Mass spectrometry-based methods for protein biomarker quantification
Klont, Frank

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER I

General introduction and scope of the thesis

Adapted from:
Frank Klont, Peter Horvatovich, Natalia Govorukhina, Rainer Bischoff. Pre- and postanalytical factors in biomarker discovery. Part of this chapter has been submitted to Brun et al. (eds.) for a book entitled ‘Proteomics for Biomarker Discovery’ in the series Methods in Molecular Biology (Springer).
The interest towards adopting indicators of the physiological state of a human being into clinical practice did not emerge in recent times as physicians millennia ago already adopted such markers of disease. Physicians of the Hippocratic School, for example, recognized specific tumors as *Karkinos*, the Greek word for crab, as corresponding swellings and ulcers resembled the shape of a crab with its claw-like projections.\(^1\)\(^2\) These same physicians, and Galen of Pergamon in particular, furthermore practiced medicine based on the theory of ‘humorism’ which proposes the clustering (and clinical assessment, accordingly) of illnesses in terms of excesses and deficiencies of the four bodily fluids, so-called humors (*i.e.* blood, yellow bile, black bile, and phlegm, which were associated with the heart, liver, spleen, and brain, respectively).\(^3\) Humoral medicine even continued to be one of the central principles of (Western) medical practice until the mid-nineteenth century, after which the humoral theory was abruptly abandoned and medicine could start to mature in times of numerous scientific discoveries and inventions.\(^2\)\(^3\)

Science and technology have been key drivers of medical advancements ever since the fall of humoral medicine and greatly expanded our knowledge of human (patho)physiology and the treatment options of physicians. In the past decades, ‘Omics’ technologies (*e.g.* genomics, proteomics) made their entrance in (bio)medical sciences and raised high expectations for the discovery of new biomarkers.\(^4\) Notable successes of genomics include *BRCA1/BRCA2* susceptibility testing for breast and ovarian cancer, viral load testing for diagnosing and monitoring human immunodeficiency virus (HIV) infection, genotyping and subtyping of chronic hepatitis C virus (HCV) infections based on the viral genome as well as guiding HCV treatment through (host) *ITPA* and *IL28A* genotype testing.\(^5\)\(^6\) Success stories of ‘proteomic medicine’ are less pronounced, though the CKD273 biomarker panel, which received a Letter of Support from the United States Food and Drug Administration (FDA) in 2016 encouraging the use of this panel in the (early) management of chronic kidney disease,\(^7\) and the OVA1 *in vitro* diagnostic multivariate index assay (IVDMIA), which received FDA-clearance for assessing the risk of ovarian cancer in women presenting with pelvic masses, are noteworthy examples in this respect.\(^8\)\(^9\) The CKD273 panel is a classifier based on 273 urinary peptides which were identified and assessed using capillary electrophoresis coupled to mass spectrometry. The OVA1 test integrates the serum levels of cancer antigen 125 (CA125), transthyretin, apolipoprotein A1, \(\beta_2\)-microglobulin, and transferrin. Evidence to support the rationale of this combination was mainly based on protein expression profiles obtained using the Surface-Enhanced Laser Desorption/Ionization-Time Of Flight (SELDI-TOF) mass spectrometric (MS) platform.\(^10\)\(^11\) While the usage and applicability of both tests may (currently) be limited as compared to the genomic tests mentioned above, both tests provide good examples of successful proteomics-based biomarker discovery and development and are testament to the opportunities of proteomics research in this area.
The combination of SELDI-TOF and ovarian cancer in the OVA1 example represents a success story of MS-based proteomics, yet this combination will also be remembered in the context of a major controversy which casted a shadow over the biomarker field of research in the first decade of this millennium. Central to this controversy was a study published in The Lancet in 2002, which reported a SELDI-TOF MS-based blood test for the detection of ovarian cancer with 100% sensitivity and 95% specificity. The test stirred hope for the prospect of early-stage cancer detection, though it also stirred concern about the test’s reliability. Reanalysis of the original data, which were deposited in an open access repository, yielded rather dissatisfactory results leading to the conclusions that “the ability to discriminate between cancer and control subjects ... reveals the existence of a significant non-biologic experimental bias between these two groups” and that “the reproducibility of the proteomic profiling approach has yet to be established.” Eventually, the reported test did not reach the clinical chemist’s ‘toolbox’ and thus was not applied for the benefit of patients. The corresponding controversy did, however, raise awareness that bias and lack of generalizability (e.g. statistical overfitting of the data) are potential threats to the validity of biomarker-based research findings and furthermore stressed the need for critically assessing pre- and postanalytical factors in biomarker development research.

