

Offset Mass Carrier Proteome Improves Quantification of Multiplexed Single Cell Proteomics

Authors

Tommy K. Cheung, Ying Zhu, and Christopher M. Rose

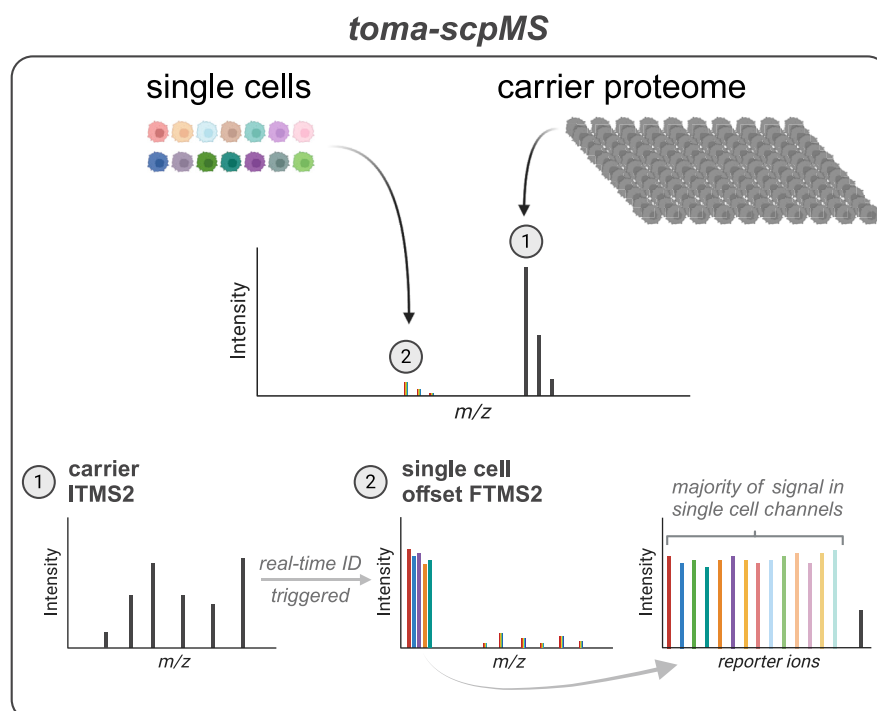
Correspondence

rose.christopher@gene.com

Graphical Abstract

In Brief

A triggered by offset mass acquisition method for scpMS (toma-scpMS) that utilizes a carrier proteome labeled with non-isobaric tags that have a different mass as single cell multiplexing reagents. Within toma-scpMS the carrier proteome and single cell proteome are separated at the precursor level, enabling separate isolation, fragmentation, and quantitation of the single cell samples. We demonstrate that toma-scpMS is more robust to high-levels of carrier proteome and offers superior quantitative accuracy as compared to traditional multiplexed scpMS approaches.



Highlights

- Offset carrier proteome for single cell proteomics using intelligent data acquisition.
- Utilize higher carrier proteome levels without decreasing accuracy of quantification.
- Number of quantified proteins due to decreased signal within quantitative channels.
- FTMS2 and SPS-MS3 analysis of single cells demonstrates minimal ratio compression.



Offset Mass Carrier Proteome Improves Quantification of Multiplexed Single Cell Proteomics

Tommy K. Cheung, Ying Zhu, and Christopher M. Rose^{*} 

Multiplexed single cell proteomics by mass spectrometry (scpMS) approaches currently offer the highest throughput as measured by cells analyzed per day. These methods employ isobaric labels and typically a carrier proteome, a sample added at 20 to 500× the single cell level that improves peptide sampling and identification. Peptides from the carrier and single cell proteomes exist within the same precursor isotopic cluster and are co-isolated for identification and quantification. This represents a challenge as high levels of carrier proteome limit the sampling of peptide ions from single cell samples and can potentially lead to decreased accuracy of quantitative measurements. Here, we address this limitation by introducing a triggered by offset mass acquisition method for scpMS (toma-scpMS) that utilizes a carrier proteome labeled with nonisobaric tags that have the same chemical composition but different mass as the labels used for quantitative multiplexing. Within toma-scpMS, the carrier proteome and single cell proteome are separated at the precursor level, enabling separate isolation, fragmentation, and quantitation of the single cell samples. To enable this workflow, we implemented a custom data acquisition scheme within inSeqAPI, an instrument application programming interface program that performed real-time identification of carrier proteome peptides and subsequent triggering of offset single cell quantification scans. We demonstrate that toma-scpMS is more robust to high levels of carrier proteome and offers superior quantitative accuracy as compared to traditional multiplexed scpMS approaches when similar carrier proteome levels are employed.

Single cell proteomics by mass spectrometry (scpMS) enables the unbiased analysis of proteomes within heterogeneous cell populations without the requirement of antibody-targeted reagents (1). There are two general approaches to scpMS: label-free and isobaric label-based quantitation methods (1).

