transferase | 34 | 20 | 35 |
| ENSRNOG00000008996 | Dpysl5 p, Dihydropyrimidinase-rela | 47 | 33 | 31 |
| ENSRNOG00000010482 | Impa1 p, inositol monophosphatase | 12 | 49 | |
| ENSRNOG00000010452 | Cycs p, Cytochrome c, somatic | 123 | 3 | 47 |
| ENSRNOG00000011438 | Ywhag p, 14-3-3 protein gamma 1 | 20 | 36 | 17 |
| ENSRNOG00000009982 | Npy p, | 29 | 41 | 32 |
| ENSRNOG00000011751 | Ncdn p, Neurochondrin | 11 | 48 | 35 |
| ENSRNOG00000019164 | Uba1 p, Ubiquitin-like modifier-acti | 7 | 47 | 44 |
| ENSRNOG00000007834 | Cand1 p, Cullin-associated NEDD8- | 6 | 50 | 1 |
| ENSRNOG00000012477 | Eef1a2 p, Elongation factor 1-alpha | 11 | 17 | 41 |
| ENSRNOG00000044402 | Ig gamma-2B chain C region | 111 | 4 | 23 |
| ENSRNOG00000039726 | Gstm1 p, | | | 35 |
| | Gstm1 p, glutathione S-transferase | 53 | 19 | 40 |
| ENSRNOG00000009439 | Eef1a1 p, Elongation factor 1-alpha | 25 | 48 | |
| ENSRNOG00000016724 | band 4.1-like protein 3 | | 2 | 28 |
| | Epb41l3 p, band 4.1-like protein 3 | 11 | 48 | |
| ENSRNOG00000014745 | no protein information available | | | 12 |
| | no protein text annotation available | 10 | 36 | |
| ENSRNOG00000023095 | Npepps p, puromycin-sensitive ami | 5 | 46 | 23 |
| ENSRNOG00000013520 | Prdx2 p, Peroxiredoxin-2 | 16 | 36 | 29 |
| ENSRNOG00000018992 | Dihydropyrimidinase-related protei | 11 | 36 | |
| | no protein information available | | | 34 |
| ENSRNOG00000002051 | no protein information available | | | 30 |
| | Synaptojanin-1 | 4 | 45 | 5 |
| ENSRNOG00000033155 | Aif3 p, AIP--dependent protein factor 3 | 18 | 28 | 31 |
| ENSRNOG00000012490 | Amph p, amphiphysin | 5 | 35 | 15 |
| ENSRNOG00000014160 | Tcp1 p, T-complex protein 1 subuni | 6 | 46 | 40 |
| ENSRNOG00000005364 | Dld p, Dihydrolipoyl dehydrogenase | 42 | 22 | 21 |
| ENSRNOG00000018813 | Dynamin-like protein | 7 | 32 | 34 |
| ENSRNOG00000018237 | Gstp1 p, Glutathione S-transferase | 49 | 20 | 29 |
| ENSRNOG00000010580 | Acot7 p, Cytosolic acyl coenzyme A | 21 | 28 | 29 |
| ENSRNOG00000021125 | Prdx5 p, Peroxiredoxin-5, mitochon | 34 | 19 | 31 |
| ENSRNOG00000049829 | no protein information available | 80 | 7 | 24 |
| ENSRNOG00000046654 | no protein information available | 18 | 28 | 25 |
| ENSRNOG00000013971 | Psat1 p, phosphoserine aminotrans | 5 | 40 | 26 |
| ENSRNOG00000013979 | Pfn1 p, Profilin-1 | 39 | 8 | 29 |
| ENSRNOG00000015987 | Rap1gds1 p, rap1 GTPase-GDP diss | 11 | 36 | 33 |
| ENSRNOG00000027582 | Dpysl4 p, Dihydropyrimidinase-rela | 23 | 36 | 24 |
| ENSRNOG00000026243 | Ap2a1 p, AP-2 complex subunit alp | 7 | 43 | 27 |
| ENSRNOG00000002671 | Nme2 p, Nucleoside diphosphate ki | 46 | 30 | 37 |
| ENSRNOG00000020660 | Cfl1 p, Cofilin-1 | 13 | 28 | 4 |
| ENSRNOG00000006178 | Dync1h1 p, cytoplasmic dynein 1 h | 1 | 65 | |
| ENSRNOG00000025083 | protein kinase C and casein kinase | 4 | 23 | 34 |
| ENSRNOG00000001592 | Cct8 p, T-complex protein 1 subuni | 3 | 39 | 26 |
| ENSRNOG00000013494 | Gsta1 p, Glutathione S-transferase | 61 | 5 | 30 |
| ENSRNOG00000023861 | clathrin coat assembly protein AP1I | 7 | 43 | 38 |
| | no protein information available | | | |