1.1. REGULATED BIOANALYSIS

Good quality analytical methods form the basis of the discovery potential of proteomics workflows and are furthermore a key success factor of efforts addressing the later stages of the biomarker development pipeline (i.e. qualification, verification, and validation). Research dealing with these later stages mostly comprise targeted proteomics endeavors following regulatory guidelines (e.g. FDA, European Medicines Agency (EMA), Clinical & Laboratory Standards Institute (CLSI)), which have become well-rooted in corresponding practices. These guidelines have the aim to minimize inter-laboratory variance by adopting consensus criteria with respect to assay performance. In particular, recommendations are provided for addressing analytical quality attributes like accuracy, precision, sensitivity, and recovery during method validation, but also some preanalytical factors, such as sample stability (e.g. storage, benchtop, freeze-thaw) and specific matrix effects (e.g. hemolysis, icterus, and lipemia, which respectively are attributed to ruptured red blood cells, bilirubin, and lipoprotein particles). Adhering to these guidelines does not guarantee the quality or usefulness of a biomarker assay, since other, non-addressed (pre)analytical variables may have a major impact on whether the method and the experimental design are suitable for addressing the study goal. Regulatory documents accordingly are living documents that are regularly updated often
based on deliberations following expert workshops and panel discussions.\textsuperscript{25} Furthermore, journals adopting requirements with respect to validation of analytical methods in an effort to raise the quality standard of published methods, is a positive development with respect to the applicability and reproducibility of scientific research and the corresponding findings, that is not limited to biomarker-related work.\textsuperscript{26,27}

While strict regulation and standardization are more difficult to implement in discovery-based proteomics as compared to targeted proteomics, it is conceivable that ‘Good Proteomics Practice’ guidelines will emerge within the foreseeable future. One example of this type of document is the Human Proteome Project Data Interpretation Guideline, which provides guidance on how to interpret fragment-ion-based mass spectrometry data for peptide and protein identification. In fact, this guideline is now mandatory for manuscripts submitted to a number of proteomics journals.\textsuperscript{28} Although recommendations and standardized procedures aiming to set quality standards for biomarker discovery research have been proposed, there is currently no consensus on their large-scale implementation.

1.2. THE PREANALYTICAL PHASE

Providing (consensus-based) guidance for adequately dealing with analytical variables during method development and validation is a complex task. To illustrate this, the FDA draft guidance document released in September 2013, which was intended to replace its predecessor from 2001, was only finalized in May 2018, more than 4.5 years after its initial release.\textsuperscript{19,29,30} Considering that these documents deal with approximately ten analytical variables which ought to be addressed during method validation, providing guidance for preanalytical variables will be considerably more challenging as these easily outnumber the analytical ones.\textsuperscript{31}

1.2.1. Presampling factors

There are dozens of physiological and environmental factors that may affect laboratory results, which are generally condensed into terms like ‘biological’, ‘inter-individual’, and ‘between-subject variation’ or ‘variability’ in biomarker development research.\textsuperscript{18,22,23,32} These terms explain the increased variation in biomarker levels that are observed when moving from early-stage, small-scale monocentric discovery studies to the advanced biomarker qualification level, where more heterogeneous and larger populations are studied across multiple clinical centers. Factors like age, gender, circadian rhythm, seasonal changes, altitude, menstruation, pregnancy, and lifestyle may play a role with some of them being rather difficult to control.\textsuperscript{31,33} While groups are generally matched with respect to gender and age in biomarker studies, other factors may be equally or even more relevant. As an example, our department found considerable changes in
peroxiredoxin 1, uteroglobin, and serpin B3 levels in pulmonary epithelial lining fluid within 3 hours after cigarette smoking on the basis of a quantitative proteomic analysis employing isobaric tags for relative and absolute quantification (iTRAQ) followed by the validation of findings using commercial immunoassays. In one of the groups included in this study (i.e. old COPD patients), cigarette smoking led to increases in uteroglobin levels of ten and three times, as determined by the proteomics- and immunoassay-based analyses, respectively. Considering these proteins as disease biomarkers for COPD would thus necessitate tight control of the smoking history prior to sampling, which may be rather difficult in practice.

