# Exploring the Human Plasma Proteome

*Edited by*
*Gilbert S. Omenn*


WILEY-VCH Verlag GmbH & Co. KGaA

**Exploring the Human Plasma Proteome**

*Edited by*
*Gilbert S. Omenn*

## *Related Titles*

Jungblut, P. R.,
Hecker, M. (Eds.)

**Proteomics of Microbial Pathogens**

2006
ISBN-13: 978–3-527–31759–2
ISBN-10: 3–527–31759–7

Liebler, D. C., Petricoin, E. F.,
Liotta, L. A. (Eds.)

**Proteomics in Cancer Research**

2005
ISBN-13: 978–0-471–44476–3
ISBN-10: 0–471–44476–6

Lion, N., Rossier, J. S.,
Girault, H. (Eds.)

**Microfluidic Applications in Biology**

**From Technologies to Systems Biology**

2006
ISBN-13: 978–3-527–31761–5
ISBN-10: 3–527–31761–9

Sanchez, J.-C., Corthals, G. L., Hochstrasser, D. F. (Eds.)

**Biomedical Applications of Proteomics**

2004
ISBN-13: 978–3-527–30807–1
ISBN-10: 3–527–30807–5

Hamacher, M., Marcus, K., Stühler, K.,
van Hall, A., Warscheid, B., Meyer, H.E.
(Eds.)

**Proteomics in Drug Research**

2006
ISBN-13: 978–3-527–31226–9
ISBN-10: 3–527–31226–9

# Exploring the Human Plasma Proteome

*Edited by*
*Gilbert S. Omenn*


WILEY-VCH Verlag GmbH & Co. KGaA

**The Editor**

***Prof. Dr. Gilbert S. Omenn***
University of Michigan
A520 MSRB I Bldg
1150 West Medical Center
Dr. Ann Arbor,
MI 48109–0626
USA



**Library of Congress Card No.:**
applied for

**British Library Cataloguing-in-Publication Data**
A catalogue record for this book is available from the British Library.

**Bibliographic information published by Die Deutsche Bibliothek**
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.



Printed in the Federal Republic of Germany
Printed on acid-free paper

**Typesetting** X Con Media AG, Bonn
**Printing** Strauss GmbH, Mörlenbach
**Binding** Litges & Dopf Buchbinderei GmbH, Heppenheim

**ISBN:** 3-527-31757-0
**ISBN 13:** 978-3-527-31757-8

## Table of Contents

# Preface

Plasma and serum are the preferred specimens for non-invasive sampling of normal individuals, at-risk groups, and patients for protein biomarkers discovered and validated to reflect physiological, pathological, and pharmacological phenotypes. These specimens present enormous challenges due to extreme complexity, representing potentially all proteins in the body and their isoforms; at least ten orders of magnitude range in protein concentrations; intra-individual and inter-individual variation from genetics, diet, smoking, hormones, and many other sources; and especially non-standardized methods of sample processing. Furthermore, the inherent limitations of incomplete sampling of peptides by mass spectrometry and high error rates of peptide identifications and protein assignments with various search algorithms and databases lead to low concordance of protein identifications even with repeat analyses of the same sample. These features complicate diagnostic comparisons of specimens.

The Human Proteome Organization (HUPO) has launched several major initiatives to explore the proteomes of liver, brain, and plasma and to generate informatics standards and large-scale antibody production. This book presents the major findings from the pilot phase of the Plasma Proteome Project (PPP). The 17 chapters embrace a combination of collaborative analyses of HUPO PPP reference specimens and several lab-specific projects, both experimental and analytical. The investigators compared PPP reference specimens of human serum and EDTA, heparin, and citrate-anti-coagulated plasma; EDTA-plasma was determined to be the preferred specimen. Together these chapters examine many features of specimen handling, depletion of abundant proteins, fractionation of intact proteins, fractionation of tryptic digest peptides, and analysis of those peptides with various MS/MS instruments. Combinations of technologies gave the most resolution. The subsequent step of matching spectra to peptide sequences with a variety algorithms has numerous, often unspecified parameters. The alignment of peptide sequences with proteins via protein or gene databases likewise is laden with uncertainties and redundancies. Especially for longitudinal and collaborative studies, the periodic issuance of modified versions of the databases creates a moving target for protein identification and annotation, let alone comparison of results from different studies. These challenges are explored in depth. As in the special issue of Proteomics (August 2005) with a total of 28 papers, the authors here provide a revealing snapshot of the output from a variety of proteomics technology platforms across laboratories.

