



## New Mass Spectrometry Technologies Contributing towards Comprehensive and High Throughput Omics Analyses of Single Cells

Journal:	<i>Analyst</i>
Manuscript ID	AN-MRV-08-2018-001574.R1
Article Type:	Minireview
Date Submitted by the Author:	08-Nov-2018
Complete List of Authors:	<p>Couvillion, Sneha; Pacific Northwest National Laboratory          Zhu, Ying; Pacific Northwest National Laboratory, Environmental Molecular Sciences Laboratory; Zhejiang University Faculty of Science, Chemistry          Nagy, Gabe; Indiana University, Chemistry; Pacific Northwest National Laboratory, Biological Sciences Division          ADKINS, JOSHUA; PNNL, BSD          Ansong, Charles; PNNL, BSD          Renslow, Ryan; Washington State University, School of Chem. Eng. &amp; Bioengineering          Piehowski, Paul ; Pacific Northwest National Laboratory          Ibrahim, Yehia; Pacific Northwest National Laboratory, Biological Sciences Division          Kelly, Ryan; Pacific Northwest National Laboratory, Environmental Molecular Sciences Laboratory          Metz, Thomas; Pacific Northwest National Laboratory, Biological Sciences Division</p>

1  
2  
3  
4 **New Mass Spectrometry Technologies Contributing towards Comprehensive**  
5 **and High Throughput Omics Analyses of Single Cells**  
6

7 Sneha P. Couvillion<sup>1</sup>, Ying Zhu<sup>1</sup>, Gabe Nagy<sup>1</sup>, Joshua N. Adkins<sup>1</sup>, Charles Ansong<sup>1</sup>, Ryan Renslow<sup>1</sup>,  
8 Paul Piehowski<sup>1</sup>, Yehia Ibrahim<sup>1</sup>, Ryan Kelly<sup>1,2</sup>, Thomas O. Metz<sup>1,\*</sup>  
9

10  
11 <sup>1</sup>Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland,  
12 WA, USA  
13

14  
15 <sup>2</sup>Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT, USA  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27

28 \*Corresponding author:  
29

30 Thomas O. Metz  
31

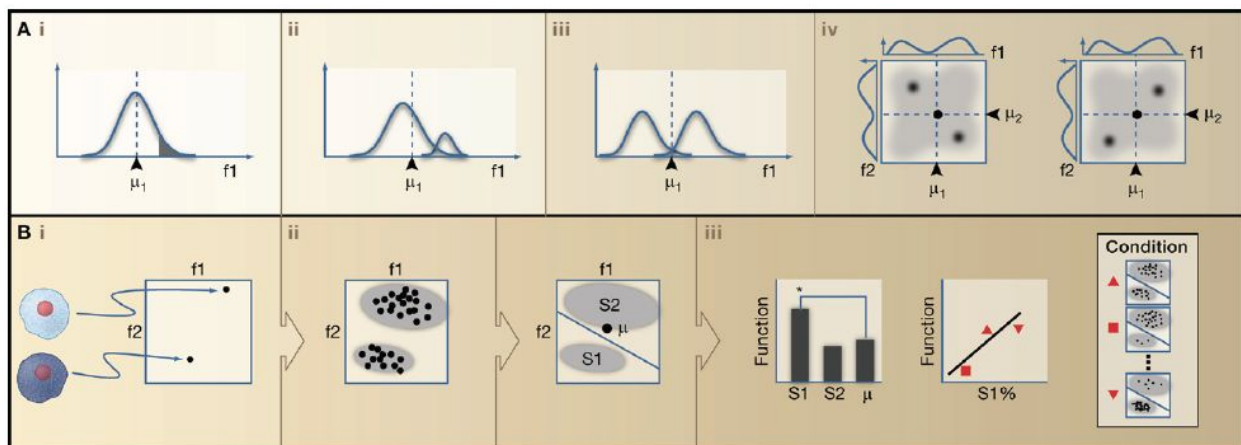
32  
33 Address: 902 Battelle Blvd.  
34 P.O. Box 999, MSIN K8-98  
35 Richland, WA 99352  
36  
37 Phone: 509-371-6581  
38 Email: thomas.metz@pnnl.gov  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**ABSTRACT**

Mass-spectrometry based omics technologies – namely proteomics, metabolomics and lipidomics – have enabled the molecular level systems biology investigation of organisms in unprecedented detail. There has been increasing interest for gaining a thorough, functional understanding of the biological consequences associated with cellular heterogeneity in a wide variety of research areas such as developmental biology, precision medicine, cancer research and microbiome science. Recent advances in mass spectrometry (MS) instrumentation and sample handling strategies are quickly making comprehensive omics analyses of single cells feasible, but key breakthroughs are still required to push through remaining bottlenecks. In this review, we discuss the challenges faced by single cell MS-based omics analyses and highlight recent technological advances that collectively can contribute to comprehensive and high throughput omics analyses in single cells. We provide a vision of the potential of integrating pioneering technologies such as Structures for Lossless Ion Manipulations (SLIM) for improved sensitivity and resolution, novel peptide identification tactics and standards free metabolomics approaches for future applications in single cell analysis.

### Why study single cells?

Routine bulk scale omics measurements which involve the analysis of a large number, often millions, of cells are population-averaging techniques that assume that cell populations are uniform. Cell to cell variability or heterogeneity is in fact very common and is a fundamental property of cellular systems.<sup>1</sup> Different cell types in the same multi-cellular organism exhibit differing phenotypes, even though they are genetically identical, due to distinct regulation of gene expression (**Figure 1**). Cells in the same tissue type may also display differences in phenotypic profiles depending on how they respond to their microenvironment. For example, cells may vary in proximity to blood vessels, other cells or oxygen gradients, which may then trigger a response from only a fraction of the population.<sup>2</sup> The ability to investigate single cells in a high throughput manner can inform us about interesting and important biological phenomena that may be obscured in ensemble measurements. This can be especially valuable in understanding diseases like cancer where heterogeneity can play a pivotal role in disease onset and progression.<sup>3</sup> Studies have shown that the presence of disseminating tumor cells or circulating tumor cells after adjuvant therapy predict a poor prognosis in cancer patients and indicates metastatic progression.<sup>4-6</sup> Although it is possible to detect these rare cell populations, advances in single cell proteomics and metabolomics could provide the much needed insight into the molecular and phenotypic nature of these cells, which is thought to cause marked inter-patient variability with regards to their dormancy or expression.<sup>7</sup>



**Figure 1. Limitations of bulk cell analysis vs benefits of single cell analysis.** (A) Population averages can mask cell heterogeneity. The mean measurement (indicated by dashed lines) of a population may not capture (i) the shaded tail of the distribution, (ii) a subpopulation or (iii) majority of the cells in case of bimodal behavior. (iv) Univariate analysis of a single measurement from individual cells may not be able to distinguish correlated (left) or anticorrelated (right) expression of cells ( $f_1$  and  $f_2$  indicate single cell measurements). (B) Interpreting functional significance from heterogeneity. (i) Individual cells can be represented as points in a feature space (ii) Cells can be partitioned into subpopulations (eg: S1 and S2) in regions of the feature space (iii) The presence of significant differences between subpopulations or ensemble averages can be tested. One can assess how informative is an entire decomposition of heterogeneity (middle and right). Reproduced from Altschuler and Wu<sup>1</sup> with permission from Elsevier, Copyright 2010.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Single cell analysis methods can also help characterize uncultivable individual microorganisms that influence community stability and resilience in microbiomes inhabiting a variety of ecosystems.<sup>8</sup> However, microbial cells are particularly challenging in that they range in size of 1/1000<sup>th</sup> - 1/10,000<sup>th</sup> the size of a eukaryotic cell. Improved methods for analysis of such small samples will allow for important insights, such as in the context of host-pathogen interactions, particularly for intracellular pathogens. Here key differences between uninfected cells, and early-, mid-, and late-infection within a population of cells are typically averaged using conventional analyses such that critical cell-cell signaling and processes of infection are lost. A major hindrance of microbiome proteomics research is the sheer number of different species at different concentrations,<sup>9</sup> which can undermine detection efficiency and control of false-discovery. The microdissection of microbiome and biofilm samples has clear benefits from very small sample processing by greatly improving homogeneity of species studied at one time.

### ***Volume of a single cell & technical challenges***

The dimensions and volume of a single cell can differ by orders of magnitude depending on the organism and cell type. Microorganisms such as *Bacillus subtilis*, *Staphylococcus aureus*, *Escherichia coli*, and *Caulobacter crescentus* typically have individual cell volumes between 0.4-3  $\mu\text{m}^3$ .<sup>3,10</sup> In humans, cell types can vary by several orders of magnitude with erythrocytes, HeLa cells and oocytes having a volume of about 100, 3000 and 4,000,000  $\mu\text{m}^3$  respectively. The size of a single cell requires suitable tools and approaches for sample handling and analysis. Depending on the cells of interest, specialized strategies have to be used for separating and isolating individual cells. Ideally these strategies will introduce minimal artifacts during the process. The limited sample volume in a single cell also calls for highly sensitive measurement methods for accurate detection and quantification of the biomolecules of interest.

The development of whole-genome and whole transcriptome amplification (WGA/WTA) has made single cell genomics and transcriptomics possible despite the miniscule amount of genetic material present in a cell.<sup>11</sup> Advancements in WGA technology such as improved performance and the availability of commercial kits and the routine use of unique molecular identifiers (UMIs) to barcode individual RNA molecules has resulted in robust genome and transcriptome sequence analysis in single cells.<sup>12</sup> In contrast, mass spectrometry-based single cell omics technologies that focus on proteins, metabolites and lipids are less mature, yet are newly emerging fields with great potential. Unlike DNA and RNA, biomolecules such as proteins, metabolites and lipids cannot be amplified, so one has to make do with the quantity actually present in the cell and work on enhancing the intrinsic sensitivity of the method of analysis.

