Serum insulin-like growth factor I quantitation by mass spectrometry: insights for protein quantitation with this technology

Richard Kin Ting Kam, Chung Shun Ho, Michael Ho Ming Chan

Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong SAR, China

ARTICLE INFO

Corresponding author:
Dr. Richard K.T. Kam
Scientific Officer
Department of Chemical Pathology
The Chinese University of Hong Kong
Prince of Wales Hospital
Shatin New Territories
Hong Kong SAR, China
Email: richardktkam@cuhk.edu.hk

Key words:
insulin-like growth factors,
liquid chromatography mass spectrometry,
MS-based protein quantitation,
clinical test methodology

ABSTRACT

Liquid chromatography mass spectrometry (LC-MS) is a widely used technique in the clinical laboratory, especially for small molecule quantitation in biological specimens, for example, steroid hormones and therapeutic drugs. Analysis of circulating macromolecules, including proteins and peptides, is largely dominated by traditional enzymatic, spectrophotometric, or immunological assays in clinical laboratories. However, these methodologies are known to be subjected to interfering substances, for example heterophilic antibodies, as well as subjected to non-specificity issues.

In recent years, there has been a growing interest in using LC-MS platforms for protein analysis in the clinical setting, due to the superior specificity compared to immunoassay, and the possibility of simultaneous quantitation of multiple proteins. Different analytical approaches are possible using LC-MS-based methodology, including accurate mass measurement of intact molecules, protein digestion followed by detection of proteolytic peptides, and in combination with immunoaffinity purification. Proteins with different complexity, isoforms, variants, or chemical alteration can be simultaneously analysed by LC-MS, either by targeted or non-targeted approaches. While the LC-MS platform offers a more specific determination of...
proteins, there remain issues of LC-MS assay harmonization, correlation with current existing platforms, and the potential impact in making clinical decision.

In this review, the clinical utility, historical aspect, and challenges in using LC-MS for protein analysis in the clinical setting will be discussed, using insulin-like growth factor (IGF) as an example.

Abbreviations (in alphabetical order):

- **EQA**: External Quality Assurance
- **GH**: Growth hormone
- **IGF**: Insulin-like growth factor
- **LC-MS**: Liquid chromatography mass spectrometry
- **m/z**: Mass-to-charge ratio
- **MALDI**: Matrix-assisted Laser Desorption/Ionization
- **MRM**: Multiple Reaction Monitoring
- **MSIA**: Mass spectrometric immunoassay
- **RCPA**: The Royal College of Pathologists of Australasia
- **SPE**: Solid-phase extraction
- **SRM**: Single Reaction Monitoring
- **TOF**: Time-of-flight
- **US FDA**: United States Food and Drug Administration

**INTRODUCTION**

Liquid chromatography mass spectrometry (LC-MS) is an analytical technique in which analytes are separated by their physical properties on a chromatographic stationary phase, followed by detection based on their specific molecular mass to charge ratio (m/z). Separation of the analytes from sample matrix interferences greatly enhances the robustness and sensitivity of the LC-MS assay. Specificity is provided by the characteristic retention times on a chromatographic column, the exact mass to charge (m/z) values of the parent ion, and m/z values of the fragment ions.

The advantage of specific detection of analytes is valuable in laboratory medicine, especially in the determination of structurally similar compounds such as steroid hormones, for therapeutic drug monitoring, and for toxicology screening (1–3). Also, immunoassay may be subjected to the heterophilic antibody interference. Indeed, thyroglobulin quantitation by immunoassay was known to be subjected to the presence of anti-thyroglobulin autoantibodies which is commonly found in patients with thyroid cancer (4). LC-MS-based assays provide a specific and alternative platform for the detection of thyroglobulin. Chromatographic separation of analytes also allows simultaneous determination of multiple molecules in a single analytical run, thereby reduces both sample consumption and turnaround time.