The extensive annotations show that present methods already are capable of detecting in plasma large numbers of low-abundance proteins of great biological interest from essentially all cellular compartments. Studies focusing on subproteomes based on glycoprotein enrichment or molecular weight yielded additional findings. As more powerful technologies are applied, we can expect ever more extensive identification, as well as quantitation, of proteins and their isoforms. The high proportion of genes which generate detectable splice isoforms further complicates protein identifications, yet helps to clarify the basis on which humans can have such complex phenotypes with a surprisingly small complement of genes (latest Human Genome Project estimate is about 22,000 protein-encoding genes).

The PPP Core Dataset has 5102 proteins identified with 2 or more peptides, of which 3020 remain after application of our integration algorithm for protein matches which cannot be distinguished with the available peptides. A special feature of the PPP is the set of independent analyses from the raw spectra or peaklists across the multiple laboratories. These independent analyses eliminate the high variability from lab-specific search algorithms, different databases, and investigators' judgments, though each independent analysis has its own peculiar attributes. We also provide comparisons with several published datasets. Meta-analysis of separate studies has similar challenges to those experienced in the integration of datasets from the collaborating PPP laboratories.

Numerous other "cuts" of the data can be made. The primary data are available for such additional analyses at the European Bioinformatics Institute (www.ebi.,ac.uk/pride); the University of Michigan (www.bioinformatics.med.umich.edu/hupo/ppp); and the Institute for Systems Biology (www.peptideatlas.org). We are keen to encourage such further analyses. Two examples have already appeared, introducing adjustments for protein length and multiple comparisons testing [1] and enhancing the characterization of the human genome from these proteomics data and gene mapping [2]. This publication presents the foundation for planning the next phases of the Plasma Proteome Project, with Young-Ki Paik, Matthias Mann, and myself as co-chairs. We will:

1. develop standardized operating procedures for specimens, protein and peptide fractionation, and analyses, with attention to replicability of results, to make proteomics practicable for clinical chemistry;
2. select priority PPP proteins for the HUPO Antibody Production Initiative, to generate reagents for biomarker and pathways studies and plasma/organ proteome comparisons;
3. collaborate on informatics, databases, annotations, and error estimation for plasma and serum studies, both HUPO-initiated and published by others;
4. stimulate proteomics technology advances, with special attention to high-resolution/higher-throughput methods and to quantitation of proteins and characterization of modified proteins (primarily glycoproteins and phosphoproteins); and
5. assure paired analyses of plasma and tissue specimens in organ-based and disease-focused proteomics initiatives.

The spirit of collaboration in the Plasma Proteome Project has been splendid. The substantial commitment of so many investigators and sponsors to this pilot phase has been admirable. As a work-in-progress the PPP has generated productive discussions at many scientific meetings. On behalf of the Executive Committee and Technical Committees, I thank everyone involved.

*Gilbert S. Omenn*
*University of Michigan, Ann Arbor*
*August 2006*

**1.** States, D. J., Omenn, G. S., Blackwell, T. W., Fermin, D., Eng, J., Speicher, D. W., Hanash, S. M. *Challenges in deriving high-confidence protein identifications from data gathered by HUPO plasma proteome collaborative study. Nature Biotech* 2006, *24*, 333–338.