In this paper, rather than providing a comprehensive review of the literature, we highlight key developments in mass spectrometry (MS)-based technologies that can contribute to omics analyses of small samples down to a single cell. Because much of the work on single cell MS-based omics performed to date has focused on proteomics, in part because that field is more mature compared to metabolomics and lipidomics, the majority of our review will in turn focus on recent developments in single cell proteomics analyses. We also focus on approaches for comprehensive, untargeted analyses of small samples, rather than approaches for targeted analyses of more limited sets of biomolecules. Finally, we provide a vision for how the

1  
2  
3 remaining challenges in single cell omics analyses can be overcome using specific technologies  
4 as the drivers.  
5

## 6 7 ***SAMPLE ISOLATION AND MANIPULATION*** 8

9  
10 Only very recently has the prospect of extending MS-based global proteome profiling to single  
11 mammalian somatic cells entered the realm of possibility.<sup>13-16</sup> Inefficiencies and sample losses  
12 associated with all steps of sample handling, separation, ionization and transmission to the  
13 high-vacuum region of the mass analyzer have historically resulted in sample requirements of  
14 micrograms or milligrams of protein isolated from millions of cells to achieve a reasonable  
15 depth of proteome coverage. These sample size limitations precluded the analysis of rare cells  
16 or required averaging over multiple cell and tissue types, greatly limiting the specificity of the  
17 measurements and obscuring important underlying heterogeneity. However, consistent  
18 improvements in each of these areas has greatly increased the feasibility of in-depth proteome  
19 measurements for single cells and other small samples for which only picogram or low-  
20 nanogram amounts of protein starting material are available.<sup>17, 18</sup> Below we highlight advances  
21 in sample handling, prior to MS analysis.  
22  
23

### 24 25 ***Isolating single cells*** 26

27  
28 Single cell techniques rely on technologies to enrich cell populations of interest and  
29 isolate rare cell types for further characterization. Although most single cell omics studies use a  
30 well-characterized cell line such as HeLa for a proof-of-concept demonstration of their method,  
31 it is important to keep in mind that actual biological samples will require robust methods to  
32 isolate cells of interest. It is imperative that the isolation technique is gentle enough to avoid  
33 altering biomolecules in the cell through physical shear forces or by harsh chemicals used  
34 during the process. Traditional methods such as filtration and centrifugation for cell isolation  
35 have made use of differences in cell density or size for a more gross level purification of  
36 samples with low specificity.<sup>19</sup> More selective separation is afforded by affinity-based methods  
37 that use antibodies against cell surface antigens that are differentially expressed in different  
38 cell types.<sup>20</sup> Flow cytometry has been widely used for rapid sorting of discrete populations of  
39 cells and individual cells based on their cell-surface markers.<sup>21</sup> Adaptations of flow cytometry  
40 include fluorescence activated cell sorting (FACS), which is a high-throughput platform for  
41 single cell analysis.<sup>22</sup> Affinity substrates and magnetic beads have been used for isolating and  
42 capturing intact cells from complex samples.<sup>23</sup> The limitation with affinity-based techniques lies  
43 in the fact that they are inherently antibody centered, targeted techniques and the user must  
44 have prior knowledge on the target to be measured. Laser capture microdissection (LCM) has  
45 been used to harvest specific cells of interest from tissues and is well suited for coupling with  
46 techniques that require small numbers of cells.<sup>24</sup> We recently used a LCM approach to gain  
47 spatially resolved insights in a proteomic study of lung development in microdissected alveolar  
48 tissue.<sup>25</sup> Also, the more traditional approach of manual isolation of cells is used by many  
49 research groups. This approach is tedious, but it ensures that rare but morphologically-  
50 identifiable cells can be isolated.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### ***Single vessel methods***

A typical bottom-up proteomics workflow involves sampling, cell lysis, extraction of proteins from sample, proteolytic digestion of proteins to peptides, separation, fractionation or enrichment of peptide mixtures and finally identification of peptides using MS. Each of these steps contributes to adsorption losses of proteins/peptides to solid supports (tubes, pipette tips etc.), which becomes a problem as the sample size approaches the analytical limits of detection. This can result in limited proteome coverage, with only the most abundant proteins being detected, and given higher analytical variability as one operates closer to the limits of detection, it can be challenging to confidently characterize differences across such samples.<sup>26</sup> To eliminate the contamination and losses typically inevitable in multi-step sample preparation workflows used in traditional MS-based omics, many groups have attempted to carry out all the necessary steps in a single vessel, thereby minimizing the number of surfaces with which the samples comes into contact.

In 2005, Wang et al. reported a 'single-tube' protocol that used the organic solvent trifluoroethanol (TFE) for cell lysis and protein denaturation.<sup>27</sup> By evaporating the organic solvent during lyophilization, cleanup steps typically carried out for traditional detergent based protocols and transfer to other tubes were avoided. Using the TFE protocol, 246 peptides and 104 proteins were identified from 5000 MCF-7 human breast cancer cells (~500 ng total protein content). Liang Li's group used a NP-40 surfactant to lyse the cells, followed by acetone precipitation to profile the proteome of 500-5000 MCF-7 cells.<sup>28</sup> The gradient time for liquid chromatography (LC) separation was optimized for the number of cells being analyzed, resulting in the identification of 167 and 619 proteins from 500 and 5000 cells, respectively. Kulak et al. used a filter-aided sample preparation (FASP)-based method but omitted the use of strong detergents that require the use of molecular weight cut-off filters for subsequent clean-up.<sup>29</sup> By choosing compatible reagents, they were able to carry out lysis, reduction and alkylation in a single step for samples as small as 1000 HeLa cells. Hughes et al. developed a platform (SP3) for proteomic sample preparation of sub-microgram amounts of material in a single tube workflow based on carboxylate-coated paramagnetic beads thus circumventing the need to use a filter membrane.<sup>30</sup> They were able to obtain an increase of ~5000 unique peptide identifications using an equivalent number of 1000 HeLa cells. This improvement was due to the ability of the SP3 approach to use 1% SDS for enhanced lysis and solubilization. More recently Li and co-workers described an approach utilizing acoustics-assisted lysis and ultra-trace LC-MS analysis to achieve zeptomole detection sensitivity and yielding ~4000 protein identifications from 100-200 MCF-7 cells.<sup>31</sup> Instead of directly using the cells of interest, the MCF-7 cells were first spiked into whole blood and then isolated using microfluidic magnetophoretic isolation, thus demonstrating the capability to isolate rare target cells from a complex matrix. Chen et al. developed an integrated device based on a two-position 10-port valve for online digestion, enrichment and analysis of 100 DLD-1 cells.<sup>32</sup> High temperature trypsin digestion was carried out and an average of 635 proteins were identified within 2 hours. Huang and colleagues designed a simple but robust online sample handling platform incorporating immobilized enzyme reactor principles to analyze nanogram (~ 440-500 ng) quantities of protein from mouse blastocysts and identified 348 proteins.<sup>33</sup>

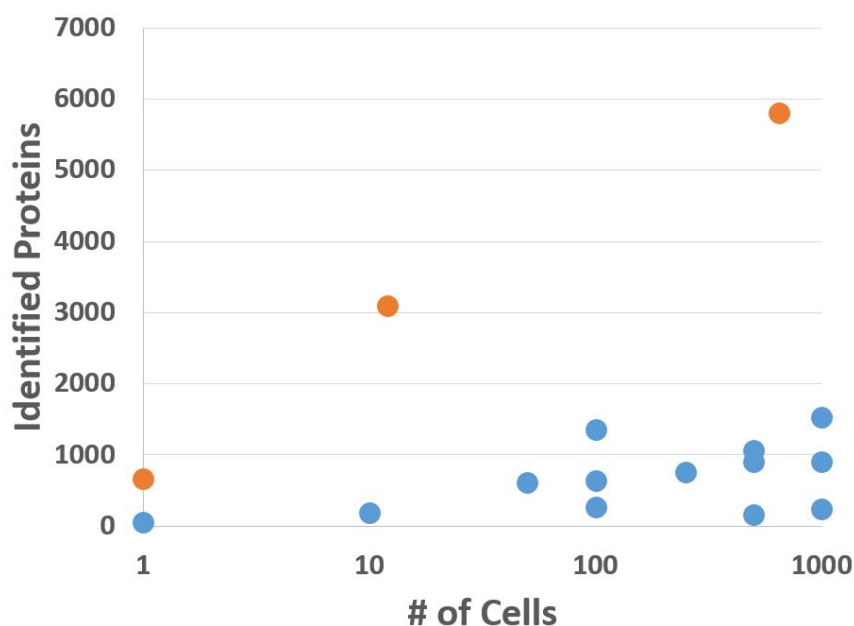
### ***Microfluidic advances***

Innovations in the field of microfluidics have enabled researchers to manipulate small quantities of fluids in ways that were not possible before. The ability to integrate multiple functions with the potential for massive parallelization has made microfluidic approaches especially attractive for single cell omics. In 2010, Mellors and coworkers developed an integrated microfluidic device for analysis of individual human erythrocytes using capillary electrophoresis (CE) coupled with MS via electrospray ionization (ESI).<sup>34</sup> The microfluidic structure incorporated continuous cell delivery, detergent-free lysis (using a buffer compatible with CE-ESI-MS), separation of lysate and ionization for detection by MS. Although this early study was used to detect hemoglobin, a highly abundant protein, it did demonstrate the power of combining microfluidics with MS.

Recently, we have utilized microfluidic sample processing in combination with nanoflow LC and a latest-generation MS system to extend label-free proteomic analyses to small cell numbers<sup>14</sup> and single mammalian cells.<sup>15</sup> A robotic nanopipetting platform with microfabricated well plates reduce the total processing volume to just ~200 nL, providing a significant reduction in potentially adsorptive surfaces during preparation relative to standard preparation protocols.<sup>14, 15</sup> Termed nanoPOTS (Nanodroplet Processing in One pot for Trace Samples), the workflow has enabled deep proteome profiling of trace samples that were previously inaccessible. NanoPOTS enables all of the required processing steps for bottom-up proteomics, including cell lysis, protein extraction, reduction, alkylation and digestion to be achieved within a single nanowell (without sample transfer between steps) through a series of 50 nL pipetting steps and incubations.<sup>14, 35, 36</sup> Evaporation is minimized during dispensing by operating the nanopipette within a humidified chamber, and extended incubations take place with a cover over the nanowell plate to minimize the headspace over the array.

Using nanoPOTS, we have reproducibly achieved a proteome coverage of >3000 protein groups from as few as 10 HeLa cells.<sup>14</sup> To our knowledge, this level of coverage was not achieved in past studies for samples comprising  $\leq 5,000$  mammalian cells.<sup>37</sup> When coupled with fluorescence-activated cell sorting to dispense single HeLa cells into each nanowell, we have profiled nearly 700 proteins on average from each cell.<sup>15</sup> We have also utilized nanoPOTS in conjunction with laser capture microdissection to isolate and analyze regions of thin tissue sections with lateral dimensions as small as 50  $\mu\text{m}$ .<sup>38</sup> The nanoPOTS platform has been applied to a variety of biological tissues, including single dissociated lung cells,<sup>15</sup> sections of single pancreatic islets from type 1 diabetic and nondiabetic donors,<sup>14</sup> rat brain,<sup>38</sup> mouse liver,<sup>39</sup> plant tissues<sup>40</sup> and spiked circulating tumor cells.<sup>41</sup> In addition, we have preliminarily isolated and profiled spiked circulating tumor cells from whole blood.<sup>39</sup> Finally, to increase the proteome coverage, we have utilized the robotic platform to fractionate nanoPOTS-prepared samples from a high-pH nanoLC separation into nanowells for a subsequent low-pH separation.<sup>42</sup> Using nanoPOTS in combination with this nanowell-mediated 2D LC workflow, we have increased proteome coverage to ~6000 protein groups for 650 HeLa cells or 10 pooled pancreatic islet thin sections. A plot of proteome coverage achieved using the nanoPOTS platform compared with other nanoscale proteome processing and analysis platforms is shown in **Figure 2**.





**Figure 2. Proteome coverage achieved using nanoPOTS (orange) and other (blue) technologies for samples comprising  $\leq 1000$  mammalian somatic cells.** Orange data points, from left to right, are from Zhu et al.,<sup>15</sup> Zhu et al.,<sup>14</sup> and Dou et al.<sup>42</sup> Blue data points represent proteome coverage reported by Wang et al.,<sup>28</sup> Zhang et al.,<sup>43</sup> Wiśniewski et al.,<sup>44</sup> Kasuga et al.<sup>45</sup> and Li et al.<sup>46</sup>

While the microfluidic nanoPOTS platform has reduced the adsorptive losses that have hindered proteomic analysis of trace samples, this is not the only approach that has been utilized to extend MS-based proteomics to the analysis of single mammalian cells. For example, Li et al.<sup>46</sup> recently reported on a microfluidic reaction chamber in an oil layer separated by an air gap minimized evaporation of sample and reagents within a droplet containing a final volume of  $\sim 550$  nL. When a single HeLa cell was analyzed, 51 proteins were identified as indicated by the blue data point on the left of **Figure 2**. In addition, Budnik, et al.<sup>16</sup> recently prepared single cells and populations of carrier cells within commercial microtubes and labeled each sample with a tandem mass tag (TMT). The labeled samples were pooled and analyzed by LC/MS, and the carrier cells within the pooled sample provided sufficient MS<sup>1</sup> sensitivity to trigger MS<sup>2</sup> peptide sequencing and identification of peptides from single cell samples based on their respective reporter ions. While the proteome coverage achieved by this method was not reported other than the aggregate number of identified proteins from multiple single cell analyses, it is likely that strategies that combine such multiplexed analyses with reduced losses achieved by microfluidic sample preparation will yield additional insights into the proteomes of single cells and other trace samples.