Although powerful in small molecule analysis, applications of the LC-MS technology for protein quantitation in clinical laboratory was until recently relatively lacking. The limitations were due to a lack of LC-MS assay harmonization and standardization, expensive LC-MS platform instrumentation and method development. Due to the availability of automated instrumentation in the clinical laboratory, quantitation of proteins has been dominated by enzymatic, spectrophotometric, and immunological assays, which also defined the reference interval of circulating proteins for clinical management. Most of the LC-MS assays in clinical laboratories were developed in-house based on a “fit-for-purpose” approach, i.e. the developed assay was considered applicable for clinical purpose based on the assay linearity, imprecision, accuracy and robustness, among other validation parameters. While LC-MS methods are considered or proposed
as reference methods for several proteins including urine albumin and whole blood Hb\textsubscript{A\textsubscript{1c}} (5,6), they are restricted to selected reference laboratories and not available in routine setting. LC-MS-based platform was considered more cost-effective compared to traditional analyzer due to the negligible reagent cost (7). However, the initial capital investment in LC-MS instrument, personnel, training, and the method development and validation process might be prohibitive. In spite of these challenges, the application of LC-MS-based protein assay in clinical laboratory, as well as in clinical trials, is becoming more popular (8).

In this review, the clinical utility and historical aspects of LC-MS-based Insulin-like Growth Factor I (IGF-I) assay in clinical laboratory are reviewed and discussed. The anticipated challenges of the platform in future application for other circulating proteins will be discussed.

**BIOCHEMISTRY AND CLINICAL UTILITY OF INSULIN-LIKE GROWTH FACTORS**

Insulin-like growth factor (IGF) is a class of circulatory peptide hormones identified in 1957 and subsequently described as “non-suppressible insulin-like activity” in 1963 (9,10). Two members of IGF family, IGF-I and IGF-II, share a similar domain structure as well as sequence homology with insulin, and are able to bind with insulin receptor on cell surface (11–13).

Insulin-like growth factor 1 (IGF-I) is a 70-amino-acid single chain polypeptide which is synthesized and secreted by the liver in response to pituitary growth hormone (GH). While the precursor, pre-pro-IGF-I, is structurally variable and may be composed of two classes of C-terminal signal peptides and three classes of N-terminal E-domains, the mature IGF-I in circulation is a conserved polypeptide with defined A, B, C and D-domains (12). Because of these well-defined properties, it is possible to generate a pure preparation of mature IGF-I as a reference material. The National Institute of Biological Standards and Control (NIBSC) has thus established the first WHO International Standard for IGF-I for immunoassay in 2008 with defined measurement uncertainty (NIBSC code 02/254).

IGF-II, on the other hand, is a 67-amino-acid single chain polypeptide which share approximately 70% homology with IGF-I (14). The circulating IGF-II level is independent of gender- and age-specific variation, and is relatively stable compared to that of IGF-I (15). The determination of IGF-II is essential for calculating IGF-II/IGF-I ratio, which is a useful parameter for the investigation of non-islet cell tumor hypoglycemia (NICTH) (16).

Circulating levels of serum total IGF-I are widely used in the diagnosis of GH disorders, including acromegaly and GH deficiency. Pituitary GH is known to be secreted on a pulsatile manner, in diurnal rhythm, and is subjected to different physiological and environmental stimuli including fasting, exercise and feeding. The short half-life of GH (20 minutes) further increases the variability of the circulating level. On the other hand, IGF-I is synthesized in a more stable manner, does not exhibit diurnal rhythm, has a longer half-time, and therefore is a more reliable biomarker of GH disorders (17). To highlight the importance of IGF-I in the management of GH disorders, it was recommended by the Endocrine Society that serum level of IGF-I should be used as a first line screening test for acromegaly, followed by GH measurement with an oral glucose loading as a confirmation test. The management goal of acromegaly was also established biochemically by normalization of serum IGF-I level. Similarly, the Endocrine Society also recommended using IGF-I level to diagnose, to document, and for treatment monitoring of persistent GH deficiency (18)(19).
IGF-I QUANTITATION BY IMMUNOASSAY

Traditionally, IGF-I was measured by using commercial immunoassay. However, due to the reported lot-to-lot reagent variation in IGF-I assay and the global supply disruption of immunoassay kit, the reliability of immunoassay for GH disorders was questioned (20). Patients with GH disorders required regular measurement of serum IGF-I for management and treatment monitoring. It is important to have a consistent and reliable assay for long term patient management.

At least 2 types of IGF-I sequence variants were reported. One of these variants was confirmed to be pathogenic (21). Apart from the issue of assay stability, immunoassay is not able to differentiate between wild-type IGF-I and genetic variants of IGF-I. Failure to differentiate between sequence variants may cause a falsely high measured value of wild-type, bioactive IGF-I. Patient management may be affected if pathogenic IGF-I variants were not promptly identified. Therefore, an alternative non-immunological platform for serum IGF-I quantitation would be valuable.