**2.** Fermin, D., Allen, B. B., Blackwell, T. W., Menon, R., Adamski, M., Xy, Y., Ulintz, P., Omenn, G. S., States, D. J. *Novel gene and gene model detection using a whole genome open reading frame analysis in proteomics. Genome Biology* 2006, *7:R35*, Published online: http://genomebiology.com/2006/7/4/R35.

## List of Contributors

***Dr. Gilbert S. Omenn***
Internal Medicine,
University of Michigan, MSRB 1,
1150 W. Medical Center Dr.
Ann Arbor,
MI 48109-0656, USA

***Dr. David J. States***
University of Michigan,
2017 Palmer Commons,
100 Washtenaw Avenue,
Ann Arbor,
MI 48109-2218, USA

***Dr. Alex J. Rai***
Assistant Professor and Director
of General Chemistry,
The Johns Hopkins University
School of Medicine,
Department of Pathology,
600 N. Wolfe St.,
Meyer B-121, Baltimore,
MD 21287-7065, USA

***Brian B. Haab***
Ph.D.,
The Van Andel Research Institute,
333 Bostwick NE,
Grand Rapids,
MI 49503, USA

***Dr. David W. Speicher***
The Wistar Institute,
3601 Spruce St. Rm. 151,
Philadelphia,
PA 19104, USA

***Dr. David W. Speicher***
The Wistar Institute,
3601 Spruce Street,
Philadelphia,
PA 19104, USA

***Professor William S. Hancock***
Barnett Institute and Department
of Chemistry and Chemical Biology,
Northeastern University,
Boston,
MA 02115, USA

***Professor Arie Admon***
Department of Biology,
Technion,
Haifa 3200, Israel

***Professor Young-Ki Paik***
Yonsei Proteome Research Center
and Biomedical Proteome
Research Center,
Yonsei University,
134 Shinchon-dong,
Sudaemoon-ku,
Seoul 120-749, Korea

***Dr. Xiaohong Qian***
Ph.D.,
Beijing Institute of Radiation Medicine,
27 Taiping Road,
Beijing 100850,
China

***Dr. Joel G. Pounds***
Biological Sciences Division,
Pacific Northwest National Laboratory,
Box 999 MSIN: P7-58,
Richland,
WA 99352, USA

***Dr. Alex J. Rai***
Assistant Professor and Director of General Chemistry,
Department of Pathology,
Division of Clinical Chemistry,
Johns Hopkins University School of Medicine,
600 N. Wolfe St., Meyer B-121,
Baltimore,
MD 21287-7065, USA

***Professor Richard J. Simpson***
Joint ProteomicS Laboratory,
Ludwig Institute for Cancer Research,
P.O. Box 2008,
Royal Melbourne Hospital,
Parkville, Victoria 3050,
Australia

***Dr. Eric W. Deutsch***
Institute for Systems Biology,
1441 N 34th Street,
Seattle,
WA 98103, USA

***Dr. Lennart Martens***
Department of Biochemistry,
Faculty of Medicine and Health Sciences,
Ghent University,
A. Baertsoenkaai 3,
B-9000 Ghent, Belgium

***Dr. Akhilesh Pandey***
McKusick-Nathans Institute of Genetic Medicine,
733 N. Broadway, BRB 569,
Johns Hopkins University,
Baltimore,
MD 21205, USA

***Thomas M. Vondriska***
Cardiovascular Research Laboratories,
Departments of Physiology and Medicine,
Division of Cardiology,
David Geffen School of Medicine at UCLA,
Room 1619, MRL Building,
Los Angeles,
CA 90095, USA

