RESEARCH NOTE

Selective degradation of serum proteins is likely responsible for the spurious differences in innate immunity proteins observed in a type 1 diabetes study [version 1; peer review: peer review discontinued]

Wenbo Zhi, Sharad Purohit, Shan Bai, Ashok Sharma, Jin-Xiong She

Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Georgia Regents University, Augusta, GA 30912, USA

V1 First published: 07 Oct 2014, 3:237
<https://doi.org/10.12688/f1000research.5384.1>

Latest published: 07 Oct 2014, 3:237
<https://doi.org/10.12688/f1000research.5384.1>

Abstract

Discovery and validation of serum protein biomarkers is of vital importance for the prediction, mechanism elucidation and monitoring response to therapy of type 1 diabetes mellitus. In this study, we attempted to replicate the results published in a 2013 issue of *The Journal of Experimental Medicine* by Qibin Zhang and colleagues described the discovery, verification and validation of several serum proteins/peptides that were drastically different between type 1 diabetes (T1D) patients and healthy controls, using label-free quantitative LC-MS-based proteomics and a multiple reaction monitoring mass spectrometry (MRM-MS) based multiplexed peptide assays. We performed the same MRM-MS assay in a large sample panel of 145 T1D patients and 156 autoantibody negative (AbN) control subjects (PANDA sample set) collected in the same geographical area, during the same period of time and by the same investigators, as well as 144 serum samples from the original authors (DASP sample set). Our measurement of 12 transitions/peptides in the DASP samples correlated very nicely with the authors' published results, indicating that the techniques used in the two laboratories yield very similar results on the same sample sets. Yet, in our PANDA samples, five of the twelve peptides (LLDSLPSDTR, FQPTLLTLPR, TNLESILSYPK, LVLLNAIYLSAK and ITQVLHFTK) whose serum levels are significantly different in the DASP verification and/or blind sample sets are not significant ($p > 0.05$). Only one peptide (TGAQELLR) showed marginal significance ($p = 0.03$). Although the remaining 6 peptides (NIQSLEVIGK, TLEAQLTPR, ELDESLQVAER, AGALNSNDAFVLK, TFTLLDPK and DIPTNSPELEETLHTITK) are significantly different between the T1D and control group in our PANDA sample set, the inter-group differences as measured by fold change (FC) are very small ($FC = 1.0 \pm 0.1$). Therefore, our results do not support the major findings in the report.

Peer review discontinued

Peer review at F1000Research is author-driven. Currently no reviewers are being invited. [What does this mean?](#)

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Jin-Xiong She (jshe@gru.edu)

Competing interests: No competing interests were disclosed.

Grant information: The study is supported by grants from the National Institutes of health (4R33HD050196, 4R33-DK069878 and 2RO1HD37800) and JDRF (1-2004-661) to JXS. Sharad Purohit was supported by a JDRF senior postdoctoral fellowship (JDRF 10-2006-792) and career development award (2-2011-153) from JDRF. Wenbo Zhi was supported by Georgia Regents University Extramural Success Award (ESA00017).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2014 Zhi W *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

How to cite this article: Zhi W, Purohit S, Bai S *et al.* **Selective degradation of serum proteins is likely responsible for the spurious differences in innate immunity proteins observed in a type 1 diabetes study [version 1; peer review: peer review discontinued]** F1000Research 2014, **3**:237 <https://doi.org/10.12688/f1000research.5384.1>

First published: 07 Oct 2014, **3**:237 <https://doi.org/10.12688/f1000research.5384.1>

Introduction

Type 1 diabetes (T1D) is an autoimmune disease that occurs when insulin-producing beta cells in the pancreas are destroyed by the body's immune system. Efforts are being made to develop biomarkers that could predict the risk of developing T1D or complications associated with the disease¹. Since pathological tissues from patients with T1D are very difficult to obtain for biomarker studies due to ethical and practical issues, proteomic analyses of serum proteins may provide useful biomarkers relevant to diseases because serum proteins may be secreted by cells at the pathological sites². Comprehensive analysis of the serum proteome is a challenging task using current proteomic technologies due to its extraordinary complexity and high dynamic range in concentration^{3,4}. Proteomic analyses are further complicated by the potentially large variation between individuals and relatively small differences between study groups. Furthermore, serum protein levels may be influenced by a large number of parameters such as subject age, sex, geographic or ethnic origin, environmental factors and disease states. Therefore, the quantification of serum proteome requires well designed strategies that incorporate different technologies and experimental designs at different phases, each addressing a different set of problems.

