



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

From hundreds to thousands: Widening the normal human Urinome



Laura Santucci^{a,1}, Giovanni Candiano^{a,1}, Andrea Petretto^b,
Maurizio Bruschi^a, Chiara Lavarello^b, Elvira Inglese^b,
Pier Giorgio Righetti^c, Gian Marco Ghiggeri^{a,*}

^a Nephrology, Dialysis, Transplantation Unit and Laboratory on Pathophysiology of Uremia, Istituto Giannina Gaslini, 16148 Genova, Italy

^b Laboratory of Mass Spectrometry – Core Facility, Istituto Giannina Gaslini, 16148 Genova, Italy

^c Politecnico di Milano, Department of Chemistry, Materials and Chemical Engineering “Giulio Natta”, Via Mancinelli 7, 20131 Milano, Italy

ARTICLE INFO

Article history:

Received 6 August 2014

Accepted 8 August 2014

Available online 22 August 2014

Keywords:

Urinary proteome

Vesicles

Combinatorial peptide ligand libraries

Mass spectrometry

ABSTRACT

The limits on protein detection in urine are unknown. Improving the analytical approach to detection would increase the number of identified proteins and potentially strengthen their predictive potential in diseases.

Here, we present the data that resulted from a combination of analytical procedures for maximizing sensitivity and reproducibility of normal human urinary proteome analysis. These procedures are ultracentrifugation, vesicle separation, combinatorial peptide ligand libraries (CPLL) and solvent removal of pigments. Proteins were identified by an Orbitrap Velos Mass Spectrometry. 3429 proteins are characterized, 1724 of which are novel discoveries.

The data are related to Santucci et al. (in press) [1] and available both here and at ChorusProject.org under project name “From hundreds to thousands: widening the normal human Urinome”. The material supplied to Chorus Project.org includes technical MS spectra data only.

© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/3.0/>).

DOI of original article: <http://dx.doi.org/10.1016/j.jprot.2014.07.021>

* Correspondence to: Division of Nephrology, Dialysis and Transplantation and Laboratory on Pathophysiology of Uremia, Istituto G. Gaslini, Largo G. Gaslini 5, 16148 Genova, Italy. Tel.: +39 010 380742; fax: +39 010 395214.

E-mail address: GMarcoGhiggeri@ospedale-gaslini.ge.it (G. Marco Ghiggeri).

¹ These two authors have equally contributed to the work and share first authorship.

<http://dx.doi.org/10.1016/j.dib.2014.08.006>

2352-3409/© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

Specifications table

Subject area	Biology
More specific subject area	Characterization of urine proteome in normal conditions. It was developed a new strategy of analysis which allowed the extension of the number of identified proteins to 3429 (1724 were of new description here). The list of identified proteins is reported in Supplementary Table 1 .
Type of data	Excel tables
How data was acquired	<i>Ultracentrifuge</i> : Beckman Optima™ 1-90K (Ti 90 rotor). <i>Mass spectroscopy</i> : for analysis of all urinary sub-fractions was used a linear Trap Quadrupole (LTQ) Orbitrap Velos Pro Mass Spectrometry.
Data format	Raw MS files as processed with the Thermo Scientific Proteome Discoverer software: peak were searched by the MASCOT and SEQUEST against Uniprot human database, filtered for a maximum 1% FDR using Percolator; the Peptide Mass Deviation was set to 10 ppm and a minimum of six 6 amino acids per identified peptide were required [2,3]. The Database search parameters were mass tolerance precursor 20 ppm [4], mass tolerance fragment CID 0.8 Da with dynamic modification of deamidation (N, Q), oxidation (M) and static modification of alkylation with IAM (C). Label free experiments and statistical analysis were performed with Perseus software.
Experimental factors	<i>Sample Collection and Urine Preparation</i> : Second morning urine (approximately 160 mL) from healthy volunteers were collected, centrifuged at 4 °C for 10 min at 1000g. Urine were frozen and maintained at –80 until further processing [5–14]. Second step was centrifugation at 17,000g for 30 min at 16 °C. The sample, after Bradford protein assay [15], was stored at –80 °C until use.
Experimental features	Procedure used for analysis of the urinary proteome: Vesicle Isolation, butanol precipitation, Combinatorial Solid-Phase Ligand Library chromatography, Mass spectrometry and Bioinformatic analysis.
Consent	Informed consent was obtained from all the participants in the study.
Data source location	Genova, Italy.
Data accessibility	The data are available at ChorusProject.org under project name “From hundreds to thousands: widening the normal human Urinome”. Row data relative to MS spectra are available at ChorusProject.org under project name “From hundreds to thousands: widening the normal human Urinome”. Data are also directly available with this article

Value of the data

- Sub-fractionating normal urine by successive steps allowed to identify 3429 proteins, a net +50% increment compared to traditional methods of analysis.
- Vesicles separation, CPLL and solvent treatments are the basic steps.
- 1724 of the urine proteins identified here are of newly identified and described. Improved characterization of the normal urinary proteome opens doors for the analysis of urine biomarkers in human diseases.