# 1

# Overview of the HUPO Plasma Proteome Project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database*

*Gilbert S. Omenn, David J. States, Marcin Adamski, Thomas W. Blackwell, Rajasree Menon, Henning Hermjakob, Rolf Apweiler, Brian B. Haab, Richard J. Simpson, James S. Eddes, Eugene A. Kapp, Robert L. Moritz, Daniel W. Chan, Alex J. Rai, Arie Admon, Ruedi Aebersold, Jimmy Eng, William S. Hancock, Stanley A. Hefta, Helmut Meyer, Young-Ki Paik, Jong-Shin Yoo, Peipei Ping, Joel Pounds, Joshua Adkins, Xiaohong Qian, Rong Wang, Valerie Wasinger, Chi Yue Wu, Xiaohang Zhao, Rong Zeng, Alexander Archakov, Akira Tsugita, Ilan Beer, Akhilesh Pandey, Michael Pisano, Philip Andrews, Harald Tammen, David W. Speicher and Samir M. Hanash*

HUPO initiated the Plasma Proteome Project (PPP) in 2002. Its pilot phase has (1) evaluated advantages and limitations of many depletion, fractionation, and MS technology platforms; (2) compared PPP reference specimens of human serum and EDTA, heparin, and citrate-anti-coagulated plasma; and (3) created a publicly-available knowledge base (www.bioinformatics.med.umich.edu/hupo/ppp; www.ebi.ac.uk/pride). Thirty-five participating laboratories in 13 countries submitted datasets. Working groups addressed (a) specimen stability and protein concentrations; (b) protein identifications from 18 MS/MS datasets; (c) independent analyses from raw MS-MS spectra; (d) search engine performance, subproteome analyses, and biological insights; (e) antibody arrays; and (f) direct MS/SELDI analyses. MS-MS datasets had 15 710 different International Protein Index (IPI) protein IDs; our integration algorithm applied to multiple matches of peptide sequences yielded 9504 IPI proteins identified with one or more peptides and 3020 proteins identified with two or more peptides (the Core Dataset). These proteins have been characterized with Gene Ontology, InterPro, Novartis Atlas, OMIM, and immunoassay-based concentration determinations. The database permits examination of many other subsets, such as 1274 proteins identified with three or more peptides. Reverse protein to DNA matching identified proteins for 118 previously unidentified ORFs.

* Originally published in Proteomics 2005, 13, 3226–3245

We recommend use of plasma instead of serum, with EDTA (or citrate) for anticoagulation. To improve resolution, sensitivity and reproducibility of peptide identifications and protein matches, we recommend combinations of depletion, fractionation, and MS/MS technologies, with explicit criteria for evaluation of spectra, use of search algorithms, and integration of homologous protein matches.

This Special Issue of PROTEOMICS presents papers integral to the collaborative analysis plus many reports of supplementary work on various aspects of the PPP workplan. These PPP results on complexity, dynamic range, incomplete sampling, false-positive matches, and integration of diverse datasets for plasma and serum proteins lay a foundation for development and validation of circulating protein biomarkers in health and disease.

## 1.1 Introduction

A comprehensive, systematic characterization of circulating proteins in health and disease will greatly facilitate development of biomarkers for prevention, diagnosis, and therapy of cancers and other diseases [1]. Proteomics technologies now permit extensive fractionation of proteins in complex specimens, analysis of peptides by MS, and matching of peptide sequences to protein "hits" through gene and protein databases generated directly and indirectly from the sequencing of the human genome [2, 3], as well as other methods for identifying proteins.

The HUPO, formed in 2001, aims to accelerate the development of the field of proteomics and to stimulate and organize international collaborations in research and education [4]. HUPO has launched major initiatives focused on the plasma, liver, and brain proteomes, proteomics standards and databases, and large-scale antibody production. The plasma proteome is linked with these other initiatives (see Fig. 1).

The long-term scientific goals of the HUPO Plasma Proteome Project (PPP) are (1) comprehensive analysis of the protein constituents of human plasma and serum; (2) identification of biological sources of variation within individuals over time due to physiology (age, sex, menstrual cycle, exercise, stress), pathology (various diseases, special cohorts), and treatments (common medications); and (3) determination of the extent of variation across individuals within populations and across populations due to genetic, nutritional and other factors. The pilot phase aims to (1) compare advantages and limitations of many technology platforms; (2) contrast reference specimens of human plasma (EDTA, heparin, or citrate-anticoagulated) and serum in terms of numbers of proteins identified and any interferences with various technology platforms; and (3) create a global, open-source knowledge base/data repository.