### **MOLECULAR IONIZATION, SEPARATION, DETECTION, AND IDENTIFICATION**

Because mass spectrometers are mass sensitive detectors and have detection biases towards the most abundant molecules in a sample, the MS and omics communities have placed significant effort into developing methods and technologies that result in maximizing the separations of peptide, metabolite and lipid samples, the efficiency of ion generation,

1  
2  
3 transmission, and utilization inside the MS, and the confident identification of detected  
4 molecules. In the below sections, we highlight key advances in these areas as they might be or  
5 have been developed for or utilized in the analysis of single cells or small sample sizes.  
6  
7

### 8 ***Advances in molecular ion generation and transmission***

9

10 One of the most commonly used configurations for the analysis of biomolecules is LC  
11 coupled to MS via ESI. Two decades ago, ESI-MS instrumentation involved a high-flow ESI  
12 source and highly inefficient ion optics such that very few analyte molecules from solution were  
13 ionized and detected at the mass analyzer, resulting in poor analytical sensitivity. Ionization  
14 efficiency, the extent of conversion of solution-phase molecules into gas-phase ions, depends  
15 on the size of the initial charged droplets emanating from the electrospray source, with smaller  
16 droplets being more readily desolvated and producing more ions.<sup>47, 48</sup> By reducing the flow rate  
17 of sample delivered to the electrospray source to a few tens of nanoliters, smaller and more  
18 readily desolvated initial charged droplets are produced, which greatly improves ionization  
19 efficiency.<sup>48</sup> Such nanoESI sources and their coupling with flow-compatible LC or CE separations  
20 have been made possible through substantial advances in ESI emitter technology and  
21 miniaturization of the accompanying separations. For example, creating an emitter with the  
22 appropriate narrow orifice and high aspect ratio to support an electrospray at low nanoliters  
23 per minute<sup>48, 49</sup> or even picoliters per minute<sup>50</sup> required the development of heating and pulling  
24 techniques or chemically etching of fused silica capillaries.<sup>49-51</sup> Similarly, new methods needed  
25 to be developed to prepare LC columns that delivered analyte at the appropriate low flow  
26 rates, which has been achieved through either packing of media into narrow-bore capillaries<sup>52-</sup>  
27 <sup>54</sup> or preparing porous layer open tubular LC columns<sup>31, 55, 56</sup>. With these advances and similar  
28 developments for capillary electrophoresis (CE)-MS, nanoscale samples could be effectively  
29 separated and ionized.<sup>57, 58</sup>  
30  
31  
32  
33  
34

35 Similar inefficiencies in MS ion optics needed to be overcome to effectively transmit  
36 ions from the electrospray ionization source to the mass analyzer that requires high vacuum to  
37 operate effectively.<sup>59</sup> A typical ESI-MS instrument interfaced an electrospray emitter with a very  
38 narrow orifice or heated capillary inlet such that only a small fraction of the electrospray-  
39 generated plume of ions arrived at the first pumping stage of the instrument, which is typically  
40 operated at a few Torr. From there, a skimmer orifice again sampled only a tiny portion of the  
41 ions for transmission to the high-vacuum region where conventional ion optics such as multi-  
42 pole ion guides could efficiently operate. To address the skimmer-related losses, our group  
43 developed the electrodynamic ion funnel,<sup>59-61</sup> which could operate in the rough-pumped  
44 vacuum regime and captured the entire cloud of ions entering in the first vacuum stage to  
45 transmit them to high vacuum. Related advances have also addressed the losses at the  
46 atmospheric pressure interface including multicapillary inlets,<sup>62</sup> high-conductance single  
47 inlets,<sup>61</sup> and low-pressure electrospray sources.<sup>63, 64</sup> Today, many of these advances have been  
48 commercialized and are standard issue on biological mass spectrometers.  
49  
50  
51  
52

### 53 ***SLIM: a new frontier for separations and ion manipulations in the gas-phase***

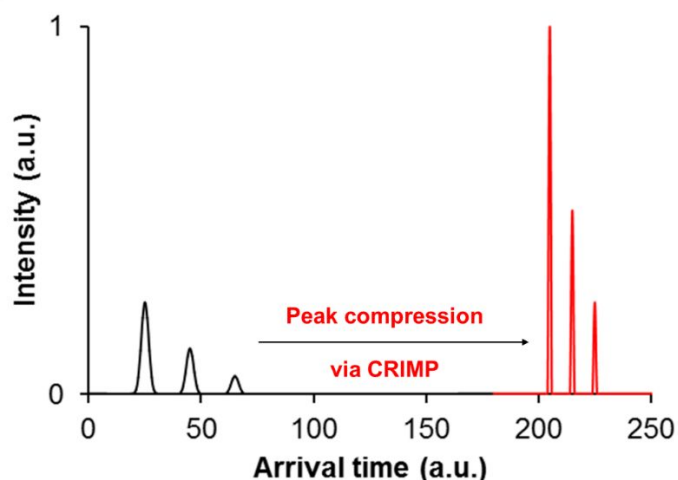
54  
55  
56  
57  
58  
59  
60

1  
2  
3 An appealing technique for enhancing single cell analysis methods is the incorporation  
4 of ion mobility (IM) spectrometry.<sup>65</sup> IM-MS is a rapid gas-phase technique, whereby ions are  
5 separated on the basis of their mass-to-charge ( $m/z$ ) and their shapes (mobilities) in a given  
6 buffer gas.<sup>66</sup> This technique provides a viable, and attractive, alternative to solution-phase  
7 methods since separations are performed on the order of milliseconds in IM, while on the order  
8 of minutes, to potentially hours, in conventional chromatographic methods.<sup>66</sup> Numerous IM-MS  
9 separations approaches have been documented for a wide variety of global, as well as targeted  
10 applications.<sup>67-70</sup> Despite their success in many applications, all of these prior approaches fall  
11 under the same joint bottleneck: lack of sensitivity and poor resolution. Structures for lossless  
12 ion manipulations (SLIM) is a recent traveling wave (TW)-based technology that has successfully  
13 been applied to overcome this resolution barrier by greatly extending the separation path  
14 length as compared to conventional IM-MS platforms.<sup>71-76</sup> In TWIM-MS-based separations, ions  
15 will separate based on their ability to keep up with the motion of the traveling wave, as related  
16 to their intrinsic molecular properties, such as  $m/z$  as well as their shapes (i.e. mobilities).  
17 Higher mobility ions (faster/more compact) will move more similarly to the speed of the TW,  
18 while lower mobility ions (slower/more elongated) will get passed by the TW and thus fall  
19 behind the wave. While much has been discussed in recent literature on the tremendous  
20 increases in resolution and resolving power that are enabled by SLIM, here we will discuss its  
21 potential in addressing the sensitivity associated with sample-limited applications. In these  
22 types of applications, every ion counts and efforts to eliminate ion losses are core to SLIM  
23 performance.  
24

25  
26 This raises the question of: how can more analyte ions be introduced to the mass  
27 analyzer. Paradoxically, conventional/commercial IM platforms cannot accommodate a large  
28 influx of ions, as it overwhelms the instrument's ability to utilize and quantify them. Thus, the  
29 challenge becomes: how can we work with large ion populations? Further, how can we  
30 integrate all the required technologies not only into a single instrument platform, but  
31 ultimately in a single, standalone, experiment? To begin to address these challenges, we  
32 recently reported on the ability to accumulate ions in the SLIM device itself, termed 'in-SLIM ion  
33 accumulation' by halting the traveling wave in the second region and thus permitting ions to  
34 accumulate at the interface between the first and second regions<sup>71</sup>. This technological  
35 advancement has enabled the introduction of over a billion ions, representing a 2-3 orders of  
36 magnitude increase in charge capacity as compared to ion introduction via other standard  
37 approaches (such as the ion funnel trap).  
38

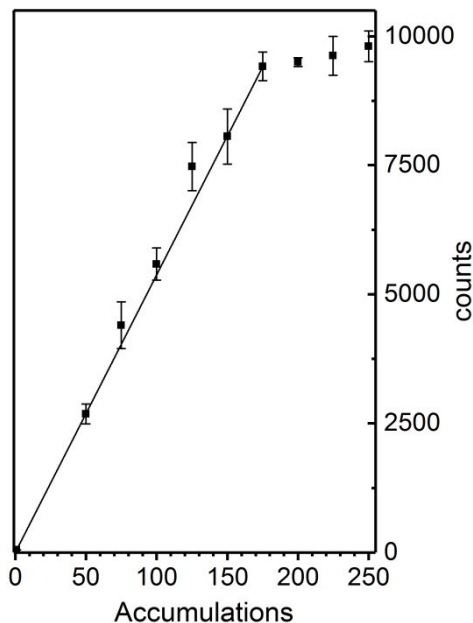
39  
40 With this new technology for ion introduction comes the issue of how to work with such  
41 large populations. Specifically, when ions are accumulated for several seconds in the SLIM  
42 module, the total width of the initial ion packet will be significantly larger, and much broader,  
43 than when ions are introduced for a few milliseconds by e.g. an ion funnel trap. Another recent  
44 capability for SLIM is the ability to spatially compress ion mobility peaks through compression  
45 ratio ion mobility programming (CRIMP).<sup>71, 75</sup> CRIMP successfully permits a broad, diffuse, ion  
46 packet (containing multiple unique mobility peaks/features) to be compressed into ones with  
47 much higher signal intensity and signal-to-noise (S/N), thereby increasing the sensitivity of  
48 measurements. **Figure 3** demonstrates an example of how several broad initial peaks can be  
49 compressed into much narrower ones. This combination of introducing massive ion populations  
50 through 'in-SLIM ion accumulation', followed by a subsequent CRIMP step, enables us to  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 manipulate small sample sizes and greatly increase the sensitivity of measurements. Presently,  
4 we have demonstrated limits of detection (LOD) of 50 pM for phosphoproteomics  
5 measurements<sup>77</sup> and as low as 1 pM for the analyses of other peptide mixtures.<sup>71</sup> Future  
6 improvements to the in-SLIM ion accumulation process (e.g., variation of traveling wave  
7 parameters during accumulation) will enable us to further improve upon our measurement  
8 sensitivity and overcome the limitations associated with volume-limited samples.  
9  
10



29 **Figure 3. Depiction of the CRIMP step, where three broad ions are successfully compressed into much**  
30 **narrower ones following CRIMP, thereby increasing signal intensity, and thus S/N.**  
31

32  
33 Even with the combination of in-SLIM ion accumulation and a subsequent CRIMP step,  
34 other steps may be necessary to further increase the sensitivity of small sample analyses. Akin  
35 to fraction collection and enrichment protocols that are prevalent in conventional  
36 chromatography-based applications, our SLIM platform enables similar manipulations to be  
37 performed in the gas phase. One such potential manipulation is that of 'ion enrichment',<sup>78</sup>  
38 where a specific region of the SLIM module can be used as an ion trapping region to accumulate  
39 low abundance ions, and thus build up their population while simultaneously resolving them  
40 from unwanted background interferences (**Figure 4**). This capability permits mobility-selected  
41 trapping of desired ions on the order of seconds, in a standalone experiment, as opposed to the  
42 tedious, time consuming, and decoupled process of chromatographic fractionation followed by  
43 an enrichment step prior to detection.  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



**Figure 4. Demonstration of ion enrichment in SLIM.** As an ion is enriched longer (more accumulations), its signal intensity will continue to linearly increase until reaching a horizontal asymptote. Reproduced from Chen et al<sup>79</sup> with permission from American Chemical Society, Copyright 2016.