| ENSRNOG00000017311 | NADP-dependent malic enzyme, mi | 7 | 22 | 36 |
| ENSRNOG00000013492 | Cap1 p, Adenylyl cyclase-associate | 10 | 28 | 29 |
| ENSRNOG00000009513 | Akr1b1 p, Aldose reductase | 3 | 40 | 22 |
| ENSRNOG00000047565 | no protein information available | | | 19 |
| | no protein text annotation available | 76 | 4 | 26 |
| ENSRNOG00000016596 | Hspa4 p, heat shock 70 kDa protei | 1 | 34 | |
| ENSRNOG00000010176 | Mapk2k1 p, Dual specificity mitogen | 8 | 27 | 26 |
| ENSRNOG00000017194 | Prdx1 p, Peroxiredoxin-1 | 36 | 22 | 22 |
| ENSRNOG00000009923 | Cct6a p, T-complex protein 1 subu | 9 | 33 | 25 |
| ENSRNOG00000013177 | Cct2 p, T-complex protein 1 subuni | 9 | 35 | 20 |
| ENSRNOG00000020075 | Eef1g p, Elongation factor 1-gamm | 4 | 31 | 34 |
| ENSRNOG00000008656 | Snca p, Alpha-synuclein | 27 | 14 | 29 |
| ENSRNOG00000011136 | Pebp1 p, Phosphatidylethanolamin | 14 | 18 | 9 |
| ENSRNOG00000049075 | no protein information available | 40 | 4 | 34 |
| ENSRNOG00000037188 | murinoglobulin-1 precursor | | 13 | 39 |
| | no protein information available | | | 14 |
| ENSRNOG00000015420 | Stxbp1 p, Syntaxin-binding protein | 12 | 24 | 22 |
| ENSRNOG00000016924 | Acly p, | | | 26 |
| | Acly p, ATP-citrate synthase isoform | 3 | 33 | 16 |
| ENSRNOG00000010945 | Ywhab p, 14-3-3 protein beta/alpha | 9 | 29 | 21 |
| ENSRNOG00000016727 | Akr1a1 p, Alcohol dehydrogenase | 12 | 24 | 16 |
| ENSRNOG00000009994 | Dlat p, Dihydrolipoyllysine-residue | 10 | 28 | |
| ENSRNOG00000008157 | synaptic 2 isoform 2 | 13 | 27 | 30 |
| ENSRNOG00000003597 | Tubb4a p, Tubulin alpha-4A chain | 14 | 25 | 27 |
| ENSRNOG00000036688 | Arhgdia p, Rho GDP-dissociation in | 5 | 22 | 30 |
| ENSRNOG00000017777 | Ahcy p, Adenosylhomocysteinase | 30 | 13 | 28 |
| ENSRNOG00000015511 | RGD1559750 p, Bifunctional purine | 4 | 28 | 26 |
| ENSRNOG00000011632 | Cct5 p, T-complex protein 1 subuni | 3 | 40 | 15 |
| ENSRNOG00000026119 | Ywhah p, 14-3-3 protein eta | 11 | 26 | |
| ENSRNOG00000015048 | Sod2 p, Superoxide dismutase, mi | 49 | 6 | |
| ENSRNOG00000019433 | Rab3a p, Ras-related protein Rab-3 | 14 | 22 | 8 |
| ENSRNOG00000009902 | Hspb1 p, Heat shock protein 105 kDa | | 33 | 15 |
| ENSRNOG00000021090 | Pygm p, glycogen phosphorylase, r | 3 | 23 | 25 |
| ENSRNOG00000015409 | Uqcrc1 p, ubiquitin carboxyl-termina | 8 | 28 | 30 |
| ENSRNOG00000019875 | Matr3 p, Matrin-3 | 5 | 26 | 20 |
| ENSRNOG00000017730 | Psmd1 p, 26S proteasome non-ATP | 1 | 36 | 19 |
| ENSRNOG00000008570 | Cadps p, | | | 25 |
| | Cadps p, calcium-dependent secret | 1 | 34 | 15 |
| ENSRNOG00000032002 | Hapln1 p, Hyaluronan and proteogl | 4 | 24 | 28 |
| ENSRNOG00000018019 | Hspa1a p, heat shock 70 kDa prot | 9 | 29 | 25 |
| ENSRNOG00000018536 | Actc1 p, Actin, alpha cardiac muscl | 9 | 29 | 20 |
| ENSRNOG00000033253 | Qdpr p, Dihydropteridine reductase | 32 | 15 | 20 |
| ENSRNOG00000023536 | Aldh5a1 p, Succinate-semialdehyd | 10 | 25 | 25 |
| ENSRNOG00000013520 | Cs p, citrate synthase, mitochondri | 10 | 14 | 19 |
| ENSRNOG00000015599 | alpha-1-inhibitor 3 precursor | 1 | 52 | 11 |