The collaborative nature of this Project permitted exploration of many variables and adoption during the study phase of emerging technologies. Planning proceeded expeditiously from the organizing meeting of HUPO in Bethesda in

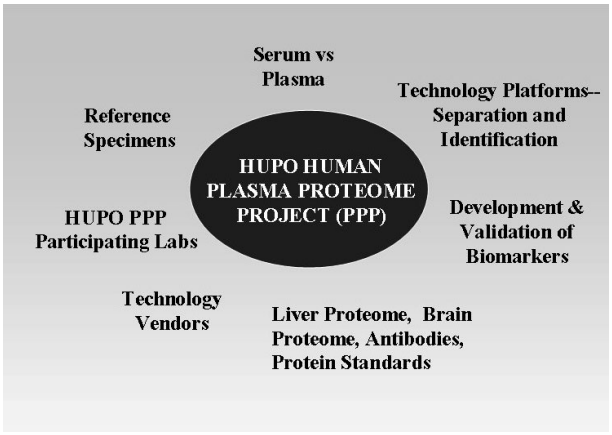


**Fig. 1** Schema showing relationship of HUPO Plasma Proteome Project (PPP) to other HUPO initiatives and components of the PPP.

April 2002, to the first PPP meeting in Ann Arbor in September 2002, the expression of interest by numerous investigators at the 1st HUPO World Congress on Proteomics in Versailles in November 2002, and then the PPP Workshop for Technical Committees and participating laboratories in Bethesda in July 2003 to launch the pilot phase. PPP reference specimens were prepared and distributed, beginning in September 2003, and first data were submitted, analyzed, and presented at a workshop at the 2nd HUPO World Congress in Montreal in November 2003. An intensive 4 day Jamboree Workshop was organized for Ann Arbor in June 2004, at which numerous work groups pursued cross-laboratory analyses and proposed further work. Investigators were advised to adopt more stringent criteria for high confidence peptide and protein identifications, and a commitment was made to collect raw spectra from the 18 laboratories that had submitted MS/MS or FT-ICR/MS datasets for independent analyses by three different groups. The datasets were moving targets, as some, but not all, labs submitted expanded or updated analyses, and about 15 laboratories completed "special projects" stimulated by HUPO PPP with a competition for small grants following the Montreal workshop.

The PPP provided participating laboratories with 1.0 mL of reference specimens of serum and plasma by three different methods of anticoagulation for plasma (EDTA, citrate, heparin) from specific donor pools. Investigators utilized their established and emerging technologies for fractionation and analysis of proteins. Investigators were encouraged to "push the limits" of their methods to detect and identify low abundance proteins. Comparisons of findings across laboratories provide a special opportunity for confirmation of protein identifications. Results were submitted to centralized bioinformatics functions at the University of Michigan and the European Bioinformatics Institute to create an integrated data repository from which PPP and other investigators could initiate further analyses and annotations. The approaches and core results have been presented at the US HUPO inaugural meeting in March 2005, the HUPO World Congress in Munich in August 2005, and at other meetings.

Here we present a comprehensive account of the major findings from the pilot phase of the Human Plasma Proteome Project, including the many associated special projects.

