Annual Report for DOE

Title: Molecular Assemblies, Genes, and Genomics Integrated Efficiently (MAGGIE)
DOE #: DE-FG02-07ER64325
TSRI #: 5-75538
PI: Gary Siuzdak, Ph.D.

MAGGIE: Component 4 - Mass Spectrometry

Research Objectives:

The objectives of MAGGIE include the provision of robust GTL technologies and comprehensive characterization to elucidate functional relationships and pathways. Toward this goal Component 4 has generated unique technologies as well as helped identified new biological interactions. The informatics and analytical technology platforms that we've developed as well as the biochemistry that it has been developed for, are presented below:

Research Progress and Implications:

As of year 4 of 4 year project, the following work has been performed.

Development of bioinformatic Platforms (Anal. Chem. 2006, 2008 & 2010). The success of bioinformatic platforms is based on their ability to identify, quantify and statistically assess features of interest. XCMS (***XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification, Smith et al. 2006***) provides a comprehensive software package for untargeted metabolomics. Recently, new algorithms for feature detection and alignment were incorporated into XCMS (**Highly sensitive feature detection for high resolution LC/MS, Tautenhahn et al. 2008**). This has also been developed for meta-analysis in a new program called ***metaXCMS*** (**Anal. Chem. 2010**). Furthermore, algorithms for peak detection and comparison of FT/ICR spectra were integrated into XCMS. The next step in data analysis is the structural characterization of metabolites representing these features. To accomplish this, MS/MS experiments are used to ascertain structural information about the molecule. $XCMS^2$ (***XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization, Benton et al. 2008***) is an extension of XCMS which facilitates the import, processing and automatic searching of MS/MS spectra against the METLIN database. Included in this platform is a "similarity search" functionality that analyzes MS/MS patterns for structural classification of unknown molecules. Current developments focus on the structural elucidation of unknown metabolites. A graphical user interface was developed, that allows for an easy and intuitive search in the METLIN database, using the combined information from MS/MS fragments, neutral losses, and the exact mass.

Novel multiprotein and metal protein complexes identified in the hyperthermophilic archaeon Pyrococcus furiosus by non-denaturing fractionation of the native proteome (Nature 2010, BMC Systems Biology 2010, BMC Bioinformatics 2011, Mol. Cell. Proteomics 2009). In this collaborative work with Component 1, a high throughput nano- LC-ESI-MS/MS platform developed by Component 4 was used to identify 106 potential heteromeric protein complexes containing 243 individual proteins as well as metal containing proteins. We attempted to directly determine the identity of PCs from native microbial biomass using Pyrococcus furiosus, a hyperthermophilic archaeon that grows optimally at 100 degrees C, as the model organism. Novel PCs were purified by large scale fractionation of the native proteome using non-denaturing, sequential column chromatography under anaerobic, reducing conditions.

Identification of a New Endogenous Metabolite and Characterization of Metabolite-Protein Interactions (J. Am. Chem Soc. 2009). The emerging field of global mass-based metabolomics provides a platform for discovering unknown metabolites and their specific biochemical pathways. We report the identification of a new endogenous metabolite, N(4)-(N-acetylaminopropyl)spermidine and the use of a novel proteomics based

method for the investigation of its protein interaction using metabolite immobilization on agarose beads. The metabolite was isolated from the organism Pyrococcus furiosus, and structurally characterized through an iterative process of synthesizing candidate molecules and comparative analysis using accurate mass LC-MS/MS. The biochemical role of the novel metabolite as well as that of two other polyamines: spermidine and agmatine is investigated through metabolite immobilization and incubation with native proteins. The identification of eleven proteins that uniquely bind with N(4)-(N-acetylaminopropyl)spermidine, provides information on the role of this novel metabolite in the native organism. The immobilization approach demonstrated here has the potential for application to other newly discovered endogenous metabolites found through untargeted metabolomics, as a preliminary screen for generating a list of proteins that could be further investigated for specific activity.