In a 2013 issue of *The Journal of Experimental Medicine*, Qibin Zhang and colleagues⁵ described the discovery, verification and validation of several serum proteins/peptides that were drastically different between type 1 diabetes (T1D) patients and healthy controls. The authors used label-free quantitative LC-MS-based proteomics to discover 24 proteins that showed significant changes between T1D and healthy control samples and further verified these differences using a MRM-MS based multiplexed peptide assays with spiked stable isotope-labeled standard (SIS) for normalization purpose. It would be of great scientific and clinical importance and interest if these biomarker candidates were to be further validated in independent samples and in different laboratories. Therefore, we attempted to replicate the results by using the same methods to analyze a large panel of 145 T1D patients and 156 autoantibody negative (AbN) control subjects (PANDA data set) collected in the same geographical area, during the same period of time and by the same investigators. Unfortunately, our results do not support the major findings in the report.

Materials and methods

Chemicals and materials

All chemicals were purchased from Sigma-Aldrich. Peptide-desalting plates (96-Well Macro SpinColumns™) were obtained from The Nest Group, Inc. Sequencing-grade trypsin was purchased from Promega. All solvents used were HPLC-grade or higher. SIS peptide (NIQSLEVIG[Lys] (13C6; 15N2, purity>95%) was custom synthesized by Thermo Fisher Scientific at the purity level of AQUA Basic (Purity>95%). The SIS peptides were received lyophilized and used as is without further purification.

Human subjects and serum samples

This study was approved by the institutional review board of the Georgia Regents University and informed consent was obtained from every subject or legally authorized representatives (parent/guardian).

Peripheral blood collected in serum separator tubes (BD Biosciences, San Jose, CA, USA) was allowed to clot for 30–120 minutes and then spun to obtain sera, which was immediately aliquoted and stored at -80°C freezers. To create serum assay plates, frozen serum samples were thawed and further aliquoted into 96-well v-bottom plates (80 samples per plate, roughly the same number of T1D patients and AbN controls) and stored at -80°C freezers. From blood collection to assay, none of the serum samples had more than three freeze/thaw cycles.

Digestion of serum samples for LC-MRM-MS analyses

5 µl of each neat serum sample (300~400 µg of proteins) was diluted with 45 µl of 8 M urea in 50 mM Tris-HCl (pH 8.0). Denatured samples were reduced with 20 mM of dithiothreitol at 56°C for 30 minutes, and then the reduced samples were alkylated with 55 mM iodoacetamide at room temperature for 30 minutes in the dark. Samples were then diluted with 450 µl 50 mM NH₄HCO₃ (pH 8.1) to reduce the urea concentration below 1 M, followed by addition of 1 M CaCl₂ solution to a final concentration of 1 mM and trypsin in the ratio of 1:50 enzyme/substrate (wt/wt). Digestion was performed overnight at 37°C. Serum digests were acidified by adding formic acid to a final concentration of 1% formic acid to stop the digestion. Next, 5 µl of SIS peptide was added to each serum digest. Samples were then desalted on C18 peptide-desalting plates and eluted peptides were freeze-dried and reconstituted with 50 µl 2% acetonitrile in water with 0.1% formic acid.

LC-MRM-MS analysis of serum digests

A Nexera UHPLC system (Shimadzu) equipped with a Kinetex UPLC column (2.1 mm × 100 mm, 1.7-µm particle size) was used for separation of tryptic digests. 5 µl (corresponding to 0.5 µl of original serum) of each sample was injected onto the column before the start of gradient LC separation. The LC flow rate was set at 0.3 ml/min with the following mobile phases: A, 0.1% formic acid (FA) in water; B, 0.1% FA in acetonitrile. The following gradient was used: 0 min, 2% B; 8 min, 35% B; 8.5 min, 95% B; 9.5 min, 95% B; 11 min, 2% B; 16 min, 2% B. The effluent from the LC column was ionized using a spray voltage of 5,500 V and peptides were detected using a triple quadrupole-ion trap hybrid mass spectrometer (4000 QTRAP, ABSCIEX). Other acquisition parameters were as follows: curtain gas of 30, temperature of 450°C, gas 1 of 50, declustering potential of 60, dwell time of 40 ms for each transition, Q1 and Q3 resolution of low and unit, respectively. The sequences/identities of the MRM transition peaks for all 12 peptides were further confirmed by MS/MS fragmentation analysis using 4000 QTRAP before data analysis.