1. Data, experimental design, materials and methods

3429 Non-redundant proteins discovered in our urine proteomic analysis are characterized and are noted in [Supplementary Table 1](#) [1]. 1615 of these proteins were contained in vesicles while the remaining 1794 were equally distributed among CPLL (1488) and butanol insoluble fractions (322). Several proteins were detected exclusively in one of the phases of the procedure, suggesting that each step is crucial in the fractionation strategy. Many (1724) proteins are here described whose presence in urine have never been reported and represent a potential source of information considering that urine is the unique site of excretion of products of interaction of metabolic processes.

1.1. Urinary vesicles isolation

The 17,000g urinary supernatant (80 mL) was ultracentrifuged at 48,000 rpm for 75 min at 18 °C. The ultracentrifugation step was repeated by adding the same volume used before until vesicles [16] were isolated. The pellets, washed in DTT (200 mg/mL) and Tris–HCl 65 mM pH 8.8, was centrifuged a 14,000 rpm for 10 min at 4 °C and stored at –80 °C until mass spectrometry analysis.

1.2. Butanol precipitation

The 48,000 rpm supernatant was dialyzed versus water; aliquots of 50 mL were added with 100 μ L of acetic acid (about pH 3–4) and 10 mL of n-butanol and centrifuged at 4,000 rpm for 10 min at 18 °C. Three different phases were obtained: protein pellet, supernatant (for ProteoMiner™, see below) and pigments (discarded), yielding two fractions: CPLL-beads chromatography and unbound.

1.2.1. Combinatorial peptide ligand library (CPLL)

The phase deriving from butanol extraction was lyophilized and loaded onto a column of 150 μ L peptide library beads equilibrated in 25 mM phosphate buffer, pH 7.4 as already described by Candiano et al. [17,18]. The eluate and the unbound fraction were preserved at –80 °C until analysis by mass spectrometry.

1.2.2. Mass spectrometry

Samples for mass spectrometry were solubilized in 0.1 mL of 4% SDS, 50 mM DTT, and 0.1 M Tris/HCl, pH 7.6, at 90 °C for 5 min and briefly sonicated and were processed by the FASP procedure using 30k Vivacom filtration devices (Sartorius) [19].

The mass spectrometer LTQ-Orbitrap Velos Pro was operated in positive ionization mode.

Single MS survey scans were performed in the Orbitrap, recording a mass window between 350 and 1650 m/z using a maximal ion injection time of 250 ms. The resolution was set to 60,000 and the automatic gain control was set to 1,000,000 ions. The experiments were done in data-dependent acquisition mode with alternating MS and MS/MS experiments. A maximum of 10 MS/MS experiments were triggered per MS scan.

Raw MS files were processed with Thermo Scientific Proteome Discoverer software version 1.3. Peak list files were searched by the MASCOT and SEQUEST search engine against Uniprot human database (Release 2012_07) containing both forward and reversed protein sequences. The Database search parameters are mass tolerance precursor 20 ppm, mass tolerance fragment CID 0.8 Da and a dynamic modification of deamidation (N, Q), oxidation (M). For all searches the option trypsin with two missed cleavages was selected. Proteins were grouped by applying the maximum parsimony rule. Resulting peptide hits were filtered for a maximum 1% FDR using Percolator and a Peptide Mass Deviation of 10 ppm per identified peptide were required. To determine the area for any identified peptides, we used precursor ions area detector node of Proteome Discoverer.

1.2.3. Bioinformatic analysis

All the statistical analyses of the identified protein tables were done with the Perseus program (J. Cox, Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany). For hierarchical clustering, we filtered the data according to the technical replicate and kept proteins with a minimum of 3 values repeated in the different analysis. Logarithmized areas were z-scored and clustered using Euclidean distances between averages. Fisher exact tests were done with a Benjamini–Hochberg FDR threshold of 0.02.

Conflict of interest

All authors declare no conflicts of interest.