LC-MS-BASED PLATFORM FOR IGF-I MEASUREMENT

An alternative platform considered for serum IGF-I quantitation is LC-MS. While the use of LC-MS in IGF-I quantitation in the clinical setting is relatively new, the method was widely investigated in the past decade for sport science and doping control. The assays were first reported in 2001 by de Kock et al. and by Bobin et al. Two approaches exist for LC-MS analysis of proteins. In “top-down” approach, the ions entering the MS instrument carry the complete amino acid sequence information of the respective intact protein. In most cases the ions being analyzed are intact protein without proteolysis by enzymatic or chemical method. On the other hand, for “bottom-up” approach, the ions entering the MS instrument only carry partial amino acid sequence of the intact protein. The ions are usually peptides generated by protease digestion, each representing a fragment of the intact proteins (Figure 1). Both determination of proteolytic peptides after enzymatic digestion (“bottom-up” approach) and analysis of intact protein (“top-down” approach) were described for the characterization and quantitation of IGF-I respectively (22,23).

BOTTOM-UP APPROACH

The bottom-up approach is based on the assumption that the generation of proteolytic peptides are stoichiometrically related to the parent proteins. By quantitating the proteolytic peptides, the concentration of parent proteins can be derived. de Kock et al. reported the use of endoproteinase Glu-C and Asp-N for the generation of peptide mass fingerprint and subsequent MS/MS analysis of peptide fragments for the characterization of IGF-I. The bottom-up approach was also reported by Kirsch et al. in 2007, and Kay et al. in 2009 (24,25). In the report by Kirsch et al., tryptic peptides were generated from human plasma, with the addition of a stable isotope-labeled peptide as the internal standard, followed by analysis in a single reaction monitoring (SRM) experiment. Kay et al. adopted a similar approach, but introduced an acetonitrile precipitation procedure in order to enrich IGF-I prior to digestion. This approach was further elaborated by the same group in 2013 (24). Instead of acetonitrile precipitation, an offline SPE device was used for IGF-I enrichment, followed by trypsin digestion and SRM experiment.

The generation and measurement of proteolytic peptides offered a unique advantage over top-down method in terms of analytical simplicity. Proteins with diverse physical properties and
molecular weights were converted into a pool of peptide mixture, which was relatively similar in terms of solubility, chromatographic behavior, molecular weight and ionization efficiency. A general analytical approach was capable of simultaneously determining multiple proteolytic peptides, and thus the parent proteins. This was reported by Such-Sanmartín et al. in 2015 in which five proteins (IGF-I, IGF-II, two IGF binding proteins and leucine-rich alpha-2-glycoprotein), with diverse molecular weights and glycosylation states, were simultaneously quantitated for anti-doping analysis and for a cancer study (26).

To reliably quantitate an analyte by LC-MS, several information was necessary including retention time on chromatographic column, accurate mass of parent ion, and accurate mass of fragment ions after fragmentation. The bottom-up approach utilized triple quadrupoles MS in MRM mode for the detection of proteolytic peptides. Similar to small molecule analysis, this approach has been well characterized and the instrument can accommodate the m/z range of peptides and their fragment ions. In terms of instrumentation, clinical laboratories equipped with triple quadrupoles MS can adopt the bottom-up approach readily.

However, while the bottom-up approach provides a relatively general analytical platform for diverse range of protein, it requires a more stringent quality control protocol to monitor analytical variation arising from enzymatic digestion to peptide purification. This was highlighted by the observation that all three groups generated the calibration curve from pooled plasma via standard addition approach in order to ensure the reproducibility of enzymatic digestion in
the endogenous matrix. Since the endogenous concentration of IGF-I cannot be certified externally, the spike-in method may also undermine the accuracy of the bottom-up assay.

**TOP-DOWN APPROACH**

Bobin et al. reported the use of ESI-ion-trap with deconvolution to measure the intact IGF-I neutral mass for quantitation, as well as the use of matrix assisted laser desorption ionization (MALDI) - time of flight (TOF) for the identification of an oxidized IGF internal standard by the “top-down” approach (14). In plasma sample analysis, the group utilized immunoaffinity column for sample purification, and successfully quantitated IGF-I level in equine plasma samples with +/- 10 ng/ml deviation from nominal value obtained from immunoassay. The work demonstrated a comparable quantitative performance between immunoassay and MS assays, and was also the first application of mass spectrometric immunoassay (MSIA) for IGF-I analysis in biological samples.