MRM data processing and statistical analysis

The acquired datasets were imported into MultiQuant™ Software (ABSCIEX). LC-MS peak area integration for each transition in each sample was manually reviewed, and the peak area ratios between endogenous and SIS peptide, were exported. The peak areas of the endogenous peptides were normalized against the SIS peptide. A Student's t-test was performed on the normalized data to determine the significances between different groups.

Dataset 1. Peak areas of MRM transitions in PANDA and DASP sample set

<http://dx.doi.org/10.5256/f1000research.5384.d36443>

AbN represents autoantibody negative control group in PANDA sample cohort; T1D denotes type 1 diabetes patients in PANDA cohort; DASP_B_Control is control group in DASP blind sample set; DASP_B_Patient is T1D patient group in DASP blind sample set; DASP_V_Control is control group in DASP verification sample set; DASP_V_Patient is T1D patient group in DASP verification sample set; Plate # indicates the # of storage plate the sample locates for PANDA samples; IS column is the peak area of the transition for SIS peptide; NA indicates not detected.

Results and discussion

A total of four plates of serum samples from our PANDA sample cohort (145 T1D patients and 156 AbN control subjects) were analyzed using MRM assays for twelve peptides that have good MS signals ($s/n > 5$), including NIQSLEVIGK, LLDSLPSDTR, FQPTLLTLPR, TNLESILSYPK, LVLLNAIYLSAK, ITQVLHFTK, TLEAQLTPR, ELDESQVAER, AGALNSNDAFVLK, TFTLLDPK, DIPTNSPELEETLHTITK and TGAQELLR. As displayed in [Figure 1](#), five of the twelve peptides (LLDSLPSDTR, FQPTLLTLPR, TNLESILSYPK, LVLLNAIYLSAK and ITQVLHFTK) whose serum levels are significantly different in the DASP verification and/or blind sample sets are not significant at all in our large PANDA sample set ($p > 0.05$). TGAQELLR

showed marginal significance ($p = 0.03$). Although the remaining 6 peptides (NIQSLEVIGK, TLEAQLTPR, ELDESQVAER, AGALNSNDAFVLK, TFTLLDPK and DIPTNSPELEETLHTITK) are significantly different between the T1D and control group in our PANDA sample set, the inter-group differences as measured by fold change (FC) are very small ($FC = 1.0 \pm 0.1$). These results are in sharp contrast to the dramatically altered concentration for several peptides observed in the T1D patients in the initial report, including NIQSLEVIGK (FC of -1.1 in PANDA versus 8.5 in DASP_V), FQPTLLTLPR (FC of 1.0 versus 2.6), TNLESILSYPK (FC of 1.0 versus 2.2), and LLDSLPSDTR (FC of 1.1 versus 2.1).

In order to exclude the possibility of different MRM assay platforms causing the discrepancies between the published and our results, we obtained the samples used in the initial report (kindly provided by the authors). The samples include part of their DASP verification cohort (DASP_V, 99 controls and 25 T1D patients) and their entire DASP blind (DASP_B) validation cohort (10 controls and 10 T1D patients). We processed these samples and performed MRM assay using the same methods used to analyze our PANDA samples. The dot plots for DASP sample cohorts are also shown in [Figure 1](#) (DASP_V and DASP_B) together with the ones for our PANDA samples. The corresponding FC and p values are listed in [Table 1](#). The FC and p values from our measurement in the DASP samples correlated very nicely with the authors' published results, clearly indicating that the techniques used in the two laboratories yield very similar results on the same sample sets.

Table 1. Fold change (FC) and p values for the 12 peptides between T1D and control group of the three sample sets. The values in the parentheses are the published results from the original paper⁵. Values in red indicate significant changes. Red double-asterisk and single-asterisk denote fold change with a p value < 0.01 and < 0.05 , respectively, in our results.