Acknowledgments

The Giannina Gaslini Institute provided financial and logistic support to the study. This work was also supported by the Ministry of Health, Italy 'Ricerca Corrente' and from contributions derived from 'Cinque per mille dell'IRPEF'. We also acknowledge contributions from the Renal Child Foundation, Fondazione La Nuova Speranza ('Progetto integrato per la definizione dei meccanismi implicati nella glomerulo sclerosi focale') and Italian Society of Nephrology (Progetto Ricercando).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2014.08.006>.

References

- [1] L. Santucci, G. Candiano, A. Petretto, M. Bruschi, C. Lavarello, E. Inglese, et al., From hundreds to thousands: widening the normal human Urinome (1), *J. Proteomics* (2014) <http://dx.doi.org/10.1016/j.jprot.2014.07.021>. (in press).
- [2] R.M. Branca, L.M. Orre, H.J. Johansson, V. Granholm, M. Huss, A. Perez-Bercoff, et al., HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics, *Nat. Methods* 11 (2014) 59–62.
- [3] N. Gupta, P.A. Pevzner False, discovery rates of protein identifications: a strike against the two-peptide rule, *J. Proteome Res.* 8 (2009) 4173–4181.
- [4] E.J. Hsieh, M.R. Hoopmann, B. MacLean, M.J. MacCoss, Comparison of database search strategies for high precursor mass accuracy MS/MS data, *J. Proteome Res.* 9 (2010) 1138–1143.
- [5] J. Adachi, C. Kumar, Y. Zhang, J.V. Olsen, M. Mann, The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins, *Genome Biol.* 7 (2006) R80.
- [6] M. Court, N. Selevsek, M. Matondo, Y. Allory, J. Garin, C.D. Masselon, et al., Toward a standardized urine proteome analysis methodology, *Proteomics* 11 (2011) 1160–1171.
- [7] A. Kentsis, F. Monigatti, K. Dorff, F. Campagne, R. Bachur, H. Steen, Urine proteomics for profiling of human disease using high accuracy mass spectrometry, *Proteomics Clin. Appl.* 3 (2009) 1052–1061.
- [8] R.S. Lee, F. Monigatti, A.C. Briscoe, Z. Waldon, M.R. Freeman, H. Steen, Optimizing sample handling for urinary proteomics, *J. Proteome Res.* 7 (2008) 4022–4030.
- [9] Q.R. Li, K.X. Fan, R.X. Li, J. Dai, C.C. Wu, S.L. Zhao, et al., A comprehensive and non-prefractionation on the protein level approach for the human urinary proteome: touching phosphorylation in urine, *Rapid Commun. Mass Spectrom.* 24 (2010) 823–832.
- [10] A. Marimuthu, R.N. O'Meally, R. Chaerkady, Y. Subbannayya, V. Nanjappa, P. Kumar, et al., A comprehensive map of the human urinary proteome, *J. Proteome Res.* 10 (2011) 2734–2743.
- [11] W. Sun, F. Li, S. Wu, X. Wang, D. Zheng, J. Wang, et al., Human urine proteome analysis by three separation approaches, *Proteomics* 5 (2005) 4994–5001.
- [12] V. Thongboonkerd, K.R. McLeish, J.M. Arthur, J.B. Klein, Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation, *Kidney Int.* 62 (2002) 1461–1469.
- [13] A. Vlahou, J. Schanstra, J. Frokiaer, M. El Nahas, G. Spasovski, H. Mischak, et al., Establishment of a European network for urine and kidney proteomics, *J. Proteomics* 71 (2008) 490–492.
- [14] P.G. Zerefos, M. Aivaliotis, M. Baumann, A. Vlahou, Analysis of the urine proteome via a combination of multi-dimensional approaches, *Proteomics* 12 (2012) 391–400.
- [15] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [16] T. Pisitkun, R.F. Shen, M.A. Knepper, Identification and proteomic profiling of exosomes in human urine, *Proc. Natl. Acad. Sci. USA* 101 (2004) 13368–13373.
- [17] G. Candiano, L. Santucci, M. Bruschi, A. Petretto, C. D'Ambrosio, A. Scaloni, et al., Cheek-to-cheek urinary proteome profiling via combinatorial peptide ligand libraries: a novel, unexpected elution system, *J. Proteomics* 75 (2012) 796–805.
- [18] L. Santucci, G. Candiano, M. Bruschi, C. D'Ambrosio, A. Petretto, A. Scaloni, et al., Combinatorial peptide ligand libraries for the analysis of low-expression proteins: validation for normal urine and definition of a first protein MAP, *Proteomics* 12 (2012) 509–515.
- [19] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, *Nat. Methods* 6 (2009) 359–362.