Other top-down quantitative approaches were subsequently reported by Nelson et al. in 2004 (27), and Bredehoft et al. and Popot et al. in 2008 (28, 29). The work by Nelson et al. and Bredehoft et al. further expanded the workflow of MSIA by modifying it into a pipette tip format and antibody-coated magnetic beads, respectively. Popot et al. derived the deconvoluted peak area in the mass chromatogram using an in-house protein transposing model for the quantitation of IGF-I in horse plasma, with better precision and accuracy compared to the use of deconvoluted peak height alone. An overview of LC-MS-based IGF-I assays revealed that immunoaffinity extraction was necessary for the top-down approach to achieve sufficient sensitivity and precision for quantitation, potentially due to the limitation of instrumentation at that period. However, the use of antibodies may affect the recovery of the analyte for LC-MS detection, as well as introducing a variable to the analysis.

In 2011, a dilute-and-shoot top-down approach, without the use of immunoaffinity extraction, was reported by Bystrom et al., using a simple acid-alcohol extraction process and online solid-phase-extraction (SPE) as the sample preparation method (30). By using a high resolution TOF-MS, a mono-isotopic peak of IGF-I within the [M+7H]$^+$ cluster was monitored by narrow mass extraction, which generated a specific mass chromatogram for quantitation. A limit of quantitation of 15.6 ng/ml was achieved without using immunoaffinity purification, and interference was detected by using the isotope ratio of IGF-I. Comparatively, the limit of quantitation was reported as 50 ng/ml by Bredehoft et al. in 2008. These were the first few reports on the use of IGF-I LC-MS assay for clinical purpose. LC-MS-based age- and gender-specific reference interval, as well as the assay comparison with immunoassay, was subsequently published by the same group in 2012, thereby providing the fundamentals of LC-MS-based IGF-I assay in clinical laboratory (15). This reported method was also adopted by Hines et al. in 2015 using Orbitrap, a high resolution mass analyzer, for clinical application (31).

The use of top-down approach in routine clinical laboratory setting is preferred since it utilizes a simpler sample preparation procedure and therefore introduces less variation, compatible with liquid handling robot for automation, and ultimately can accommodate a higher sample volume. However, the accurate mass information on fragment ions is usually absent. The instrument’s relatively slower MS/MS duty cycle may not be able to provide sufficient data points for quantitation, and the m/z range of intact protein and their fragment ions may exceed that of the instrument. Multiplex analysis was also limited to proteins with similar physical properties,
for example IGF-I and IGF-II. Therefore, the top-down approach requires a high-resolution mass analyzer to provide high mass accuracy of parent ions, as well as to resolve single isotopic peaks of multiple charged proteins to provide specificity of quantitation. With the advance of high resolution MS in recent years including TOF and Orbitrap instrument, it is expected that the top-down approach will be more commonly used and affordable for the clinical laboratory.

An overview on the imprecision and linear range of IGF-I assays reported in the literature is provided in Table 1.

DETECTION OF SEQUENCE VARIANTS AND OXIDIZED PROTEINS

While mature IGF-I is a non-glycosylated protein and is relatively conserved in its amino acid sequence, two types of variants have been reported independently using the top-down approach. Hines et al. in 2015 reported an A70T IGF-I variant in a cohort of 1720 samples, with an estimated prevalence of 0.6% in the studied population (31). The variant was identified from the outliers in a correlation study with immunoassay platform, in which subjects with either a heterozygous or homozygous variant showed 50% and 100% reduction in wild-type IGF-I compared to immunoassay respectively. The other variant was suggested by Oran et al. in 2014 to be the A67T IGF-I variant, which was found to be present in 9 out of 1054 (0.85%) samples (32). Using MSIA in pipette tip format in combination with an automatic liquid handler, the group used MALDI-TOF MS for the top-down quantitation of IGF-I, and the A67T variant was identified as a twin peak with a shift of 30 m/z unit; however, this A67T variant was not confirmed by genetic sequencing. It should be noted that the antibodies reported was unable to differentiate between wild-type IGF-I and these two variants, as the immunoassay utilized in Hines et al. study quantified wild-type and A70T variants as a single entity, while the antibody used in MSIA in Oran et al. study captured A67T for MS analysis as well. Interestingly, by using the same principle, Oran et al. also speculated a glycosylated IGF-I in the circulation in the same study with a mass shift of approximately 700 Da, although it was not confirmed by tandem MS/MS fragmentation or other structural analysis.