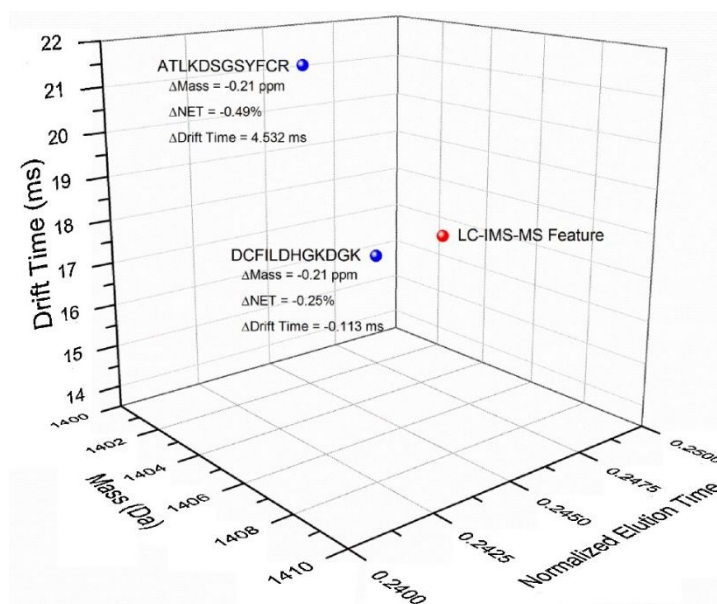
SLIM has a very promising future as part of the analytical toolbox for sample-limited applications through its ability to: accumulate, and work with, large ion populations<sup>71</sup>, increase the sensitivity of measurements through peak compression via CRIMP<sup>75</sup>, as well as enriching ions through various mobility-selective trapping regions<sup>78</sup>, all in a standalone module. These recent advancements in SLIM present an exciting time for this technology to rise to the forefront of analytical measurements in the search for increased sensitivity without sacrificing resolution.

### ***Peptide identification strategies for single cell proteomics***

A particular challenge that arises due to the sample mass limitations of single cells for omics analyses, is the need for MS/MS analysis for confident identification of the analyte of interest. In the case of peptides and proteins, gas-phase fragmentation provides necessary sequence information that is used along with the parent ion mass and the organism's genome to obtain an accurate identification. In a data-dependent acquisition (DDA) experiment, the most commonly applied approach in bottom up proteomics analysis, all ions are injected into the mass spectrometer and a full scan mass spectrum (MS1) is generated. A user-defined number of the most abundant ions from the MS1 scan are then sequentially isolated for fragmentation. As a result, the ion flux required to generate high quality tandem mass spectral (MS/MS) information exceeds that required for MS1 level detection by several fold. Thus, a number of approaches have been developed to achieve confident identification of peptides in high throughput, bottom up proteomics experiments without the need to perform fragmentation.<sup>80-82</sup> Briefly, equivalent samples with adequate sample mass are analyzed using identical LC-MS/MS conditions to identify peptide sequences and their LC elution times in order

1  
2  
3 to establish a comprehensive reference library. To further increase the experimental dynamic  
4 range and coverage of the proteome, these samples may also be fractionated prior to LC-  
5 MS/MS analysis. The measured LC retention time and exact mass information is then leveraged  
6 to identify matching MS1 signals in subsequent analyses, based on peptide mass and elution  
7 time. In the case of sample limited proteomics, the existing MaxQuant MBR algorithm has  
8 shown significant utility (any or all nanoPOTS and SNaPP references here.<sup>14, 15, 83-85</sup>  
9

10 While promising, approaches based on MS1 and elution time information do have  
11 limitations at present. First, due to the high sample complexity of bottom up proteomics  
12 measurements, accurately assessing and controlling the false discovery rate of identifications  
13 based on MS1 and elution time information is challenging.<sup>86</sup> One way that we foresee  
14 improving confidence is to add a third dimension of information for matching,<sup>87</sup> as illustrated in  
15 **Figure 5**. The high sensitivity and reproducibility of ion mobility separations,<sup>88</sup> as discussed  
16 above, make them the ideal candidate for a third measurement dimension, and the  
17 improvements obtainable are just beginning to be realized.  
18  
19  
20  
21



22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41 **Figure 5. Identification of a peptide feature based on comparison of experimental properties vs those**  
42 **contained in a reference library.** A peptide was detected in a LC-IMS-MS experiment with a measured  
43 mass and elution time (red dot) that could match to two different peptides in a reference library (blue  
44 dots), when considering only mass and elution time. The mass measurement errors between the  
45 experimental mass and library masses are identical, while the normalized elution time errors are  
46 different but not sufficiently so to enable a unique identification. Drift time, as measured by drift tube  
47 IM, was significantly different between the two library entries to enable the peptide to be identified as  
48 DCFILDHKGKDGK. Reproduced from Crowell et al.<sup>89</sup> with permission from Elsevier, Copyright 2013.  
49

50  
51 The second limitation to identifying peptides in single cell analyses using MS1 and elution time  
52 information and reference libraries, is that one can only identify peptide species that have  
53 already been identified in other analyses. The disadvantage here is two-fold. First, the number  
54 of analyses required to produce high proteome coverage for an appropriate reference library is  
55 increased, placing the burden on platform throughput. Secondly, analyses are limited to what is  
56  
57  
58  
59  
60

1  
2  
3 already known and discovery of novel peptides is not possible. To drive innovation in this area,  
4 application of machine learning technologies will play a key role. One potential direction is  
5 through the development of models that accurately predict LC elution and IMS drift times.  
6 Models, trained on ample existing data, could be used to create matching databases without  
7 the need for additional analyses.<sup>90-92</sup> Further, these models could then be extended to predict  
8 the behavior of hypothetical peptides or post-translationally modified species not previously  
9 identified, expanding the utility of these approaches as 'discovery' tools.  
10  
11

### 12 ***Standards-free metabolomics***

13  
14  
15 As described above, MS-based omics analyses of single cells will be challenged by the  
16 amount of molecules that can be isolated from the sample, and efficiently ionized and  
17 transferred through the instrument to the detector in such a way that allows for their  
18 comprehensive, unambiguous identification and accurate quantification. In this section, we  
19 focus on unambiguous identification or characterization of the chemical structures associated  
20 with detected metabolites and other small molecules. Because peptides and lipids are  
21 polymeric or otherwise comprised of well-characterized building blocks that allow for fairly  
22 accurate prediction of their tandem mass spectra and therefore algorithms and associated  
23 software that enable relatively comprehensive, confident characterizations of the proteome  
24 and lipidome,<sup>93-96</sup> we will discuss efforts in the research community that are driving the  
25 implementation of standards-free metabolomics analyses and how these might be leveraged  
26 for maximizing the amount of information that can be obtained in analysis of single cells.  
27  
28

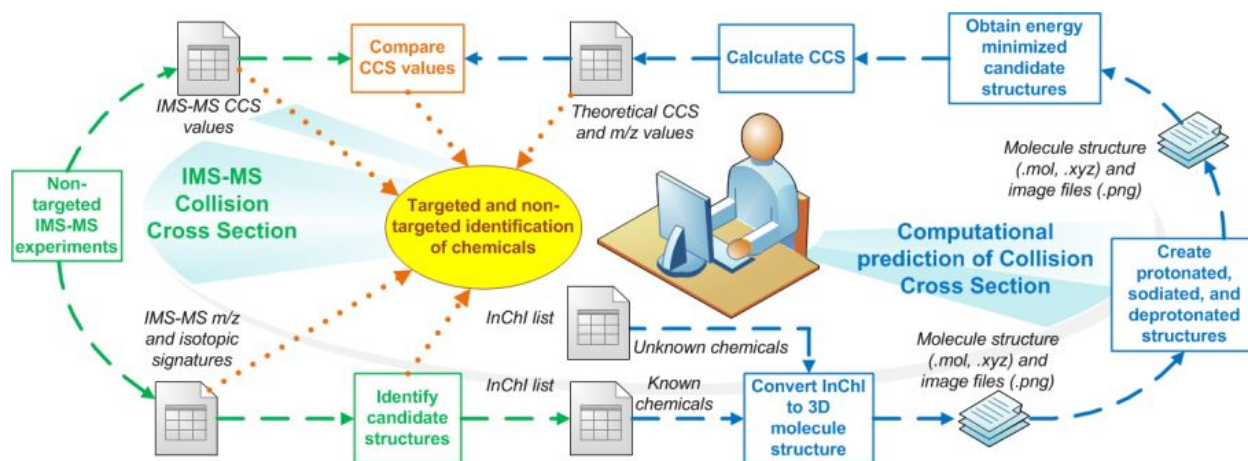
29  
30 Metabolites and related chemicals, such as those with anthropogenic origins, are not  
31 constrained to polymeric or template-like structures. Of course, there are polymers that can  
32 qualify as metabolites or be considered as part of the metabolome, such as carbohydrates (i.e.  
33 the glycome<sup>97, 98</sup>), as well as secondary metabolites that do have template-like structures, such  
34 as terpenes<sup>99</sup> and terpenoid;<sup>100</sup> however, the structures of chemicals that largely comprise the  
35 metabolome are constrained only by the laws of thermodynamics (i.e. those that determine a  
36 chemically stable structure). Because of this chemical diversity, methods have not yet been  
37 developed that allow for comprehensive and unambiguous identification of all metabolites  
38 detected in MS-based metabolomics studies. This technical gap is due in part to the paradigm  
39 typically required for a confident, chemical identification of a molecule,<sup>101-104</sup> which is based on  
40 comparison of multiple, orthogonal experimental data (e.g. MS/MS spectra and retention  
41 times) to the same data contained in a reference library that was constructed through analysis  
42 of authentic reference compounds. This represents the single biggest limitation in MS-based  
43 metabolomics studies, because authentic chemical standards are not available for the majority  
44 of metabolites or chemicals.<sup>105</sup> When microliter volume samples can yield hundreds of  
45 thousands of features,<sup>106</sup> which may represent billions of currently unknown molecules, a  
46 paradigm shift is required in order to comprehensively and unambiguously identify all the  
47 molecules in complex samples. Therefore, there is growing interest within the metabolomics  
48 community, and among funding agencies,<sup>107</sup> to develop methods and tools that accurately  
49 predict experimental properties of metabolites encountered in MS-based metabolomics  
50 analyses, such as retention times,<sup>108-112</sup> MS/MS spectra,<sup>113-117</sup> collision cross sections (CCS),<sup>118-</sup>  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 <sup>120</sup> etc. Ultimate success from these efforts will be an analytical approach that provides  
4 unambiguous identification of the complete set of detected molecules in a sample.