Phosphonium labeling for increasing metabolomic coverage of neutral lipids using electrospray ionization mass spectrometry (Rapid Commun Mass Spectrom. 2009). Mass spectrometry has become an indispensable tool for the global study of metabolites (metabolomics), primarily using electrospray ionization mass spectrometry (ESI-MS). However, many important classes of molecules such as neutral lipids do not ionize well by ESI and go undetected. Chemical derivatization of metabolites can enhance ionization for increased sensitivity and metabolomic coverage. Here we describe the use of tris(2,4,6,-trimethoxyphenyl)phosphonium acetic acid (TMPP-AA) to improve liquid chromatography (LC)/ESI-MS detection of hydroxylated metabolites (i.e. lipids) from serum extracts. Cholesterol which is not normally detected from serum using ESI is observed with attomole sensitivity. This approach was applied to identify four endogenous lipids (hexadecanoyl-sn-glycerol, dihydrotachysterol, octadecanol, and alpha-tocopherol) from human serum. Overall, this approach extends the types of metabolites which can be detected using standard ESI-MS instrumentation and demonstrates the potential for targeted metabolomics analysis.

Mass Spectrometry Technology (*Nature 2007, PNAS 2008, and Nature Prot. 2008, Anal. Chem. 2010 & 2011*). The interest in leveraging mass spectrometry for studying enzyme activities in complex biological samples derives from its high sensitivity and specificity. However, signal suppression and significant sample preparation limits its overall utility. We introduced Nanostructure-Initiator Mass Spectrometry (NIMS) as a new approach for mass spectrometry analysis. And have since developed Nanostructure-Initiator Mass Spectrometry enzymatic (Nimzyme) assay which uses the fluorous liquid-coated surface of Nanostructure-Initiator Mass Spectrometry (NIMS) to noncovalently attach enzyme substrates through fluorous tags. This ‘soft’ immobilization allows efficient desorption/ionization while also enabling the use of surface washing steps to reducing signal suppression from complex biological samples as a result of the preferential retention of the tagged products and reactants.

Comprehensive Ionization (*Analytical Chemistry 2008 and Spectroscopy 2008*). Separate LC-MS analyses in positive and negative ionization mode were performed on methanol extracts using, ESI, APCI and multimode (MM) which is simultaneous ESI and APCI ionization. The fundamental observation that was obtained for each of the investigated ionization methods was that >90% of the ions detected in a particular polarity mode were found unique to that particular polarity when compared to the opposite polarity of the same ionization method. As the entire dataset was evaluated, three categories of ions emerged. **(1)** Ions with corresponding ions detected in opposite polarity. **(2)** Ion-features for which different in-source fragmentation cannot be excluded, and hence, potentially only unique in a molecular ion sense. **(3)** Ion-features that are truly unique to a particular polarity.

Correlating the Transcriptome, Proteome and Metabolome (*J. Proteome Research 2008*) Here we present a comprehensive platform for the transcriptomic, proteomic and metabolite profiling using gene array measurements and mass spectrometry that not only allows the identification of a novel metabolite N^1-acetylthermospermine, but also the characterization of transcript, protein and metabolite level changes. Finally, our novel metabolite immobilization and proteomics approach demonstrated here with *P. furiosus,* allowed us to discover protein candidates that may play a part in the regulation of identified endogenous metabolites and have a specific interaction with metabolites of interest. The combination of these technologies have been used to monitor global molecular changes in response to environment stress, and for mapping networks that define connectivity in function between gene expression, proteins and metabolites for the model organism, *Pyrococcus furious*.

Providing Proteomics Support to Components 1, 2 and 3
High throughput proteomics methods using robotics and nano-LC-MS/MS have been developed to facilitate automated analysis of MAGGIE proteomic samples provided by the above components. These methods have been used to identify protein complexes in over 9300 samples to date.

Information Access

Patti G.J, Tautenhahn R., Fonslow B.R., Cho Y., Deutschbauer A., Arkin A., Northen T., Siuzdak G.
Meta-Analysis of Global Metabolomics and Proteomics Data to Link Alterations with Phenotype
Spectroscopy (cover article), **2011**, 26(3), 151-154.

Tautenhahn R., Patti G.J., Kalisiak E., Miyamoto T., Schmidt M., Fan Y.L., McBee J., Baliga N.S., Siuzdak G.
metaXCMS: Second-Order Analysis of Untargeted Metabolomics Data
Analytical Chemistry, **2011**, 83, 696-700.

Yanes O., Tautenhahn R., Patti G., Siuzdak G.
Expanding Coverage of the Metabolome for Global Metabolite Profiling
Analytical Chemistry, **2011**, 83(6), 2152-2161.

Greving M.P., Patti G.J., Siuzdak G.
Nanostructure-Initiator Mass Spectrometry (NIMS) Metabolite Analysis and Imaging
Analytical Chemistry, **2011**, 83(1), 2-7.