## 1.2 PPP reference specimens

The primary specimens were sets of four reference specimens prepared under the direction of the HUPO PPP Specimens Committee by BD Diagnostics for each of three ethnic groups: Caucasian-American (B1), African-American (B2), and Asian-American (B3). Each pool consisted of 400 mL of blood each from one male and one post-menopausal female healthy, fasting donor, collected into 10 mL tubes in a prescribed sequence (see Supplementary Protocol) after informed consent. Very large pools were rejected as requiring too prolonged specimen handling and processing unlike the collection of individual specimens; even a protocol for two males and two females proved to require more than the 2 h limit we set. Equal numbers of tubes and aliquots were generated with appropriate concentrations of $K_2$-EDTA, lithium heparin, or sodium citrate for plasma or permitted to clot at room temperature for 30 min to yield serum (with micronized silica as clot activator). The additives were dry-sprayed on the inner walls of the tubes, except for 1.0 mL of 0.105 M buffered sodium citrate, which gave a final ratio of 9:1 for blood to citrate in a 10 mL final volume, causing an 11% dilution of the blood. No protease inhibitor cocktails were used. This procedure required 2 h, mostly at 2 to 6°C. After centrifugation, volumes from the male and female donors in each donor pair for each specimen type were pooled and then aliquoted into numerous 250 μL portions in vials which were frozen and stored at −70°C. The centrifugation conditions with citrate consistently produced platelet-poor plasma (platelet count $<10^3/\mu L$). Aliquots tested negative for HIV, HBV, HCV, HTLV-1, and syphilis. We supplied four × 250 μL aliquots for each of the four plasma/serum specimens in each set. These vials were shipped on dry ice *via* courier in early May 2003 (and later to additional laboratories which petitioned to join the project, some of which could no longer be supplied the B1 set). No reshipping was permitted.

The Chinese Academy of Medical Sciences (CAMS) used a variant of the BD protocol to generate similar reference serum and plasma specimens, as described by Li *et al.* [5] and He *et al.* [6]. Pools were prepared after review by the CAMS Ethics Committee and informed consent by ten male and ten female donors in Beijing. Donors were fasting and avoided taking medicines or drinking alcohol for the 12 h before sampling. A subsequent pooling of 20 mL from each of the male and female serum or plasma specimens created the C1-CAMS PPP reference specimens which were sent to the 15 laboratories requesting these specimens after storage at −80°C. They were shipped on dry ice using the same courier in September 2003. C1-CAMS specimens were centrifuged originally, and then again upon thawing, at 4°C [6].

Finally, the UK National Institute of Biological Standards and Control (NIBSC) made available to the PPP their lyophilized citrated plasma standard prepared for hemostasis and thrombosis studies from a pool of 25 donors [1].

A standard questionnaire was sent to all laboratories expressing interest. Of 55 laboratories that originally committed to participate, 41 received the BD B1 specimens, 27 the B2 and B3 specimens, 15 the CAMS specimens, and 45 the NIBSC specimens. Laboratories varied on how many of the specimens they actually analyzed.

## 1.3 Bioinformatics and technology platforms

As intended, laboratories used a wide variety of methods, including multiple LC-MS/MS instruments, MALDI-MS, and FT-ICR-MS; depletion of abundant proteins; fractionation of intact proteins on 2-D gels or with LC or IEF methods; protein enrichment or labeling methods; immunoassays or antibody arrays; and direct (SELDI) MS. They also varied on choice of search algorithm and database, and criteria for declaring high or lower confidence identification of peptide sequences and matching proteins (Tab. 1). In general, the numbers of proteins reported individually by the labs do not have the integration feature which was applied to the whole PPP dataset. In several cases, much more extensive analyses were reported. Thus, many of the individual papers in this special issue have additional protein identifications not included in the project-wide dataset(s).

### 1.3.1 Constructing a PPP database for human plasma and serum proteins

Data management for this project included guidance and protocols for data collection, then centralized integration, analysis, and dissemination of findings worldwide *via* a communications infrastructure. As described in great detail by Adamski *et al.* [7, 8], key challenges were integration of heterogeneous datasets, reduction of redundant information to minimal identification sets, and data annotation. Multiple factors had to be balanced, including when to "freeze" on a particular release of the ever-changing database selected for the PPP and how to deal with "lower confidence" peptide identifications. Freezing of the database was essential to conduct extensive comparisons of complex datasets and annotations of the dataset as a whole. However, it complicates the work of linking findings of the current study to evolving knowledge of the human genome and its annotation. Many of the entries in the protein sequence database(s) available at the initiation of the project or even the analytical phase were revised, replaced, or withdrawn over the course of the project, and continue to be revised. Our policies and practices anticipated the guidelines issued recently by Carr *et al.* [9], as documented by Adamski *et al.* [7].

