Pacific Northwest National Laboratory

Operated by Battelle for the U.S. Department of Energy

Application of LC-MS-Based Proteomics in the Discovery of Candidate Protein Biomarkers of Impaired Glucose Tolerance and Type 2 Diabetes

Thomas O. Metz[1], Wei-Jun Qian[1], Jon M. Jacobs[1], Ashoka D. Polpitiya[1], Marina A. Gritsenko[1], Hongliang Jiang[1], Therese R. W. Clauss[1], Rui Zhang[1], Anil K. Shukla[1], Aleksey V. Tolmachev[1], David G. Camp II[1], Lawrence S. Phillips[2], and Richard D. Smith[1]

[1]Biological Sciences Division, Pacific Northwest National Laboratory, [2]Department of Medicine, Emory University

Identification of a common anti-proteinase and human complement proteins as biomarkers

Overview

- Samples from the Screening for Impaired Glucose Tolerance (SIGT)[1] study investigated
- MS-based proteomics used in conjunction with immunodepletion of abundant proteins
- Accurate mass and time (AMT) tag approach[2] and stable isotope-labeling[3] employed for peptide identification and quantitation, respectively
- 84 peptides corresponding to 25 proteins identified as significantly different (ANOVA, $p < 0.05$) among NGT, IGT, and T2DM

Introduction

Diabetes mellitus affects approximately 5-10% of the world's population. Type 2 diabetes mellitus (T2DM) represents 90% of all reported cases and is reaching epidemic proportions. It is estimated that over 12 million individuals in the United States have been diagnosed with T2DM, while another 5 million are unaware of their disease. Pre-diabetes is the progenitor to T2DM and is characterized by elevated blood glucose levels insufficient to be classified as T2DM. Current estimates are that 47 million people in the United States are affected by pre-diabetes.

Currently, the oral glucose tolerance test (OGTT) is the best method to diagnose individuals as pre-diabetic, or having impaired glucose tolerance (IGT), and as having T2DM. However, the OGTT is not feasible in large populations and suffers from high cost.

In this study, we tested the hypothesis that application of proteomics in the study of plasma from individuals with normal glucose tolerance (NGT), IGT, and T2DM would result in the identification of novel protein biomarkers. We used mass spectrometry-based proteomics, immunodepletion of abundant plasma proteins, stable isotope-labeling, and the accurate mass and time (AMT) tag approach (Figure 1) to identify peptides, and therefore proteins, differentially expressed in plasma samples from individuals with NGT, IGT, and T2DM.

Methods

- Plasma samples were received frozen on dry ice and stored at -80°C until preparation.
- The top 6 abundant plasma proteins were depleted from pooled and individual NGT, IGT, and T2DM plasma samples by immunoaffinity chromatography (MARS-6, Agilent Technologies).
- Immunodepleted plasma proteins were enzymatically digested using trypsin.
- Plasma peptides from pooled NGT, IGT, and T2DM samples were fractionated by strong-cation exchange (SCX) chromatography then analyzed by LC-MS/MS
- Detected peptides were identified using the SEQUEST algorithm and added to an existing human plasma AMT tag database
- Plasma peptides from individual NGT, IGT, and T2DM samples were subjected to trypsin-catalyzed $^{16}O/^{18}O$ labeling. Individual samples were labeled with ^{16}O water, while a pooled reference sample was labeled with ^{18}O water.
- Aliquots of $^{16}O/^{18}O$-labeled plasma peptides were mixed 1:1 and analyzed by capillary LC-Fourier transform ion cyclotron resonance (FTICR) MS
- LC-FTICR MS data were processed using the Proteomics Research Information Storage and Management (PRISM) system
- Detected peptides were matched to the AMT tag database using initial mass and LC normalized elution time (NET) tolerances of ±6 ppm and 0.025 NET, respectively.
- Peptide abundances from multiple datasets were normalized by the Lowess method and comparative data analysis was performed using ANOVA
- Relative protein abundances were estimated using the RRollup method[4]