Other pathogenic changes in the IGF-I sequence have been identified previously. A partial deletion of IGF-I in exons 4 and 5 was described in 1996, causing growth failure and mental retardation (33). The deletion rendered the mutated IGF-I undetectable by radioimmunoassay. IGF-I with a single nucleotide polymorphism of T-to-A transversion in the untranslated region of exon 6 was described in 2003, which causes dysregulation of IGF-I mRNA processing and IGF-I deficiency (34). A V44M variant was identified in 2005 which reduced the affinity of IGF-I to IGFBP-3 by 90 folds, leading to severe growth and mental retardation (21). Immunoassay was not able to differentiate between the V44M variant and wild-type IGF-I.

Protein and peptide were known to be susceptible to chemical modification, including oxidation of methionine, and carbamylation of lysine and N-terminal in the present of urea (35,36). Cory et al. in 2012 reported the present of significant amount of oxidized IGF-I in the commercial QC pools, which exhibited a shift of 2.285 m/z value and was not observed in human serum samples (15). Our group also observed the presence of oxidized IGF-I in The Royal College of Pathologist of Australasia (RCPA) External Quality Assurance (EQA) materials, which was absent in human serum samples (data not showed).

The presence of protein variants and chemically altered proteins represented an important group of analytes of unknown significance. The
### Table 1: Overview on the imprecision and linear range of LC-MS-based IGF-I assay reported in literature

<table>
<thead>
<tr>
<th>Year</th>
<th>MSIA</th>
<th>Quantitative approach</th>
<th>Calibration curve</th>
<th>Max. imprecision of quantitation</th>
<th>Linear range (ng/ml)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>N</td>
<td>Top-down</td>
<td>Pure standard</td>
<td>Not available</td>
<td>30 - 500</td>
<td>(23)</td>
</tr>
<tr>
<td>2001</td>
<td>N</td>
<td>Top-down</td>
<td>Pure standard and standard addition in plasma</td>
<td>Not available</td>
<td>310 - 1480</td>
<td>(22)</td>
</tr>
<tr>
<td>2004</td>
<td>Y</td>
<td>Top-down</td>
<td>Pure standard</td>
<td>Not available</td>
<td>7.8 - 1000</td>
<td>(27)</td>
</tr>
<tr>
<td>2007</td>
<td>N</td>
<td>Bottom-up</td>
<td>Standard addition in serum or plasma</td>
<td>30.00%</td>
<td>2000 - 8000</td>
<td>(24)</td>
</tr>
<tr>
<td>2008</td>
<td>Y</td>
<td>Top-down</td>
<td>Standard addition in serum or plasma</td>
<td>12.00%</td>
<td>50 - 1000</td>
<td>(28)</td>
</tr>
<tr>
<td>2008</td>
<td>Y</td>
<td>Top-down</td>
<td>Known plasma samples</td>
<td>8.05%</td>
<td>318 - 898</td>
<td>(29)</td>
</tr>
<tr>
<td>2009</td>
<td>N</td>
<td>Bottom-up</td>
<td>Standard addition in serum or plasma</td>
<td>17.00%</td>
<td>15.6 - 2000</td>
<td>(25)</td>
</tr>
<tr>
<td>2011</td>
<td>N</td>
<td>Top-down</td>
<td>Standard addition in artificial matrix</td>
<td>5.20%</td>
<td>15.6 - 2000</td>
<td>(30)</td>
</tr>
<tr>
<td>2012</td>
<td>N</td>
<td>Top-down</td>
<td>Standard addition in artificial matrix</td>
<td>6.50%</td>
<td>15 - 2000</td>
<td>(15)</td>
</tr>
<tr>
<td>2012</td>
<td>Y</td>
<td>Top-down</td>
<td>Standard addition in urine matrix</td>
<td>18.50%</td>
<td>Not available</td>
<td>(43)</td>
</tr>
<tr>
<td>2013</td>
<td>N</td>
<td>Bottom-up</td>
<td>Standard addition in serum or plasma</td>
<td>Not available</td>
<td>254 - 4230</td>
<td>(44)</td>
</tr>
<tr>
<td>2013</td>
<td>Y</td>
<td>Bottom-up</td>
<td>Pure standard</td>
<td>7.36%</td>
<td>1 - 1500</td>
<td>(45)</td>
</tr>
<tr>
<td>2013</td>
<td>N</td>
<td>Top-down</td>
<td>Standard addition in human or chicken whole blood</td>
<td>11.00%</td>
<td>50 - 600</td>
<td>(46)</td>
</tr>
<tr>
<td>2014</td>
<td>N</td>
<td>Bottom-up</td>
<td>Standard addition in rat serum</td>
<td>5.60%*</td>
<td>100 - 1000</td>
<td>(42)</td>
</tr>
<tr>
<td>2014</td>
<td>N</td>
<td>Top-down</td>
<td>Standard addition in rat serum</td>
<td>4.80%</td>
<td>50 - 1000</td>
<td>(47)</td>
</tr>
<tr>
<td>2014</td>
<td>Y</td>
<td>Top-down</td>
<td>Standard addition in artificial matrix</td>
<td>9.75%</td>
<td>5 - 1500</td>
<td>(32)</td>
</tr>
<tr>
<td>2015</td>
<td>N</td>
<td>Bottom-up</td>
<td>Standard addition in serum or plasma</td>
<td>&lt; 15%</td>
<td>50 - 3200</td>
<td>(26)</td>
</tr>
<tr>
<td>2015</td>
<td>N</td>
<td>Top-down</td>
<td>Not available</td>
<td>Not available</td>
<td>Not available</td>
<td>(31)</td>
</tr>
</tbody>
</table>