5  
6 Because any experimental manipulation of samples, molecules, and ions during MS-  
7 based metabolomics analyses will be associated with losses, it is important to minimize such  
8 manipulations in analyses of small samples, such as from single cells, in order to maximize the  
9 overall analysis efficiency. As such (and as discussed above), approaches that rely on MS1-level  
10 information for identification and quantification are preferred, and are characterized by  
11 retention time, CCS, molecular ion  $m/z$ , and isotopic distribution. Of these,  $m/z$ , isotopic  
12 signature, and CCS are those that are most accurately calculated and measured. Calculation of  
13  $m/z$  and isotopic signatures based on molecular formula is a trivial calculation when the  
14 chemical structure is known or proposed. To take advantage of this information, modern MS  
15 instrumentation is capable of mass measurement accuracies of a few ppm to sub-ppm,<sup>121, 122</sup>  
16 depending on type of mass analyzer.<sup>123</sup> However, until mass analyzer precision can differentiate  
17 isomers based on differences in the mass equivalence of their chemical bond binding energies  
18 (sub-ppb level),  $m/z$  measurements will, at best, only ever be able to provide a molecular  
19 formula for each detected feature. For a given molecular formula, hundreds of known  
20 molecules can be represented (e.g.,  $C_8H_{16}O_2$ ,  $C_{10}H_{12}O_2$ , and  $C_8H_8O_3$  all represent over 100  
21 known molecules in PubChem,<sup>124</sup> and theoretically, each formula can represent billions of  
22 isomers (i.e., well over  $10^{60}$  possible molecules are possible).<sup>125</sup> The future of small molecule  
23 identification must rely upon consistently measurable and accurately calculable attributes that  
24 are (semi)-orthogonal to  $m/z$ . As the number of attributes increases, the effective distance  
25 between features increases, enabling unambiguous identification of molecules in small-volume  
26 complex biological samples.

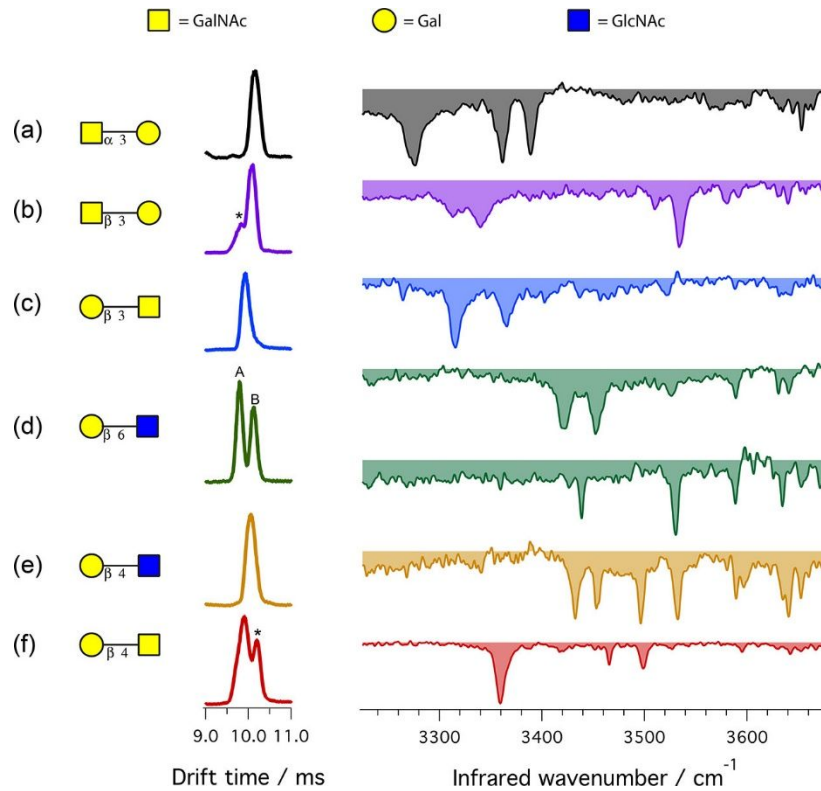
27  
28 As an example semi-orthogonal attribute to  $m/z$ , CCS can be experimentally measured  
29 with high reproducibility and accuracy (condition dependent),<sup>88</sup> since IM separations do not  
30 involve interactions with surfaces, and therefore do not suffer from degradation of  
31 chromatographic stationary phases as do GC and LC. CCSs can also be very accurately predicted  
32 based on three-dimensional chemical structures using quantum chemistry methods,<sup>126-129</sup> and  
33 with lesser accuracy based on two-dimensional structures or chemical properties (e.g. pKa)  
34 using machine learning approaches.<sup>130, 131</sup> Accurate prediction of multiple molecular attributes  
35 such as  $m/z$  and CCS will enable more comprehensive identifications of the metabolome  
36 through matching to experimental data (**Figure 6**). Forthcoming applications of IMS separation  
37 methods should include consideration of multiple adduct ion types simultaneously, CCS peak  
38 shape (as governed by the molecular conformer distribution, see **Figure 7**), and CCS of adduct  
39 ion multimers (e.g., molecular ion hetero- and homo-dimers) to further increase the  
40 dimensionality of the IMS data, and thus increase the uniqueness of each molecular  
41 identification.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60





**Figure 6. Representative workflow for identifying molecules based on matching of experimental  $m/z$  and CCS to computationally predicted values.**

An intriguing option for the identification of molecules from single cells is the incorporation of infrared spectroscopy (IR), or other spectrophotometry techniques, into MS-based measurement approaches. Rizzo and colleagues have demonstrated the coupling of IR with ion traps,<sup>132</sup> and more recently, with IMS in order to generate both CCS and IR spectra for use in characterizing biomolecules (**Figure 7**).<sup>133</sup> Infrared spectra are almost completely orthogonal to  $m/z$ , with each molecule having a spectra that is unique (with exception to enantiomers).<sup>134</sup> Critically, IR spectra are also calculable using similar computational pipelines already being developed for quantum chemistry-based CCS calculations.<sup>135</sup> IR combined with ion trap mass spectrometry has been successfully used for small molecule identification by Martens et al.,<sup>136, 137</sup> with examples of acetylhexosamines, glutaric acid, and ethylmalonic acid. By combining a method like IMS, which can separate stable conformers of molecular ions, and cryogenic cooling systems, which reduce the temperature of the molecular ions, uncongested (narrow absorption band and non-convoluted) infrared spectra can be reproducibly obtained for each  $m/z$ -CCS feature pair. This has successfully been utilized by Masson et al.,<sup>132</sup> demonstrating separation and identification of *cis/trans*-proline versions of the peptide GPGG, with confirmation via density functional theory (quantum chemical electronic structure) calculations. By relying on multiple orthogonal properties such as  $m/z$ , CCS, and IR spectra, future metabolomics analyses of single cells will likely be completely free from reliance upon authentic reference materials.

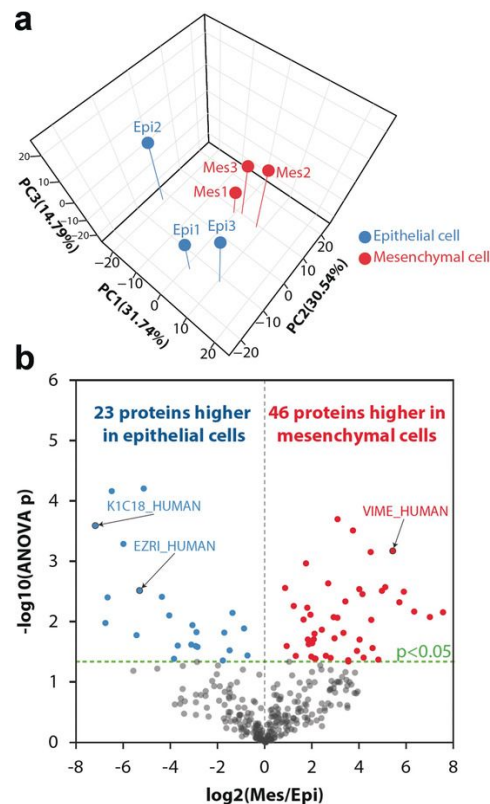


**Figure 7. Drift times and infrared spectra for six isomeric disaccharides.** The drift time distributions for the six isobaric disaccharides overlap, and so they would not be distinguishable by IMS alone. However, their vibrational spectra are very different, providing a means for their identification. Reproduced from Masellis et al.<sup>138</sup> with permission from Springer Nature, Copyright 2017.

### **SINGLE CELL PROTEOMICS APPLICATION: CELL TYPING**

A significant capability offered by single cell transcriptomics is the ability to delineate different cell types present in a tissue sample based on transcript abundance alone. However since transcript levels do not always correlate with protein levels, the utility of the markers identified for specific cell types for enriching or isolating these cells by magnetic-activated cell sorting (MACS) or FACS is usually an unknown. Protein markers for specific cell types by definition should readily allow isolation of specific cell populations of interest more efficiently. We have recently evaluated the feasibility of our nanoPOTS platform to differentiate human cell types from a clinical specimen based on proteome expression.<sup>15</sup> Here primary human lung epithelial and mesenchymal cells were isolated and the cryopreserved lung epithelial and mesenchymal cells were FACS-sorted into nanowells at a count of one cell per well (three single-cell samples were analyzed). 485 proteins were identified across the single-cell samples, and of the 485 proteins, 328 proteins were quantifiable. The levels of all 328 proteins quantified in single cells were projected onto their principal components. Principal component analysis (PCA) revealed the single-cell proteomes separated by cell type along component 2 (**Figure 8a**), indicating that our label-free single-cell proteomics platform has the potential to identify cell types based on protein expression alone. To identify features facilitating the

distinction of the two cell types, analyses of variance (ANOVA) were performed. ANOVA comparing single-cell epithelial and mesenchymal proteomes revealed 20% of quantified proteins to be significantly differential ( $p < 0.05$ ) (**Figure 8b**). Among the significantly differential proteins were those expected to be cell type specific, such as vimentin (VIME\_human) a mesenchymal marker, which was higher in abundance in mesenchymal cells, as well as ezrin (EZRI\_human) and keratin 18 (K1C18\_human), epithelial cell markers, which were higher in abundance in epithelial cells.



**Figure 8. Principal component analysis (PCA) of label-free proteomics data from FACS-sorted cells.** a) Unsupervised PCA based on label-free quantification of proteins expressed in epithelial and mesenchymal cells from human lung. b) Volcano plot of differentially expressed proteins. Epithelial cell Replicate 2 was excluded for this analysis. Reproduced from Zhu et al.<sup>139</sup> with permission from John Wiley and Sons, Copyright 2018.