Transition	Peptide	Protein	PANDA		DASP_V		DASP_B	
			FC	p value	FC	p value	FC	p value
550.8/873.5	NIQSLEVIGK	PPBP	-1.1**	4.92E-03	6.8 (8.5)**	2.07E-06 (6.62E-18)	26.0 (30.4)**	5.76E-10 (1.08E-05)
558.8/575.3	LLDSLPSDTR	SERPING1	1.1	3.02E-01	2.9 (2.1)*	2.12E-02 (1.39E-04)	6.2 (7.5)**	9.19E-10 (1.08E-05)
593.3/455.8	FQPTLLTLPR	SERPING1	1.0	8.74E-01	5.0 (2.6)**	7.54E-03 (1.71E-03)	7.0 (7.4)**	6.56E-06 (1.08E-05)
632.8/807.5	TNLESILSYPK	SERPING1	1.0	6.25E-01	1.6 (2.2)	6.15E-02 (2.49E-05)	4.0 (7.1)**	1.82E-05 (1.08E-05)
659.4/1105.7	LVLLNAIYLSAK	SERPING1	1.1	4.68E-01	5.4 (1.6)*	1.30E-02 (6.16E-03)	8.6 (3.3)**	1.35E-04 (2.06E-04)
444.3/729.4	TGAQELLR	GSN	-1.1*	2.70E-02	1.0 (-1.5)	9.2E-01 (3.41E-09)	-1.6 (-1.4)*	3.70E-02 (2.88E-02)
660.3/1007.5	AGALNSNDAFVLK	GSN	-1.1**	1.24E-04	-1.5 (-1.4)**	1.41E-06 (9.64E-11)	-1.4 (-1.5)**	1.09E-03 (2.88E-03)
514.8/814.4	TLEAQLTPR	SERPIND1	-1.2**	3.74E-05	-1.4 (-1.5)**	9.42E-05 (1.17E-11)	-1.3 (-1.2)	7.92E-02 (1.23E-01)
543.8/487.3	ITQVLHFTK	C4A	-1.1	8.67E-02	-1.8 (-1.6)**	1.05E-08 (3.63E-10)	-1.3 (-1.1)	7.75E-02 (9.71E-01)
644.8/602.3	ELDESQVAER	CLU	-1.1**	3.18E-04	-1.7 (-1.2)**	2.29E-13 (7.85E-05)	-1.1 (-1.1)	1.65E-01 (4.81E-01)
467.8/686.4	TFTLLDPK	PGLYRP2	-1.2**	2.07E-04	-1.8 (-1.5)**	5.74E-13 (1.31E-11)	-1.1 (-1.4)	2.88E-01 (6.84E-03)
713.7/956.0	DIPTNSPELEETLHTITK	KNG1	-1.2**	3.29E-05	-1.2 (-1.2)**	5.77E-03 (2.68E-07)	1.2 (-1.0)	2.20E-01 (5.29E-01)

Figure 1. Relative abundance for 12 peptides. Dot plots showing the peptide abundance in each study subjects. Three sample sets (PANDA, DASP_V and DASP_B) were analyzed in our laboratory using identical procedures. The black horizontal line in each group represents the mean value for the group. Relative peptide abundance (y axis) was calculated as the ratio between the endogenous peptide and its spiked-in, heavy isotope-labeled synthetic peptide. The numbers in the parentheses are the subject numbers in each group. Red double-asterisk and single-asterisk denote a p value <0.01 and <0.05, respectively.

We further compared the results for the two sample cohorts and found out that 6 of the 7 transitions/peptides whose serum levels are decreased in T1D groups of DASP_V and/or DASP_B sample sets are also significantly lower in the T1D group of our PANDA sample set (TGAQELLR, TLEAQLTPR, ELDESQVAER, AGALNSNDAFVLK, TFTLLDPK and DIPTNSPELEETLTHITTK, corresponding to 5 different proteins). In contrast, the transitions/peptides whose serum levels are significantly increased in T1D group from the DASP sample set either showed a significantly opposite trend (FC=-1.1, $p=4.92E-03$ versus FC=6.8, $p=2.07E-06$ in DASP_V for NIQSLEVIGK, corresponding to protein PPBP) or did not change significantly in our PANDA sample set ($p>0.05$ for LLDSLPSDTR, FQPTLLTLPR, TNLESILSYPK and LVLLNAIYLSAK, corresponding to the sample protein SERPING1).