Powers M.E., Smith P.A., Roberts T.C., Fowler B.J., King C.C., Trauger S.A., Siuzdak G., Romesberg F.E.
Type I Signal Peptidase and Protein Secretion in *Staphylococcus epidermidis*
Journal of Bacteriology, **2011**, 193(2), 340-348

Long J.Z., Cisar J.S., Miliken D., Niessen S., Wang C., Trauger S.A., Siuzdak G., Cravatt B.F.
Metabolomics annotates ABHD3 as a physiologic regulator of medium-chain phospholipids
Nature Chemical Biology, **2011**, 7, 763-765.

Lancaster WA, Praissman JL, Poole FL II, Cvetkovic A, Menon AL, Scott JW, Jenney FE Jr, Thorgersen MP, Kalisiak E, Apon JA, Trauger SA, Siuzdak G, Tainer JA, Adams MWW
A Computational Framework for Proteome-Wide Pursuit and Prediction of Metalloproteins using ICP-MS and MS/MS Data
BMC Bioinformatics, **2011**, 12, 1-12.

Patti G.J., Woo H.K., Yanes O., Shriver L., Thomas D., Uritboonthai W., Apon J.V., Steenwyk R., Manchester M., Siuzdak G.
Detection of Carbohydrates and Steroids by Cation-Enhanced Nanostructure-Initiator Mass Spectrometry (NIMS) for Biofluid Analysis and Tissue Imaging
Analytical Chemistry, **2010**, 82(1), 121-128

Facciotti M.T., Pang W.L., Lo F-Y., Whitehead K., Koide T., Masumura K., Pan M., Kaur A., Larsen D.J., Reiss D.J., Hoang L., Kalisiak E., Northen T., Trauger S.A., Siuzdak G., Baliga N.S.
Large scale physiological readjustment during growth enables rapid, comprehensive and inexpensive system analysis
BMC Systems Biology, **2010**, 4, 64

Cvetkovic A., Menon A.L., Thorgersen M.P., Scott J.W., Poole F.L. II, Jenney F.E. Jr., Lancaster W.A., Praissman J.L., Shanmukh S., Vaccaro B.J., Trauger S.A., Kalisiak E., Apon J.V., Siuzdak G., Yannone S.M., Tainer J.A., Adams M.W.W.
Microbial metalloproteomes are largely uncharacterized
Nature, **2010**, 466, 779-784

Lowe R., Szili E., Kirkbride P., Thissen H., Siuzdak G., Voelcker N.
Combined Immunocapture and Laser Desorption/Ionization Mass Spectrometry on Porous Silicon
Analytical Chemistry, **2010**, 82(10), 4201-8

Wikoff W.R., Anfora A.T., Liu J.W., Shultz P.G., Lesley S.A., Peters E.C., Siuzdak G.
Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites
Proc Natl Sci USA, **2009**, 106(10), 3698-703

Wikoff W.R., Kalisiak E., Trauger S., Manchester M., Siuzdak G.
Response and recovery in the plasma metabolome tracks the acute LCMV-induced immune response
J Proteome Res, **2009**, 8(7), 3578-87

Menon A.L., Poole F.L., Cvetkovic A., Trauger S.A., Kalisiak E., Scott J.W., Shanmukh S., Praissman J., Jenney F.E., Wikoff W.R., Apon J.V., Siuzdak G., Adams M.W.W.
Novel Multiprotein Complexes Identified in the Hyperthermophilic Archaeon *Pyrococcus furiosus* by Non-denaturing Fractionation of the Native Proteome
Molecular and Cellular Proteomics, **2009**, 8(4), 735-751

Woo H.K., Go E.P., Hoang L., Trauger S.A., Bowen B.,Siuzdak G., Northen T.R. **Phosphonium labeling for increasing metabolomic coverage of neutral lipids using electrospray ionization mass spectrometry** *Rapid Commun Mass Spectrom.* **2009**, 23 (12): 1849-55.

Kalisiak J., Trauger S.A., Kalisiak E., Morita H., Fokin VV., Adams M.W., Sharpless K.B., Siuzdak G. **Identification of a new endogenous metabolite and the characterization of its protein interactions through an immobilization approach.** *J Am Chem Soc.* **2009**, 131(1):378-86.