## PROSPECT FOR SINGLE-CELL OMICS ANALYSES

### Proteomics

Proteins are important because they dictate cell function and therefore changes in the proteome can be a reliable indicator of functional heterogeneity across single cells. Even if the genome of an organism does not typically vary between cells, protein expression can differ amongst cell types and physiological or environmental changes can alter expression patterns

1  
2  
3 even between cells of the same type. Many attempts have been made to achieve deep  
4 proteome coverage using increasingly smaller numbers of cells with the goal of making single  
5 cell analysis a possibility. While proteome analysis has been preliminarily extended to the single  
6 cell level, the coming years will see significant additional gains in terms of proteome coverage,  
7 and, critically, measurement throughput. Incremental coverage gains are expected to result  
8 from further miniaturization of LC separations and further improvements in ion transmission  
9 efficiency and MS sequencing speed. Additionally, multiplexed sample processing and analysis  
10 utilizing, e.g., tandem mass tags such as in the recently reported SCOPE-MS<sup>16</sup> workflow will  
11 potentially provide further gains in both throughput and proteome coverage for single cells,  
12 particularly when implemented within small-volume sample processing workflows. Novel  
13 sample acquisition and data analysis workflows such as the recently reported BoxCar MS may  
14 additionally enhance single cell analyses.<sup>140</sup> Finally, increasing the separation power of ion  
15 mobility-based gas-phase separations through extended pathlengths, as with the SLIM  
16 platform,<sup>141, 142</sup> may ultimately supplant liquid-phase separations altogether, resulting in an  
17 orders-of-magnitude increase in analysis speed. Given the numerous ongoing developments in  
18 this area, we expect to see single cell proteomics rapidly extended from basic proof-of-concept  
19 studies to become an important and widely used tool in biological and medical research.  
20  
21  
22  
23  
24

### 25 ***Metabolomics and Lipidomics***

26  
27 Single cell metabolomics and lipidomics is still a budding field when compared to proteomics.  
28 Zhang and coworkers recently described a combined droplet extraction and a pulsed direct  
29 current ESI-MS method to record over 600 tandem mass spectra from a single human astrocyte  
30 cell and identified over 300 phospholipids.<sup>143</sup> Yin et al. extracted picoliter volumes of material  
31 from live *Allium cepa* cells using electroosmotic extraction and a hydrophobic MS compatible  
32 electrolyte followed by nano-ESI analysis to identify more than 50 metabolites.<sup>144</sup> Efficient  
33 partitioning of hydrophilic and hydrophobic analytes was observed into the aqueous and  
34 hydrophobic electrolyte phase respectively providing additional insight on the molecular  
35 properties. The field of single cell metabolomics and lipidomics can greatly benefit by  
36 leveraging the recent headway made in single-cell proteomics given that many principles of  
37 sample handling, preparation and analysis using MS will be relevant across disciplines. This will  
38 open up the exciting prospect of multi-omic analyses of single cells.  
39  
40  
41  
42  
43

### 44 ***Acknowledgments***

45  
46 The authors would like to acknowledge the following funding for support. The National  
47 Institutes of Health (NIH), National Institute of Allergy and Infectious Diseases grant  
48 U19AI106772 (T.O.M.); National Aeronautics and Space Administration, Exobiology Program  
49 (T.O.M.); NIH National Cancer Institute grant R33 CA225248 (R.T.K.); NIH National Institute of  
50 Biomedical Imaging and Bioengineering grant R21 EB020976 (R.T.K.); NIH National Heart, Lung,  
51 and Blood Institute grant U01 HL122703 (C.A.); National Institute of Diabetes and Digestive and  
52 Kidney Diseases grant UC4DK108101 (C.A.); and the Pacific Northwest National Laboratory  
53 (PNNL) Laboratory Directed Research and Development Program via the Microbiomes in  
54 Transition Initiative. A portion of results presented in this paper was based on data from  
55  
56  
57  
58  
59  
60

1  
2  
3 analyses performed in the Environmental Molecular Sciences Laboratory, a national scientific  
4 user facility sponsored by the U.S. Department of Energy (DOE) Office of Biological and  
5 Environmental Research, and located at PNNL. PNNL is a multi-program national laboratory  
6 operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.  
7  
8

## 9 **References**

- 10
- 11
- 12 1. S. J. Altschuler and L. F. Wu, *Cell*, 2010, **141**, 559-563.
- 13 2. S. Huang, *Development*, 2009, **136**, 3853-3862.
- 14 3. I. Dagono-Jack and A. T. Shaw, *Nature Reviews Clinical Oncology*, 2018, **15**, 81.
- 15 4. S. Braun, F. D. Vogl, B. Naume, W. Janni, M. P. Osborne, R. C. Coombes, G. Schlimok, I. J.  
16 Diel, B. Gerber and G. Gebauer, *New England journal of medicine*, 2005, **353**, 793-802.
- 17 5. S. Braun, C. Kantenich, W. Janni, F. Hepp, J. de Waal, F. Willgeroth, H. Sommer and K.  
18 Pantel, *Journal of clinical oncology*, 2000, **18**, 80-80.
- 19 6. D. F. Hayes, M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, M. C. Miller, J. Matera, W.  
20 J. Allard, G. V. Doyle and L. W. Terstappen, *Clinical Cancer Research*, 2006, **12**, 4218-  
21 4224.
- 22 7. K. Pantel, R. H. Brakenhoff and B. Brandt, *Nature Reviews Cancer*, 2008, **8**, 329.
- 23 8. *Nature Biotechnology*, 2016, **34**, 1077.
- 24 9. P. Y. Lee, S. F. Chin, H. M. Neoh and R. Jamal, *J Biomed Sci*, 2017, **24**, 36.
- 25 10. P. A. Levin and E. R. Angert, *Cold Spring Harbor perspectives in biology*, 2015, a019216.
- 26 11. I. C. Macaulay and T. Voet, *PLoS genetics*, 2014, **10**, e1004126.
- 27 12. L. Deleye, L. Tilleman, A.-S. Vander Plaetsen, S. Cornelis, D. Deforce and F. Van  
28 Nieuwerburgh, *Scientific Reports*, 2017, **7**, 3422.
- 29 13. C. Lombard-Banek, S. Reddy, S. A. Moody and P. Nemes, *Molecular & Cellular*  
30 *Proteomics*, 2016, mcp. M115. 057760.
- 31 14. Y. Zhu, P. D. Piehowski, R. Zhao, J. Chen, Y. Shen, R. J. Moore, A. K. Shukla, V. A. Petyuk,  
32 M. Campbell-Thompson and C. E. Mathews, *Nature communications*, 2018, **9**, 882.
- 33 15. Y. Zhu, G. Clair, W. B. Chrisler, Y. Shen, R. Zhao, A. K. Shukla, R. J. Moore, R. S. Misra, G.  
34 S. Pryhuber, R. D. Smith, C. Ansong and R. T. Kelly, *Angew Chem Int Ed Engl*, 2018, **57**,  
35 12370-12374.
- 36 16. B. Budnik, E. Levy, G. Harmange and N. Slavov, *bioRxiv*, 2018, 102681.
- 37 17. R. D. Smith, Y. Shen and K. Tang, *Accounts of chemical research*, 2004, **37**, 269-278.
- 38 18. L. Yi, P. D. Piehowski, T. Shi, R. D. Smith and W.-J. Qian, *Journal of Chromatography A*,  
39 2017.
- 40 19. M. J. Tomlinson, S. Tomlinson, X. B. Yang and J. Kirkham, *Journal of tissue engineering*,  
41 2013, **4**, 2041731412472690.
- 42 20. Y. Zhang, V. Lyons and D. Pappas, *Electrophoresis*, 2018, **39**, 732-741.
- 43 21. H. M. Davey and D. B. Kell, *Microbiological reviews*, 1996, **60**, 641-696.
- 44 22. M. H. Spitzer and G. P. Nolan, *Cell*, 2016, **165**, 780-791.
- 45 23. A. Sarkar, H. W. Hou, A. E. Mahan, J. Han and G. Alter, *Scientific reports*, 2016, **6**, 23589.
- 46 24. V. Espina, J. D. Wulfkuhle, V. S. Calvert, A. VanMeter, W. Zhou, G. Coukos, D. H. Geho, E.  
47 F. Petricoin III and L. A. Liotta, *Nature protocols*, 2006, **1**, 586.
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

- 1
  - 2
  - 3
  - 4
  - 5
  - 6
  - 7
  - 8
  - 9
  - 10
  - 11
  - 12
  - 13
  - 14
  - 15
  - 16
  - 17
  - 18
  - 19
  - 20
  - 21
  - 22
  - 23
  - 24
  - 25
  - 26
  - 27
  - 28
  - 29
  - 30
  - 31
  - 32
  - 33
  - 34
  - 35
  - 36
  - 37
  - 38
  - 39
  - 40
  - 41
  - 42
  - 43
  - 44
  - 45
  - 46
  - 47
  - 48
  - 49
  - 50
  - 51
  - 52
  - 53
  - 54
  - 55
  - 56
  - 57
  - 58
  - 59
  - 60
25. G. Clair, P. D. Piehowski, T. Nicola, J. A. Kitzmiller, E. L. Huang, E. M. Zink, R. L. Sontag, D. J. Orton, R. J. Moore, J. P. Carson, R. D. Smith, J. A. Whitsett, R. A. Corley, N. Ambalavanan and C. Ansong, *Scientific reports*, 2016, **6**, 39223.
26. H. B. Gutstein, J. S. Morris, S. P. Annangudi and J. V. Sweedler, *Mass spectrometry reviews*, 2008, **27**, 316-330.
27. H. Wang, W.-J. Qian, H. M. Mottaz, T. R. Clauss, D. J. Anderson, R. J. Moore, D. G. Camp, A. H. Khan, D. M. Sforza and M. Pallavicini, *Journal of proteome research*, 2005, **4**, 2397-2403.
28. N. Wang, M. Xu, P. Wang and L. Li, *Analytical chemistry*, 2010, **82**, 2262-2271.
29. N. A. Kulak, G. Pichler, I. Paron, N. Nagaraj and M. Mann, *Nature methods*, 2014, **11**, 319.
30. C. S. Hughes, S. Foehr, D. A. Garfield, E. E. Furlong, L. M. Steinmetz and J. Krijgsveld, *Molecular systems biology*, 2014, **10**, 757.
31. S. Li, B. D. Plouffe, A. M. Belov, S. Ray, X. Wang, S. K. Murthy, B. L. Karger and A. R. Ivanov, *Molecular & Cellular Proteomics*, 2015, mcp. M114. 045724.
32. Q. Chen, G. Yan, M. Gao and X. Zhang, *Analytical chemistry*, 2015, **87**, 6674-6680.
33. E. L. Huang, P. D. Piehowski, D. J. Orton, R. J. Moore, W.-J. Qian, C. P. Casey, X. Sun, S. K. Dey, K. E. Burnum-Johnson and R. D. Smith, *Endocrinology*, 2016, **157**, 1307-1314.
34. J. S. Mellors, K. Jorabchi, L. M. Smith and J. M. Ramsey, *Analytical chemistry*, 2010, **82**, 967-973.
35. Y. Zhu, Y.-X. Zhang, L.-F. Cai and Q. Fang, *Analytical chemistry*, 2013, **85**, 6723-6731.
36. Y. Zhu, Y.-X. Zhang, W.-W. Liu, Y. Ma, Q. Fang and B. Yao, *Scientific reports*, 2015, **5**, 9551.
37. W. Chen, S. Wang, S. Adhikari, Z. Deng, L. Wang, L. Chen, M. Ke, P. Yang and R. Tian, *Anal Chem*, 2016, **88**, 4864-4871.
38. Y. Zhu, M. Dou, P. D. Piehowski, Y. Liang, F. Wang, R. K. Chu, W. Chrisler, J. N. Smith, K. C. Schwarz and Y. Shen, *Molecular & Cellular Proteomics*, 2018, mcp. TIR118. 000686.
39. K. Xu, Y. Zhu, Y. Liang, P. D. Piehowski, M. Dou, K. C. Schwarz, R. Zhao, R. L. Sontag, R. Moore and R. T. Kelly, *Analytical chemistry*, 2018, **Submitted**.
40. Y. Liang, Y. Zhu, M. Dou, K. Xu, R. K. Chu, B. William, R. Zhao, K. K. Hixson and R. T. Kelly, *Analytical chemistry*, 2018, **Submitted**.
41. Y. Zhu, J. Podolak, R. Zhao, A. K. Shukla, R. J. Moore, G. V. Thomas and R. T. Kelly, *Anal Chem*, 2018, **90**, 11756-11759.
42. M. Dou, Y. Zhu, A. Liyu, Y. Liang, J. Chen, P. Piehowski, K. Xu, R. Zhao, R. J. Moore and M. A. Atkinson, *Chemical Science*, 2018, DOI: 10.1039/C8SC02680G.
43. Q. Chen, G. Yan, M. Gao and X. Zhang, *Anal Chem*, 2015, **87**, 6674-6680.
44. J. R. Wisniewski, P. Ostasiewicz and M. Mann, *J Proteome Res*, 2011, **10**, 3040-3049.
45. K. Kasuga, Y. Katoh, K. Nagase and K. Igarashi, *Proteomics*, 2017, **17**.
46. Z. Y. Li, M. Huang, X. K. Wang, Y. Zhu, J. S. Li, C. C. L. Wong and Q. Fang, *Anal Chem*, 2018, **90**, 5430-5438.
47. M. S. Wilm and M. Mann, *International Journal of Mass Spectrometry and Ion Processes*, 1994, **136**, 167-180.
48. M. Wilm and M. Mann, *Analytical chemistry*, 1996, **68**, 1-8.