**MSIA**: mass spectrometric immunoassay; *Average value calculated from 4 laboratories.
reported IGF-I mRNA with mutations in exon 4, 5 or 6 encoded mature wild-type IGF-I, as only the untranslated regions were affected. The A70T and V44M IGF-I variants (and the genetically unconfirmed A67T variants) in circulation were chemically different to wild-type IGF-I in terms of amino acid sequences, but were recognized by immunological method as wild-type IGF-I in the original reports. Immunoassay relies on the recognition of epitope by antigen binding sites on the antibody for quantitation. Unless the variant region or the chemically modified residue were recognized by the antibody or interfere with the antibody binding, they will be detected as wild-type proteins, and may contribute to clinically discordant results.

LC-MS-based platforms provide an ideal methodology for the simultaneous detection and quantitation of protein variants and chemically altered proteins, since these protein species are detected unambiguously as distinct peaks with specific m/z values on the mass spectrum. However, it should be noted that LC-MS-based platforms are also limited by their inability to differentiate isobaric compounds. As in the case of A67T and A70T IGF-I variants, protein variants with different amino acid sequences but with the same empirical formula will have the same observed m/z value on the mass spectrum, and their detection relies solely on their difference in chromatographic behaviours.

The detection of protein variants and chemically altered proteins is also limited by the choice of proteolytic peptides in the bottom-up approach, especially for targeted triple quadrupole instrument, as the quantitative peptides may not cover the variable amino acid sequence. In such cases the protein variants will be detected as wild-type proteins, similar to the immunoassay. Depending on the measurand of interest, the non-specific detection of protein variants or chemically altered proteins may be advantageous when it is the total protein species, instead of the specific form, that is of clinical interest. On the other hand, if the quantitative peptides cover the variant region, the measured native peptides will reflect the level of wild-type proteins in the circulation, but the information carried by the variant or chemically modified peptides may be lost in SRM mode unless its presence was anticipated and monitored for simultaneously with the native peptide. Care must be taken when developing LC-MS-based protein assays for clinical applications, as well as when interpreting the results for clinical decision.

**INSIGHT FOR OTHER PROTEIN QUANTITATION WITH LC-MS PLATFORM**

*Comparison of the two approaches and MSIA for different proteins*

Other complex circulating proteins are expected to be technically challenging for LC-MS-based quantitation, for example proteins with diverse post-translational modification (PTM) such as transferrin and haptoglobin; proteins composed of different subunits such as human chorionic gonadotrophin (hCG) and insulin; and proteins with a combination of the above such as immunoglobulins and thyroglobulin. These proteins may exist as multiple glycoforms or isoforms with different molecular weights, similar to that described above for the IGF-I variants. Multimeric proteins are prone to dissociation during sample handling and LC separation, and the analytical approach should consider all subunits of interest. As demonstrated by the IGF-I assays, different information can be obtained by the selective use of different analytical approaches and strategies.