1.2.2. Sampling factors
A group of experts in the field of clinical proteomics recommended in a perspective paper in 2010 that detailed descriptions of (appropriate and consistently applied) sampling parameters ought to be provided in publications, since the quality of samples and corresponding results may otherwise be compromised. While this recommendation is rather difficult to comply with, notably for already acquired, biobanked samples, it puts the focus on potential sampling errors which may lead to spurious findings. When, for example, studying the HUPO Plasma Proteome Project specimen collection and handling recommendations published in 2005, it becomes apparent that the list of critical sampling factors, in this case related to blood-based samples, is quite extensive. For each of these sampling factors, either related to venipuncture (e.g. needle gauge), phlebotomy (e.g. tourniquet technique, patient position), or collection device (e.g. tube versus bag, glass versus plastic, presence versus absence of protease inhibitors), there are numerous examples of biomarkers that are affected by corresponding changes in sampling conditions. Such variables are often not controlled or standardized in proteomics research, since many projects target (long-term) stored samples for which sampling conditions were fixed when designing the study or have not been documented in the necessary detail.

1.2.3. Sample processing and storage factors
Preanalytical factors related to sample processing and storage are more tangible compared to factors addressed in the previous sections. Unintentional mistakes can be made after samples have been taken (e.g. sample contamination, sample spills, improper labeling, inadequately following protocols, ‘forgetting’ to process the samples in time, ‘losing’ samples), and some conditions may lead to unintentionally compromised sample integrity (e.g. exposure to sunlight or moisture, the use of secondary vials with unfavorable (adsorptive) properties, temperature fluctuations in case of sample shipment, power outages, or freezer break-down and maintenance). In particular when dealing with large numbers of samples that are stored
for extended time periods (e.g. in biobanks), it is important to have a tight quality control and to document events that may affect sample quality.36,37

1.3. PRACTICAL CONSIDERATIONS

The complexity of Omics-based biomarker studies is daunting and repeated failures of so-called ‘biomarker candidates’ to translate into useful, robust clinical assays resulted in a certain skepticism and sometimes even an outright negative attitude towards performing such studies at all. While there is probably no single study that is perfect in all respects, these challenges should motivate researchers to establish and subsequently work according to standards through which the risk of bias is mitigated. Adopting a ‘Quality-by-Design’ (QbD) concept, as originally proposed by Joseph M. Juran, may function as a safeguard against potential errors in biomarker discovery and development research and thus increase the success rate.38 Lessons may furthermore be learned from the pharmaceutical industry, where stringent standards on documenting, managing, and reporting deviations are the rule. Documenting protocol deviations and violations yet also potential weaknesses in experimental design are also very helpful in biomarker research, for example for adequately interpreting (unexpected) findings. Openness to reporting such information or to sharing all experimental details should be advocated, as this will allow other scientists as well as reviewers and readers of the scientific literature to adequately draw conclusions. In case of the ovarian cancer example, more openness regarding the samples, experimental design and data processing procedures in the initial article might have limited the extent of the ensuing controversy, or at least would have prevented the assay developers from stating that “inappropriate conclusions drawn … could have been avoided by communication between the producers and consumers of the data”13,39”.

1.4. SCOPE OF THE THESIS

This thesis aims to contribute to the advancement of promising protein biomarker candidates as well as mass spectrometry (MS-)based methodologies towards clinical implementation. In addition, this thesis puts focus on establishing sample preparation methods for protein biomarkers based on a rational design, but also addresses limitations and potential sources of bias arising from sample preparation methodologies.

The recognition of mass spectrometry as relevant clinical assay platform is still hampered by the complexity and the relatively low throughput of corresponding workflows and instrumentation. Admittedly, matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometers are easy to use and have considerable high-throughput
capabilities, though this technique is often considered to be non-quantitative in nature. MALDI-TOF can, however, be used in quantitative biomarker assays as is shown in Chapter 2 which presents a fully validated, semi-automated MALDI-TOF MS-based method for quantification of insulin-like growth factor 1 (IGF1) in the range of 10-1000 ng/mL in human plasma. IGF1 is a well-established biomarker for growth hormone-related disorders and is traditionally quantified using ligand binding assays in clinical laboratories. In the past decade, concerns were raised with respect to the reliability of these IGF1 assays, which set the stage for the development of mass spectrometric alternatives. Several liquid chromatography-mass spectrometry (LC-MS) methods for reliable IGF1 quantification were accordingly developed, though their roles in clinical practice are yet limited due to protocol complexity and throughput constraints. Simplicity and high throughput capabilities are, in turn, key quality attributes of MALDI-TOF mass spectrometers, hence MALDI-TOF mass spectrometry would represent a truly relevant clinical assay platform, as is outlined in this chapter. This chapter furthermore discusses the conditions that should be met to quantify biomarkers using MALDI-TOF MS and proposes strategies to ensure that high quality data are acquired by this type of mass spectrometer.