- 1
- 2
- 3
- 4 49. R. T. Kelly, J. S. Page, Q. Luo, R. J. Moore, D. J. Orton, K. Tang and R. D. Smith, *Analytical chemistry*, 2006, **78**, 7796-7801.
- 5
- 6 50. I. Marginean, K. Tang, R. D. Smith and R. T. Kelly, *Journal of the American Society for Mass Spectrometry*, 2014, **25**, 30-36.
- 7
- 8 51. B. R. Reschke and A. T. Timperman, *Journal of the American Society for Mass Spectrometry*, 2011, **22**, 2115-2124.
- 9
- 10 52. Y. Shen, R. Zhao, S. J. Berger, G. A. Anderson, N. Rodriguez and R. D. Smith, *Analytical chemistry*, 2002, **74**, 4235-4249.
- 11
- 12 53. Y. Shen, N. Tolić, C. Masselon, L. Paša-Tolić, D. G. Camp, K. K. Hixson, R. Zhao, G. A. Anderson and R. D. Smith, *Analytical chemistry*, 2004, **76**, 144-154.
- 13
- 14 54. Y. Shen, R. J. Moore, R. Zhao, J. Blonder, D. L. Auberry, C. Masselon, L. Paša-Tolić, K. K. Hixson, K. J. Auberry and R. D. Smith, *Analytical chemistry*, 2003, **75**, 3596-3605.
- 15
- 16 55. M. Rogeberg, T. Vehus, L. Grutle, T. Greibrokk, S. R. Wilson and E. Lundanes, *Journal of separation science*, 2013, **36**, 2838-2847.
- 17
- 18 56. Q. Luo, G. Yue, G. A. Valaskovic, Y. Gu, S.-L. Wu and B. L. Karger, *Analytical chemistry*, 2007, **79**, 6174-6181.
- 19
- 20 57. L. Sun, G. Zhu, Y. Zhao, X. Yan, S. Mou and N. J. Dovichi, *Angewandte Chemie International Edition*, 2013, **52**, 13661-13664.
- 21
- 22 58. C. Lombard-Banek, S. A. Moody and P. Nemes, *Angewandte Chemie International Edition*, 2016, **55**, 2454-2458.
- 23
- 24 59. R. T. Kelly, A. V. Tolmachev, J. S. Page, K. Tang and R. D. Smith, *Mass spectrometry reviews*, 2010, **29**, 294-312.
- 25
- 26 60. S. A. Shaffer, K. Tang, G. A. Anderson, D. C. Prior, H. R. Udseth and R. D. Smith, *Rapid Commun. Mass Spectrom.*, 1997, **11**, 1813-1817.
- 27
- 28 61. S. A. Shaffer, D. C. Prior, G. A. Anderson, H. R. Udseth and R. D. Smith, *Analytical chemistry*, 1998, **70**, 4111-4119.
- 29
- 30 62. T. Kim, H. R. Udseth and R. D. Smith, *Analytical chemistry*, 2000, **72**, 5014-5019.
- 31
- 32 63. J. S. Page, K. Tang, R. T. Kelly and R. D. Smith, *Analytical chemistry*, 2008, **80**, 1800-1805.
- 33
- 34 64. I. Marginean, J. S. Page, A. V. Tolmachev, K. Tang and R. D. Smith, *Analytical chemistry*, 2010, **82**, 9344-9349.
- 35
- 36 65. H. Suhr, *Plasma Chromatography*, Plenum Press, New York, 1984.
- 37
- 38 66. V. Gabelica and E. Marklund, *Curr Opin Chem Biol*, 2018, **42**, 51-59.
- 39
- 40 67. X. Zhang, K. Quinn, C. Cruickshank-Quinn, R. Reisdorph and N. Reisdorph, *Curr Opin Chem Biol*, 2018, **42**, 60-66.
- 41
- 42 68. G. Ben-Nissan and M. Sharon, *Current Opinion in Chemical Biology*, 2018, **42**, 25-33.
- 43
- 44 69. C. Manz and K. Pagel, *Current Opinion in Chemical Biology*, 2018, **42**, 16-24.
- 45
- 46 70. C. Hinz, S. Liggi and J. L. Griffin, *Current Opinion in Chemical Biology*, 2018, **42**, 42-50.
- 47
- 48 71. L. Deng, S. V. B. Garimella, A. M. Hamid, I. K. Webb, I. K. Attah, R. V. Norheim, S. A. Prost, X. Zheng, J. A. Sandoval, E. S. Baker, Y. M. Ibrahim and R. D. Smith, *Anal Chem*, 2017, **89**, 6432-6439.
- 49
- 50 72. L. Deng, M. Ibrahim Yehia, S. Baker Erin, A. Aly Noor, M. Hamid Ahmed, X. Zhang, X. Zheng, V. B. Garimella Sandilya, K. Webb Ian, A. Prost Spencer, A. Sandoval Jeremy, V. Norheim Randolph, A. Anderson Gordon, V. Tolmachev Aleksey and D. Smith Richard, *ChemistrySelect*, 2016, **1**, 2396-2399.
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

- 1  
2  
3 73. Y. M. Ibrahim, A. M. Hamid, L. Deng, S. V. B. Garimella, I. K. Webb, E. S. Baker and R. D.  
4 Smith, *Analyst*, 2017, **142**, 1010-1021.  
5  
6 74. L. Deng, I. K. Webb, S. V. B. Garimella, A. M. Hamid, X. Zheng, R. V. Norheim, S. A. Prost,  
7 G. A. Anderson, J. A. Sandoval, E. S. Baker, Y. M. Ibrahim and R. D. Smith, *Anal Chem*,  
8 2017, **89**, 4628-4634.  
9  
10 75. S. V. Garimella, A. M. Hamid, L. Deng, Y. M. Ibrahim, I. K. Webb, E. S. Baker, S. A. Prost,  
11 R. V. Norheim, G. A. Anderson and R. D. Smith, *Anal Chem*, 2016, **88**, 11877-11885.  
12  
13 76. L. Deng, Y. M. Ibrahim, A. M. Hamid, S. V. Garimella, I. K. Webb, X. Zheng, S. A. Prost, J.  
14 A. Sandoval, R. V. Norheim, G. A. Anderson, A. V. Tolmachev, E. S. Baker and R. D. Smith,  
15 *Anal Chem*, 2016, **88**, 8957-8964.  
16  
17 77. C. D. Chouinard, G. Nagy, I. K. Webb, T. Shi, E. S. Baker, S. A. Prost, T. Liu, Y. M. Ibrahim  
18 and R. D. Smith, *Anal Chem*, 2018, **90**, 10889-10896.  
19  
20 78. T.-C. Chen, Y. M. Ibrahim, I. K. Webb, S. V. B. Garimella, X. Zhang, A. M. Hamid, L. Deng,  
21 W. E. Karnesky, S. A. Prost, J. A. Sandoval, R. V. Norheim, G. A. Anderson, A. V.  
22 Tolmachev, E. S. Baker and R. D. Smith, *Analytical Chemistry*, 2016, **88**, 1728-1733.  
23  
24 79. T. C. Chen, Y. M. Ibrahim, I. K. Webb, S. V. Garimella, X. Zhang, A. M. Hamid, L. Deng, W.  
25 E. Karnesky, S. A. Prost, J. A. Sandoval, R. V. Norheim, G. A. Anderson, A. V. Tolmachev,  
26 E. S. Baker and R. D. Smith, *Anal Chem*, 2016, **88**, 1728-1733.  
27  
28 80. L. Pasa-Tolic, C. Masselon, R. C. Barry, Y. Shen and R. D. Smith, *Biotechniques*, 2004, **37**,  
29 621-624, 626-633, 636 passim.  
30  
31 81. J. Cox and M. Mann, *Nat Biotechnol*, 2008, **26**, 1367-1372.  
32  
33 82. X. Shen, S. Shen, J. Li, Q. Hu, L. Nie, C. Tu, X. Wang, D. J. Poulsen, B. C. Orsburn, J. Wang  
34 and J. Qu, *Proc Natl Acad Sci U S A*, 2018, **115**, E4767-E4776.  
35  
36 83. M. Murgia, N. Nagaraj, A. S. Deshmukh, M. Zeiler, P. Cancellara, I. Moretti, C. Reggiani,  
37 S. Schiaffino and M. Mann, *EMBO Rep*, 2015, **16**, 387-395.  
38  
39 84. C. S. Hughes, S. Foehr, D. A. Garfield, E. E. Furlong, L. M. Steinmetz and J. Krijgsveld, *Mol*  
40 *Syst Biol*, 2014, **10**, 757.  
41  
42 85. Y. Zhu, R. Zhao, P. D. Piehowski, R. J. Moore, S. Lim, V. J. Orphan, L. Paša-Tolić, W.-J.  
43 Qian, R. D. Smith and R. T. Kelly, *International journal of mass spectrometry*, 2018, **427**,  
44 4-10.  
45  
46 86. J. R. Stanley, J. N. Adkins, G. W. Slysz, M. E. Monroe, S. O. Purvine, Y. V. Karpievitch, G. A.  
47 Anderson, R. D. Smith and A. R. Dabney, *Anal Chem*, 2011, **83**, 6135-6140.  
48  
49 87. F. Meier, S. Beck, N. Grassl, M. Lubeck, M. A. Park, O. Raether and M. Mann, *J Proteome*  
50 *Res*, 2015, **14**, 5378-5387.  
51  
52 88. S. M. Stow, T. J. Causon, X. Zheng, R. T. Kurulugama, T. Mairinger, J. C. May, E. E. Rennie,  
53 E. S. Baker, R. D. Smith, J. A. McLean, S. Hann and J. C. Fjeldsted, *Anal Chem*, 2017, **89**,  
54 9048-9055.  
55  
56 89. K. L. Crowell, E. S. Baker, S. H. Payne, Y. M. Ibrahim, M. E. Monroe, G. W. Slysz, B. L.  
57 LaMarche, V. A. Petyuk, P. D. Piehowski, W. F. Danielson, 3rd, G. A. Anderson and R. D.  
58 Smith, *Int J Mass Spectrom*, 2013, **354-355**, 312-317.  
59  
60 90. L. Moruz, D. Tomazela and L. Kall, *J Proteome Res*, 2010, **9**, 5209-5216.  
91. O. V. Krokhin, S. Ying, J. P. Cortens, D. Ghosh, V. Spicer, W. Ens, K. G. Standing, R. C.  
Beavis and J. A. Wilkins, *Anal Chem*, 2006, **78**, 6265-6269.