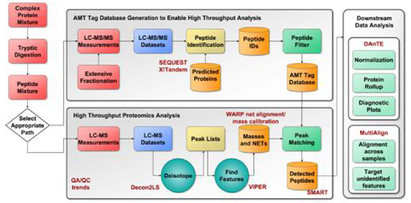

Figure 1. Overview of the AMT tag approach

Results

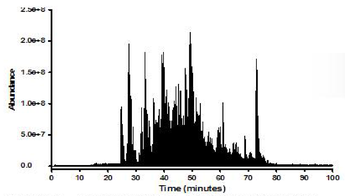

Figure 2. Representative chromatogram from LC-MS/MS analysis of peptides. Shown are data from an immunodepleted plasma protein digest of pooled T2DM plasma fractionated offline using SCX chromatography.

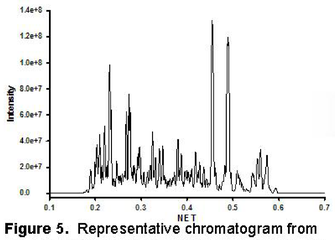

Figure 5. Representative chromatogram from LC-FTICR MS analysis of peptides. Shown are data from an immunodepleted plasma protein digest of T2DM plasma.

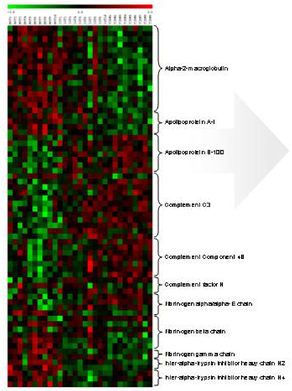

Figure 8. Subset of 25 proteins identified as candidate biomarkers of IGT and T2DM. Each row represents a unique peptide, and each column represents a unique sample. Plotted are z-scores of peptide abundances.

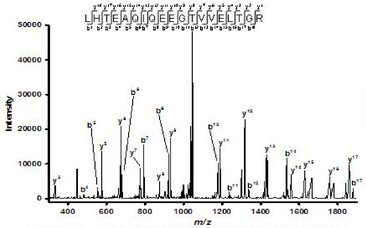

Figure 3. Tandem mass spectrum of detected peptide. Based on the fragment information, this peptide was identified as LHTEAQIQEEGTVVELTGR

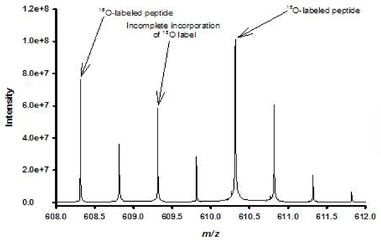

Figure 6. Representative mass spectrum from LC-FTICR MS analysis of peptides. Shown are $^{16}O/^{18}O$ data from the sample in Figure 3.

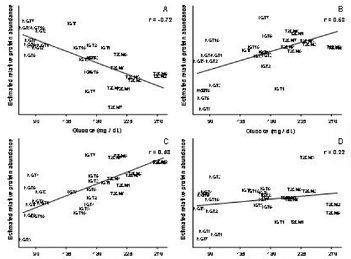

Figure 9. Correlation of estimated relative protein abundances for A) Alpha-2-macroglobulin, B) Complement C3, C) Complement component 4B, and D) Complement factor H to blood glucose concentrations measured 2 hr post-OGTT.

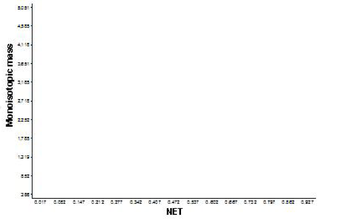

Figure 4. Calculated monoisotopic masses and observed NETs of identified peptides stored as AMT tags.