| ENSRNOG00000022693 | Nme1 p, Nucleoside diphosphate ki | 7 | 8 | 20 |
| ENSRNOG00000019372 | Pc p, Pyruvate carboxylase, mitoch | 2 | 31 | 16 |
| ENSRNOG00000019328 | Phgdh p, D-3-phosphoglycerate de | 7 | 24 | 15 |
| ENSRNOG00000017519 | Psmd2 p, 26S proteasome non-ATP | 1 | 21 | 21 |
| ENSRNOG00000018404 | Aars p, Alanine--tRNA ligase, cytop | 4 | 36 | 21 |
| ENSRNOG00000018367 | Taldo1 p, Transaldolase | 38 | 5 | 23 |
| ENSRNOG00000002378 | Bpnt1 p, 3 (2 ),5 -bisphosphate nu | 5 | 18 | |
| ENSRNOG00000009275 | Kpnb1 p, importin subunit beta-1 | 3 | 28 | 10 |
| ENSRNOG00000012996 | Prdx6 p, Peroxiredoxin-6 | 1 | 28 | 17 |
| ENSRNOG00000047098 | Hbb-b1 p, hemoglobin, beta adult s | 36 | 6 | 26 |
| ENSRNOG00000015630 | Cct7 p, T-complex protein 1 subuni | 5 | 31 | 24 |
| ENSRNOG00000010362 | Anxa2 p, Annexin A2 | 66 | | 26 |
| ENSRNOG00000011849 | Mapk1 p, Mitogen-activated protei | 8 | 19 | |
| ENSRNOG00000046889 | no protein information available | 34 | 1 | 23 |
| ENSRNOG00000001709 | AP-2 complex subunit mu | 1 | 28 | 9 |
| | no protein information available | | | 15 |
| ENSRNOG00000030317 | LOC100360180 p, Corticotatin Corti | 21 | 14 | 19 |
| ENSRNOG00000019113 | Heterogeneous nuclear ribonucleoprotein K | | | 19 |

Acknowledgement: We thank Dr. Haiyan Zheng at the Biological mass spectrometry facility at RWJ Medical School and Rutgers, for directing the LC-MSMS analysis and her guidance with reporting and interpreting the data.

**References**

1.Radivoiac P. et al. Intrinsic Disorder and Functional Proteomics. *Biophysical Journal* Volume 92, 1439–1456 (2007).
2.Buyong Ma and Ruth Nussinov. Minireview: Regulating highly dynamic unstructured proteins and their coding mRNAs. Genome Biology 10:204 (2009). doi:10.1186/gb-2009-10-1-204.
3.Gyenis L., et al. Chemical proteomics and functional proteomics strategies for protein kinase inhibitor validation and protein kinase substrate identification: Applications to protein kinase CK2. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. Volume 1834, Issue 7, 1352-1358 (2013).
4.Han S., et al. A blue native polyacrylamide gel electrophoretic technology to probe the functional proteomics mediating nitrogen homeostasis in *Pseudomonas fluorescens. J Microbio. Methods.* Vol. 90, Issue 3. 206-210 (2012).
5.Johanna Graßmann, Romy K. Sheerle and Thomas Letzel. Functional proteomics: Application of Mass Spectrometry to the study of enzymology in complex mixtures. *Anal. Bioanal. Chem*. Volume 402, Issue 2, pp 625-645 (2012).
6.Jeffrey DA, Bogyo M. Chemical proteomics and its application to drug discovery. *Curr. Opin. Biot.* 14, 87-95 (2003).
7.Oda Y et al. Quantitative chemical proteomics for identifying candidate drug targets. *Anal. Chem.* 75, 2159-2165 (2003).
8.Piggott AM, Karuso P. Quality, not quantity: the role of natural products and chemical proteomics in modern drug discovery. *Comb. Chem. High. Throughput Screen* 7,607- 630 (2004).
9.Heal WP, et al. Activity-based probes: discovering new biology and new drug targets. *Chem. Soc. Rev*. 10, 246-257 (2011).