It is obvious that proteins with diverse PTM cannot be easily quantitated by the top-down approach due to the variable molecular weights. In this regard the use of bottom-up approach may be advantageous since a consensus peptide...
can be used, similar to the case of protein variants. An excellent example is the reference method for glycated hemoglobin HbA1c, which utilized the endoproteinase to generate a terminal hexapeptide for LC-MS/MS quantitation (6). Another recent application was the quantitation of thyroglobulin, which is a 660 kDa dimeric glycoprotein with multiple glycoforms. LC-MS/MS quantitation of total thyroglobulin has been achieved by using a combination of the bottom-up approach, MSIA, and enrichment of peptides without N-glycosylation sites (37). Immunoaffinity capture involves the use of antibody binding to concentrate and purify the protein of interest before either top-down or bottom-up analysis. While it improved the sensitivity and specificity of the LC-MS assay, it also introduces variation in the analytical process. The immunoaffinity process is subjected to temperature, pH and ionic strength of the reaction buffer, stability of the antibody, and conjugation of the antibody to the solid phase support. It is well-known that circulating proteins in serum covers a large dynamic range, and immunoaffinity purification becomes inevitable when the protein of interest is present at picomolar to femtomolar concentration levels, as in the case of thyroglobulin or growth factors (38). At this concentration, the use of the top-down or bottom-up approach alone is not feasible due to the matrix suppression effect from highly abundant proteins. There has been rigorous discussion on MSIA, or hybrid ligand-binding assay (LBA)/LC-MS technologies, for biopharmaceutical compounds for regulatory purpose (39,40). Commercial immunoaffinity proteomics assays were also available for clinical application (41). It will be expected that MSIA will play a crucial role in the analysis of low abundant proteins with clinical significance.

**Standardization and harmonization**

Similar to other analytical platforms, the quality of LC-MS-based protein assays cannot be guaranteed if they are not traceable to higher reference materials or methods, or not harmonized among existing platforms. The bottom-up approach reported by Kay et al. in 2013 exhibited positive bias compared with immunoassays (Passing-Bablok regression slope = 1.37), while negative bias was observed in the top-down approach by Bystrom et al. (Deming linear regression slope = 0.81) and Hines et al. (Least square linear fit slope = 0.84). The apparent discrepancy may be attributed to the different regression models used for comparison, different reference materials used for calibration in two assays, or the use of different patient cohort.

A systematic study was carried out by Cox et al. in 2014 which compared the inter-laboratory agreement of IGF-I measurement using the bottom-up LC-MS platforms and immunoassays (42). A standardized sample and calibrator preparation procedure was used across five laboratories, and two designated tryptic peptides were measured by the different LC-MS instruments in each laboratory. Instead of using human serum, the standardized method used rat serum as the matrix for the preparation of the human IGF-I calibration curve, which was not subjected to the variation of endogenous rat IGF-I concentrations. Results showed that considerable discrepancy was observed between laboratories, especially for high concentration samples. The study concluded that the use of in-house prepared calibrators may contribute a significant degree of imprecision across different laboratories, even though the calibrators were traceable to the same reference material.

These studies not only highlighted the lack of harmonization in IGF-I measurement between
immunoassays and between different LC-MS platforms, the observation is also applicable to other LC-MS-based protein assays as well. It is expected that the between-assay discrepancy is higher in the bottom-up approach than in the top-down approach, since it is subjected to more analytical variables and especially the choice of quantitative peptides. In an effort to provide a more standardized MS-based proteomic assay, the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium has established an Assay Portal which allows open access to standard operating procedures, reagents, experiment parameters and validation data on different bottom-up assays, developed by the research community (https://assays.cancer.gov/).

FUTURE PERSPECTIVE AND CONCLUDING REMARKS

In the past decade, there has been a growing trend for the implementation of LC-MS-based technology in clinical setting, including steroid analysis, immunosuppressant therapeutic drug monitoring, and toxicology screening. Due to instrumentation advancement including increased sensitivity and resolution, quantitation of macromolecules like proteins or peptides is also possible using LC-MS.

The analysis of IGF-I provides an excellent example of macromolecule determination by LC-MS platform in the clinical setting, which is proven to be a flexible and valuable analytical tool for protein quantitation, and is able to provide additional information on proteins that was not accessible by traditional assays. Like other analytical platforms, the adaptation of LC-MS technology for protein determination requires a continuing effort in validation and standardization for clinical application.

REFERENCE