Sample preparation is a major challenge when developing MS-based biomarker assays given that most proteins of interest reside at low or sub ng/mL levels in biological samples and thereby cannot readily be detected by mass spectrometry without removing other, higher abundant proteins which hamper their detection. Strategies to enrich proteins from their complex matrix are accordingly being deployed to facilitate the mass spectrometric detection of these proteins. Many of such strategies involve the use of antibodies that are coupled to bead-based solid supports which offer protocol flexibility and automation possibilities yet also come along with considerable costs. In Chapter 3, a cheaper, faster, and simpler alternative to bead-based immunoaffinity enrichment strategies is presented that relies on passive adsorption of antibodies to microtiter plates, which belong to the standard equipment of any analytical or clinical laboratory. The potential of this strategy is demonstrated by means of an LC-MS method for quantification of the low abundant soluble receptor of advanced glycation end-products (sRAGE) in the clinically relevant range between 100 pg/mL and 10 ng/mL in human serum. sRAGE is a decoy receptor for various pro-inflammatory proteins, notably in the lungs, and is considered to be a highly promising biomarker candidate for chronic obstructive pulmonary disease (COPD) based on data from several large-scale clinical studies. This disease is characterized by chronic bronchitis (i.e. airway inflammation) and emphysema (i.e. destruction of alveoli, the air sacs in the lung) and currently represents the number three cause of death worldwide. Despite the high prevalence and staggering burden of COPD, clinicians are left without disease-specific laboratory markers to assist in (early) diagnosis and
disease management, and clinicians furthermore lack therapeutic options to treat this disease rather than treating its symptoms. Accordingly, sRAGE was a highly relevant analytical target for demonstrating the applicability of our immunoaffinity enrichment strategy.

Antibodies are considered to be the workhorses of biomedical experiments as is, for example, exemplified by their wide-scale application in quantitative protein assays. Antibodies are, however, not without limitations. Non-antibody affinity ligands are accordingly being developed to provide researchers with versatile and robust antibody alternatives. Chapter 4 presents the first-time application of affimers, a novel class of affinity binders, in regulated bioanalysis and addresses the capability of custom anti-sRAGE affimers to replace antibodies in an LC-MS method for quantification of sRAGE in human serum at clinically relevant levels.

Employing antibodies, affimers, or other affinity ligands to enrich low abundant proteins represents a convenient approach to quantify these proteins by LC-MS. For some of these proteins, affinity ligand-free procedures can turn out to be viable alternatives to affinity-based procedures which may suffer from batch-to-batch variability of the affinity ligands or which may be susceptible to interference from potential binding proteins. Chapter 5 describes the application of strong cation exchange (SCX) solid-phase extraction (SPE) for the enrichment of sRAGE, which has a neutral isoelectric point (pI) at first glance but a distinct bipolar charge distribution upon closer inspection. The presented method involves SCX-based enrichment at the protein level under highly basic conditions to achieve an adequate degree of sample cleanliness thereby allowing for quantification of sRAGE in the low to sub ng/mL range in serum.

Different sample preparation methods can yield different protein levels as measured by targeted proteomics workflows or lead to different information being acquired in discovery proteomics experiments. To highlight the impact of the selected method on the outcome of proteomic analyses, Chapter 6 presents a comparison of commonly-used sample preparation methods in mass spectrometry-based proteomics. The methods are compared on the basis of peptide and protein losses, precision of quantification, discovery potential, and the distribution of physicochemical properties of identified proteins and peptides, thereby aiming to underline the relevance of establishing sample preparation methods based on a rational design.

At last, Chapter 7 summarizes the findings and corresponding interpretations that are outlined in this thesis and discusses future perspectives of mass spectrometry-based protein analysis in a clinical setting.
1.5. REFERENCES