How can one explain the drastic discrepancies between these two studies? Comparison of the serum concentration distribution of the 5 reported significantly elevated transitions/peptides in PANDA versus DASP sample sets (Figure 1) indicates that their expression levels are similarly high in the patients (T1D) from both studies. In contrast, the PANDA controls have similarly high levels as the patients from PANDA and DASP, while the DASP controls have dichotomous distribution for these five transitions/peptides. A few of the DASP controls have comparable levels as observed in the patients and PANDA controls, with a large proportion of the DASP controls have undetectable levels of those peptides (Figure 1). Our results clearly indicate that the DASP control samples contain many outliers that are responsible for the reported differences between T1D and controls in the previous study¹. To identify factors that may account for the discrepancies between controls in the two studies, we examined various demographic parameters available to us. Sex is unlikely to explain the differences as the PANDA and DASP cohorts contain subjects of both genders. The only apparent demographic difference between the controls in the two studies is the age distribution. The mean ages are 30.1 and 20.7 for controls from the PANDA and DASP datasets, respectively. However, the peptide concentration does not significantly vary according to subject's age in either the AbN controls or the T1D groups from PANDA (data not shown), an observation consistent with the authors' conclusion. The significant fold changes for the proteins in the original report are clearly caused by the extreme low serum protein levels in the control group, rather than the increased level in the patient group, as shown by the distinct dichotomous distribution and undetectable levels in many DASP controls (Figure 1). This situation

is most likely caused by degradation of proteins that are sensitive to sample collection and storage conditions, a phenomenon well known to protein biochemists^{6,7}. Interestingly, some proteins may be more susceptible to degradation than other proteins as suggested by the dichotomous distribution of the intensities of the 5 transitions derived from two proteins (PPBP and SERPING1) and the continuous distribution of the intensities of the 7 transitions derived from 6 other proteins (GSN, SERPIND1, C4A, CLU, PGLYRP2 and KNG1).

In conclusion, our results directly question the validity of the conclusions of the initial report and highlight the importance and necessity to conduct biomarker discovery and validation studies using well collected and preserved samples as well as sufficiently powered sample sets.

Data availability

F1000Research: Dataset 1. Peak areas of MRM transitions in PANDA and DASP sample set, [10.5256/f1000research.5384.d36443](https://doi.org/10.5256/f1000research.5384.d36443)⁸

Author contributions

WZ and SP performed experiments and wrote the manuscript. JXS initiated and supervised this study and revised the manuscript. SB and AS assisted in data analysis and manuscript revision.

Competing interests

No competing interests were disclosed.

Grant information

The study is supported by grants from the National Institutes of health (4R33HD050196, 4R33-DK069878 and 2R01HD37800) and JDRF (1-2004-661) to JXS. Sharad Purohit was supported by a JDRF senior postdoctoral fellowship (JDRF 10-2006-792) and career development award (2-2011-153) from JDRF. Wenbo Zhi was supported by Georgia Regents University Extramural Success Award (ESA00017).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors would like to thank Dr. Thomas O. Metz and his colleagues for providing us the DASP serum samples.

References

- Purohit S, She JX: **Biomarkers for type 1 diabetes**. *Int J Clin Exp Med*. 2008; 1(2): 98–116.
[PubMed Abstract](#) | [Free Full Text](#)
- Zhi W, Purohit S, Carey C, et al.: **Proteomic technologies for the discovery of type 1 diabetes biomarkers**. *J Diabetes Sci Technol*. 2010; 4(4): 993–1002.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Anderson L: **Candidate-based proteomics in the search for biomarkers of cardiovascular disease**. *J Physiol*. 2005; 563(Pt 1): 23–60.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Jacobs JM, Adkins JN, Qian WJ, et al.: **Utilizing human blood plasma for proteomic biomarker discovery**. *J Proteome Res*. 2005; 4(4): 1073–85.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Zhang Q, Fillmore TL, Schepmoes AA, et al.: **Serum proteomics reveals systemic dysregulation of innate immunity in type 1 diabetes**. *J Exp Med*. 2013; 210(1): 191–203.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Traum AZ, Wells MP, Aivado M, et al.: **SELDI-TOF MS of quadruplicate urine and serum samples to evaluate changes related to storage conditions**. *Proteomics*. 2006; 6(5): 1676–80.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Timms JF, Arslan-Low E, Gentry-Maharaj A, et al.: **Preanalytic influence of sample handling on SELDI-TOF serum protein profiles**. *Clin Chem*. 2007; 53(4): 645–56.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Zhi W, Purohit S, Bai S, et al.: **Peak areas of MRM transitions in PANDA and DASP sample set**. *F1000Research*. 2014.
[Data Source](#)

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research