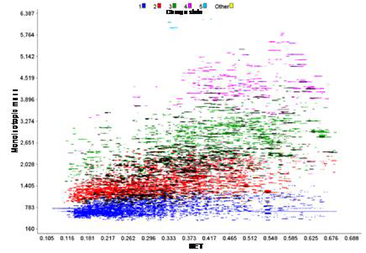

Figure 7. Deisotoped peptide peaks visualized in 2 dimensions (monoisotopic mass vs. NET). Shown are $^{16}O/^{18}O$ data from the sample in Figure 3.

Table 1. Correlation of estimated relative abundances for proteins shown in Figure 8 to various clinical data. Shown are the correlation coefficients (r).

Protein	Glucose - fasting	Glucose - 2 hr post-OGTT	Cholesterol	Triglycerides	HbA1c
Alpha-2-macroglobulin	-0.56	-0.72	-0.31	-0.42	-0.55
Apolipoprotein A-I	-0.43	-0.60	0.10	-0.27	-0.35
Apolipoprotein B-100	0.36	0.62	0.25	0.48	0.38
Complement C3	0.47	0.53	0.21	0.37	0.51
Complement component 4B	0.55	0.69	0.04	0.34	0.53
Complement factor H	0.04	0.22	0.09	0.00	0.24
Fibrinogen alpha/alpha-E chain	0.29	0.35	0.08	0.21	0.28
Fibrinogen beta chain	-0.25	-0.41	0.14	-0.24	-0.12
Fibrinogen gamma chain	-0.42	-0.43	0.04	-0.26	-0.29
Inter-alpha-trypsin inhibitor heavy chain H2	-0.25	-0.46	-0.26	-0.33	-0.33
Inter-alpha-trypsin inhibitor heavy chain H4	-0.11	-0.41	-0.35	-0.27	-0.23

Conclusions

- LC-MS-based proteomics in conjunction with the AMT tag approach successfully identified candidate protein biomarkers of IGT and T2DM
- Among 25 proteins identified by ≥ 2 unique peptides, Alpha-2-macroglobulin, Complement 3, Complement component 4B, and Complement factor H show strong perturbations during IGT and T2DM
- Apolipoproteins A-I and B-100 are also strongly altered during IGT and T2DM, but may reflect differences in HDL and LDL concentrations, respectively
- Link between Complement C3 and insulin resistance previously reported in individuals with obesity[5], T2DM[6,7], hypertension[8], and dyslipidemia[9]
- Further studies in a larger population with NGT, IGT, and T2DM are required to validate these candidate protein biomarkers

Acknowledgements

Portions of this research were supported by NIH NIDDK (DK071283 and DK066204), NCRR (RR00039 and RR018522), and the Robert W. Woodruff Health Sciences Center Fund. This research was performed at the Environmental Molecular Science Laboratory, a national scientific user facility located at Pacific Northwest National Laboratory and sponsored by the U. S. Department of Energy (DOE) Office of Biological and Environmental Research. PNNL is operated by Battelle for the DOE under Contract No. DE-AC06-76RLO-1830.

References

1. Phillips, L.S., et al. *Diabetes Care* **2006**, *29*:1405-1407.
2. Zimmer, J.S., et al. *Mass Spectrom. Rev.* **2006**, *25*:450-482.
3. Qian, W.J., et al. *Mol. Cell. Proteomics* **2005**, *4*:700-709.
4. Polpitiya, A.D., et al. *Bioinformatics* **2008**, In press.
5. Pomeroy, C., et al. *Clin. Exp. Immunol.* **1997**, *108*:507-515.
6. Mantov, S. and Raev, D. *Int. J. Cardiol.* **1996**, *56*:145-148.
7. Figueredo, A., et al. *Diabetes Care* **1993**, *16*:445-449.
8. Ylitalo, K., et al. *Atherosclerosis* **1997**, *129*:271-277.

Contact Information

Thomas O. Metz, PhD
Biological Sciences Division, K8-98
Pacific Northwest National Laboratory
P.O. Box 999, Richland, WA 99352
e-mail: thomas.metz@pnl.gov