The 18 participating laboratories using MS/MS or FT-ICR-MS submitted a total of 42 306 protein identifications using various search engines and databases to handle spectra and generate peptide sequence lists from the specimens analyzed.

**Tab. 1** Protein identifications by lab, by specimen, and by methods

| Lab ID | Specimen | Depletion | Protein separation | Reduction/ alkylation | Peptide separation | Mass spectrum | Search software | 3020 High confidence | 3020 Lower confidence | Single peptide |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | b1-cit | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 61 | 39 | 12 |
| 1 | b1-edta | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 35 | 30 | 14 |
| 1 | b1-hep | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 50 | 38 | 13 |
| 1 | b1-serum | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 21 | 6 | 5 |
| 1 | b2-cit | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 57 | 37 | 12 |
| 1 | b2-hep | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 58 | 30 | 12 |
| 1 | b2-serum | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 59 | 31 | 12 |
| 1 | b3-serum | aig | none | iam | rp/scx/rp | esi-ms/ms_decaxp | PepMiner | 17 | 6 | 7 |
| 2 | b1-cit | none | cho affinity | iam | scx/rp | esi-ms/ms_qtof | SEQUEST | 165 | 79 | 94 |
| 2 | b1-serum | none | cho affinity | iam | scx/rp | esi-ms/ms_qtof | SEQUEST | 136 | 48 | 38 |
| 2 | nibsc | none | cho affinity | iam | scx/rp | esi-ms/ms_qtof | SEQUEST | 171 | 121 | 85 |
| 11 | b1-cit | none | cho affinity | iam | rp | esi-ms/ms_decaxp | SEQUEST | 59 | 4 | 9 |
| 11 | b1-edta | none | cho affinity | iam | rp | esi-ms/ms_decaxp | SEQUEST | 64 | 6 | 4 |
| 11 | b1-hep | none | cho affinity | iam | rp | esi-ms/ms_decaxp | SEQUEST | 62 | 9 | 15 |
| 11 | b1-serum | none | cho affinity | iam | rp | esi-ms/ms_decaxp | SEQUEST | 64 | 3 | 16 |
| 12 | b1-cit | aig | none | iam | rp/scx/rp | esi-ms/ms_deca | SEQUEST | 111 | 0 | 113 |
| 12 | b1-edta | aig | none | iam | rp/scx/rp | esi-ms/ms_deca | SEQUEST | 111 | 0 | 101 |
| 12 | b1-hep | aig | none | iam | rp/scx/rp | esi-ms/ms_deca | SEQUEST | 127 | 0 | 130 |
| 12 | b1-serum | aig | none | iam | rp/scx/rp | esi-ms/ms_deca | SEQUEST | 123 | 0 | 111 |
| 17 | b1-serum | aig | 1s sds | iam | rp | esi-ms/ms_lcq | SEQUEST | 50 | 19 | 7 |
| 21 | b1-cit | top6 | rotofor-ief/rp/1d-sds | iam | rp | esi-ms/ms_qtof | MASCOT | 40 | 0 | 1 |
| 21 | b1-cit | top6 | rotofor-ief/rp/1d-sds | none | none | maldi-ms/ms[a]bi4700 | MASCOT | 51 | 0 | 3 |
| 21 | b1-cit | top6 | rotofor-ief/rp/1d-sds | none | rp | esi-ms/ms_qtof | MASCOT | 39 | 0 | 1 |
| 21 | b1-edta | top6 | rotofor-ief/rp/1d-sds | iam | rp | esi-ms/ms_qtof | MASCOT | 40 | 0 | 1 |