- 1
- 2
- 3
- 4 92. K. Petritis, L. J. Kangas, P. L. Ferguson, G. A. Anderson, L. Pasa-Tolic, M. S. Lipton, K. J.
- 5 Auberry, E. F. Strittmatter, Y. Shen, R. Zhao and R. D. Smith, *Anal Chem*, 2003, **75**, 1039-
- 6 1048.
- 7 93. J. K. Eng, A. L. McCormack and J. R. Yates, *J Am Soc Mass Spectrom*, 1994, **5**, 976-989.
- 8 94. D. N. Perkins, D. J. Pappin, D. M. Creasy and J. S. Cottrell, *Electrophoresis*, 1999, **20**,
- 9 3551-3567.
- 10 95. J. E. Kyle, K. L. Crowell, C. P. Casey, G. M. Fujimoto, S. Kim, S. E. Dautel, R. D. Smith, S. H.
- 11 Payne and T. O. Metz, *Bioinformatics*, 2017, **33**, 1744-1746.
- 12 96. T. Kind, K. H. Liu, D. Y. Lee, B. DeFelice, J. K. Meissen and O. Fiehn, *Nat Methods*, 2013,
- 13 **10**, 755-758.
- 14 97. P. Rudd, N. G. Karlsson, K. H. Khoo and N. H. Packer, in *Essentials of Glycobiology*, eds.
- 15 rd, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T.
- 16 Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger, Cold Spring
- 17 Harbor (NY), 2015, DOI: 10.1101/glycobiology.3e.051, pp. 653-666.
- 18 98. H. J. Gabius, *Biosystems*, 2018, **164**, 102-111.
- 19 99. A. M. Boutanaev, T. Moses, J. Zi, D. R. Nelson, S. T. Mugford, R. J. Peters and A. Osbourn,
- 20 *Proc Natl Acad Sci U S A*, 2015, **112**, E81-88.
- 21 100. B. M. Lange, *Annu Rev Plant Biol*, 2015, **66**, 139-159.
- 22 101. S. Neumann and S. Bocker, *Anal Bioanal Chem*, 2010, **398**, 2779-2788.
- 23 102. M. Brown, W. B. Dunn, P. Dobson, Y. Patel, C. L. Winder, S. Francis-McIntyre, P. Begley,
- 24 K. Carroll, D. Broadhurst, A. Tseng, N. Swainston, I. Spasic, R. Goodacre and D. B. Kell,
- 25 *Analyst*, 2009, **134**, 1322-1332.
- 26 103. L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. Fan, O.
- 27 Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka,
- 28 A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden and M. R.
- 29 Viant, *Metabolomics*, 2007, **3**, 211-221.
- 30 104. A. L. Castle, O. Fiehn, R. Kaddurah-Daouk and J. C. Lindon, *Brief Bioinform*, 2006, **7**, 159-
- 31 165.
- 32 105. S. Beisken, M. Eiden and R. M. Salek, *Expert Rev Mol Diagn*, 2015, **15**, 97-109.
- 33 106. K. Uppal, D. I. Walker, K. Liu, S. Li, Y. M. Go and D. P. Jones, *Chem Res Toxicol*, 2016, **29**,
- 34 1956-1975.
- 35 107. N. I. o. Health, RFA-RM-17-013: Compound Identification Development Cores (U2C)).
- 36 108. S. Hou, K. Stevenson and J. J. Harynuk, *J Sep Sci*, 2018, **41**, 2544-2552.
- 37 109. C. B. Mollerup, M. Mardal, P. W. Dalsgaard, K. Linnet and L. P. Barron, *J Chromatogr A*,
- 38 2018, **1542**, 82-88.
- 39 110. M. Hu, E. Muller, E. L. Schymanski, C. Ruttkies, T. Schulze, W. Brack and M. Krauss, *Anal*
- 40 *Bioanal Chem*, 2018, **410**, 1931-1941.
- 41 111. T. Bruderer, E. Varesio and G. Hopfgartner, *J Chromatogr B Analyt Technol Biomed Life*
- 42 *Sci*, 2017, **1071**, 3-10.
- 43 112. J. Stanstrup, S. Neumann and U. Vrhovsek, *Anal Chem*, 2015, **87**, 9421-9428.
- 44 113. S. Wolf, S. Schmidt, M. Muller-Hannemann and S. Neumann, *BMC Bioinformatics*, 2010,
- 45 **11**, 148.
- 46 114. L. Ridder, J. J. van der Hooft, S. Verhoeven, R. C. de Vos, R. van Schaik and J. Vervoort,
- 47 *Rapid Commun Mass Spectrom*, 2012, **26**, 2461-2471.
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
115. F. Allen, A. Pon, M. Wilson, R. Greiner and D. Wishart, *Nucleic Acids Res*, 2014, **42**, W94-99.
116. K. Duhrkop, H. Shen, M. Meusel, J. Rousu and S. Bocker, *Proc Natl Acad Sci U S A*, 2015, **112**, 12580-12585.
117. F. Qiu, D. D. Fine, D. J. Wherritt, Z. Lei and L. W. Sumner, *Anal Chem*, 2016, **88**, 11373-11383.
118. Z. Zhou, X. Shen, J. Tu and Z. J. Zhu, *Anal Chem*, 2016, **88**, 11084-11091.
119. Z. Zhou, J. Tu, X. Xiong, X. Shen and Z. J. Zhu, *Anal Chem*, 2017, **89**, 9559-9566.
120. G. Paglia, J. P. Williams, L. Menikarachchi, J. W. Thompson, R. Tyldesley-Worster, S. Halldorsson, O. Rolfsson, A. Moseley, D. Grant, J. Langridge, B. O. Palsson and G. Astarita, *Anal Chem*, 2014, **86**, 3985-3993.
121. D. F. Smith, D. C. Podgorski, R. P. Rodgers, G. T. Blakney and C. L. Hendrickson, *Anal Chem*, 2018, **90**, 2041-2047.
122. J. B. Shaw, T. Y. Lin, F. E. Leach, 3rd, A. V. Tolmachev, N. Tolic, E. W. Robinson, D. W. Koppenaal and L. Pasa-Tolic, *J Am Soc Mass Spectrom*, 2016, **27**, 1929-1936.
123. D. P. Jones, *Toxicol Rep*, 2016, **3**, 29-45.
124. S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang and S. H. Bryant, *Nucleic Acids Res*, 2016, **44**, D1202-1213.
125. C. M. Dobson, *Nature*, 2004, **432**, 824-828.
126. J. L. Kaszycki and A. A. Shvartsburg, *Journal of the American Society for Mass Spectrometry*, 2017, **28**, 294-302.
127. A. A. Shvartsburg, K. W. M. Siu and D. E. Clemmer, *Journal of the American Society for Mass Spectrometry*, 2001, **12**, 885-888.
128. G. J. Dear, J. Munoz-Muriedas, C. Beaumont, A. Roberts, J. Kirk, J. P. Williams and I. Campuzano, *Rapid Commun. Mass Spectrom.*, 2010, **24**, 3157-3162.
129. X. Zheng, R. S. Renslow, M. M. Makola, I. K. Webb, L. Deng, D. G. Thomas, N. Govind, Y. M. Ibrahim, M. M. Kabanda, I. A. Dubery, H. M. Heyman, R. D. Smith, N. E. Madala and E. S. Baker, *J Phys Chem Lett*, 2017, **8**, 1381-1388.
130. Z. W. Zhou, X. Xiong and Z. J. Zhu, *Bioinformatics*, 2017, **33**, 2235-2237.
131. M. T. Soper-Hopper, A. S. Petrov, J. N. Howard, S. S. Yu, J. G. Forsythe, M. A. Grover and F. M. Fernandez, *Chem Commun (Camb)*, 2017, **53**, 7624-7627.
132. A. Masson, M. Z. Kamrath, M. A. Perez, M. S. Glover, U. Rothlisberger, D. E. Clemmer and T. R. Rizzo, *J Am Soc Mass Spectrom*, 2015, **26**, 1444-1454.
133. M. Z. Kamrath and T. R. Rizzo, *Acc Chem Res*, 2018, **51**, 1487-1495.
134. J. Coates, *Encyclopedia of analytical chemistry*, 2000, **12**, 10815-10837.
135. Y. Bouteiller, J. C. Gillet, G. Gregoire and J. P. Schermann, *J Phys Chem A*, 2008, **112**, 11656-11660.
136. J. Martens, G. Berden, H. Bentlage, K. L. M. Coene, U. F. Engelke, D. Wishart, M. van Scherpenzeel, L. A. J. Kluijtmans, R. A. Wevers and J. Oomens, *J Inherit Metab Dis*, 2018, **41**, 367-377.
137. J. Martens, G. Berden, R. E. van Outersterp, L. A. J. Kluijtmans, U. F. Engelke, C. D. M. van Karnebeek, R. A. Wevers and J. Oomens, *Sci Rep*, 2017, **7**, 3363.

- 1  
2  
3 138. C. Masellis, N. Khanal, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, *J Am Soc Mass*  
4 *Spectrom*, 2017, **28**, 2217-2222.  
5  
6 139. Y. Zhu, G. Clair, W. B. Chrisler, Y. Shen, R. Zhao, A. K. Shukla, R. J. Moore, R. S. Misra, G.  
7 S. Pryhuber, R. D. Smith, C. Ansong and R. T. Kelly, *Angew Chem Int Ed Engl*, 2018, DOI:  
8 10.1002/anie.201802843.  
9  
10 140. F. Meier, P. E. Geyer, S. V. Winter, J. Cox and M. Mann, *Nature methods*, 2018, 1.  
11 141. L. Deng, Y. M. Ibrahim, E. S. Baker, N. A. Aly, A. M. Hamid, X. Zhang, X. Zheng, S. V.  
12 Garimella, I. K. Webb and S. A. Prost, *ChemistrySelect*, 2016, **1**, 2396-2399.  
13 142. L. Deng, Y. M. Ibrahim, A. M. Hamid, S. V. Garimella, I. K. Webb, X. Zheng, S. A. Prost, J.  
14 A. Sandoval, R. V. Norheim and G. A. Anderson, *Analytical chemistry*, 2016, **88**, 8957-  
15 8964.  
16  
17 143. X.-C. Zhang, Q. Zang, H. Zhao, X. Ma, X. Pan, J. Feng, S. Zhang, R. Zhang, Z. Abliz and X.  
18 Zhang, *Analytical Chemistry*, 2018.  
19 144. R. Yin, V. Prabhakaran and J. Laskin, *Analytical Chemistry*, 2018, **90**, 7937-7945.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60