Benton H.P., Wong D.M., Siuzdak G. **XCMS2: Processing tandem mass spectrometry data for metabolite identification and structural characterization** *Analytical Chemistry*, **2008,** 80 (16), 6382-9.

Woo H-K., Northen T.R., Yanes O., Siuzdak G. **Nanostructure-Initiator Mass Spectrometry (NIMS): A protocol for preparing and applying NIMS surfaces for high sensitivity mass analysis** *Nature Protocols* **2008**, 3(8),1341-9..

O'Maille G., Hoang L., Nordstrom A., Go E.P., Qin C., Siuzdak G. **Metabolomics Mass Spectrometry Profiling using Chemical Derivatization with Isotope Labeling** *Spectroscopy*, **2008**.

Yanes O., Woo H.K., Northen T.R., Oppenheimer S.R., Shriver L., Apon J., Estrada M.N., Potchoiba M.J., Steenwyk R., Manchester M., Siuzdak G. **Nanostructure Initiator Mass Spectrometry: tissue imaging and direct biofluid analysis.** 2009, 81(8), 2969-75.

Northen T.R., Lee J-C, Hoang L., Raymond J., Hwang D-R, Yannone S.M., Wong C-H, Siuzdak G. **A Nanostructure-Initiator Mass Spectrometry-Based Enzyme Activity Assay** *PNAS*, **2008**, 105(10), 3678-3683.

Nordstrom A. Want E., Northen T., Lehtio J., Siuzdak G. **Multiple Ionization Mass Spectrometry Strategy Reveals the Complexity of Metabolomics** *Analytical Chemistry*, **2008**, 80, 421-429.

Trauger S.A., Kalisiak E., Kalisiak J., Morita H., Weinbergb M.V., Menon A.L., Poole F.L. II, Adams M.W.W., Siuzdak G.
Correlating the Transcriptome, Proteome, and Metabolome in the Environmental Adaptation of a Hyperthermophile
Journal of Proteome Research, **2008**, 7, 1027-1035.

Northen T.R., Yanes O., Northen M.T., Marrinucci D., Uritboonthai W., Apon J., Golledge S.L., Nordstrom A., Siuzdak G. **Clathrate Nanostructures for Mass Spectrometry** *Nature*, **2007**, 449, 1033-1036.

Want E.J., Nordstrom A., Morita H., Siuzdak G. **From Exogenous to Endogenous: The Inevitable Imprint of Mass Spectrometry in Metabolomics** *Journal of Proteome Research*, **2007**, 6(2), 459-468

Nordstrom A., He L., Siuzdak G. **Desorption/Ionization on Silicon (DIOS)** *Encyclopedia of Mass Spectrometry*, Editor Michael Gross, **2007**.

Northen T.R., Northen M.T., Nordstrom A., Uritboonthai W., Turner K., Siuzdak G. **High surface area of porous silicon drives desorption of intact molecules.** *Journal of the American Society of Mass Spectrometry **2007**,* 18 (11), 1945-9.

Go E.P., Uritboonthai W., Apon J.A., Trauger S.A., Nordstrom A., O'Maille G., Brittain S., Peters E.C., Siuzdak G. **Selective Metabolite and Peptide Capture/Mass Detection Using Fluorous Affinity Tags** *Journal of Prot. Res.,* **2007**, 6, 1492-1499.

O'Maille G., Hoang L., Nordstrom A., Go E.P., Qin C., Siuzdak G. **Enhanced Metabolite Profiling via Chemical Derivatization and Isotope Labeling** *Spectroscopy,* **2008.**

Go E.P., Wikoff W., Shen Z., O'Maille G., Morita H., Conrads T.P., Nordstrom A., Trauger S.A., Uritboonthai W., Lucas D., Chan K.C., Veenstra T.D., Lewicki H., Oldstone M.B., Schneemann A., Siuzdak G. **Mass Spectrometry Reveals Specific and Global Molecular Transformations during Viral Infection** *Journal of Proteome Research*, **2006**, 5 (9), 2405 – 2416.

Smith C.A., Want E.J., O'Maille G., Abagyan R., Siuzdak G. **XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification** *Analytical Chemistry*, **2006**, 78, 779-787.

Nordstrom A., O'Maille G., Qin C., Siuzdak G. **Non-linear Data Alignment for UPLC-MS and HPLC-MS based Metabolomics: Quantitative Analysis of Endogenous and Exogenous Metabolites in Human Serum** *Analytical Chemistry*, **2006**, 78, 7289-3295.