Label-free methods, such as data-dependent acquisition (2) or data-independent acquisition (DIA) (3–5), offer simplified sample preparation and the ability to identify proteins uniquely

expressed in rare cell types. For these methods, each mass spectrometry (MS) experiment analyzes a single cell with the current potential to analyze 40 to 80 cells per day (3–5). Recently, multiplexed DIA (plexDIA) methods have been introduced that improve throughput by 3 to 5 fold (6, 7), but more difficult sample preparation as well as increased spectral complexity may limit further increases in multiplexing capacity.

Isobaric label-based quantification requires a more complex sample preparation procedure but currently represents the highest-throughput approach that enables deep proteome coverage through the use of a carrier proteome (8, 9). The first application of scpMS for the analysis of cancer cell lines, single cell proteomics by mass spectrometry (SCoPE-MS) (10, 11), introduced the concept of a carrier proteome to enable the analysis of single cell proteomes. A carrier proteome is a sample that is the same or similar to the cells being analyzed and is added into a multiplexed experiment at 20 to 500× the level of the single cells (8, 10–13). Precursor and fragment ions from the carrier proteome overlap, improving sampling and identification of peptides as compared to analyses without a carrier proteome. However, the co-isolation of carrier and single cell peptide ions can challenge the sampling of single cell ions and decrease quantitative accuracy when high levels of carrier proteome are utilized (8, 9). This is due to carrier proteome ions limiting the isolation of single cell ions as well as potential space charge effects that limit the detection of single cell reporter ions when a high level of carrier proteome reporter ion is present (8, 9). These effects can be somewhat overcome by increasing the number of ions sampled as well as the maximum time allowed for ion injection (9, 12, 13). However, due to co-isolation of carrier and single cell precursors, increasing the sampling of single cell peptide ions requires additional sampling of carrier proteome ions, exacerbating potential space charging effects.

An alternative approach would be to use an offset mass carrier proteome to avoid the co-isolation of peptides from the carrier and single cell proteomes (14). This can be accomplished by labeling carrier proteome peptides with nonisobaric

From the Department of Proteomic and Genomic Technologies, Genentech, Inc, South San Francisco, California, USA

^{*}For correspondence: Christopher M. Rose, rose.christopher@gene.com.

versions (*i.e.*, TMT-zero, TMT-SH, TMTpro-zero, and TMTpro-SH) of the quantitative labels used for single cell multiplexing (*i.e.*, TMT and TMTpro). This enables the carrier proteome peptides to be isolated, fragmented, and identified separately from the isolation, fragmentation, and quantification of single cell peptides. Because single cell peptides are isolated separately, more single cell ions can be sampled, leading to an increase of signal-to-noise within single cell reporter ion channels. This improves quantitative accuracy and decreases the negative effects of high levels of carrier proteome reporter ions. A brief exploration of this approach demonstrated that an offset mass carrier proteome increased signal for bulk diluted “single cell” channels, but that higher levels of carrier proteome led to a decrease in quantified proteins (14). Here, the decrease in protein identifications was likely due to the utilization of a “targeted mass difference” acquisition method that only triggered peptide analysis when both the carrier and single cell precursor ions were detected in an MS1 survey scan. Higher carrier levels make it difficult to detect both the carrier and single cell peptides within the same survey scan, leading to decreased peptide sampling and decreased protein identifications. This implementation also identified “single cell” peptides from the same offset scans used for quantification (14). While this ensures that quantified reporter ions are derived from the identified peptide, this approach does not utilize the carrier proteome fragment ions for peptide identification, a key advantage of the carrier proteome within SCoPE-MS.

We sought to address these challenges by employing a novel intelligent data acquisition (IDA) approach that utilized real-time peptide database search to identify peptides from an offset carrier proteome and trigger subsequent single cell-focused scans. IDA methods utilizing real-time peptide identification approaches were introduced more than 10 years ago (15, 16), but these implementations required custom instrument code (16) or difficult to use instrument interfaces (15). Recently, the introduction of a real-time database search within the native instrument method editor of Orbitrap Eclipse and Ascend instruments, as well as the growth of programs that utilize the instrument application programming interface (IAPI) to create custom IDA methods has made the utilization of real-time peptide identification more prevalent (17–20). These methods are often focused on multiplexed quantification utilizing isobaric labels because a rapid-ion trap scan can be used to limit the collection of slow FTMS quantification scans to peptides that will be identified in a post-run search (17, 20, 21). IDA methods have also been used for scpMS, most notably the RETICLE method that utilizes an ion-trap identification (22) scan to trigger an Orbitrap FTMS2 quantification scan, and pSCoPE which utilizes an advanced inclusion list method to prioritize low-abundance peptides and improve quantification across runs (23).

Here, we describe a novel triggered by offset mass scpMS (toma-scpMS) method within inSeqAPI (20) that utilizes a

real-time search to identify carrier proteome peptides and trigger offset mass quantitative analysis of multiplexed single cells peptides. The resulting data enables the connection of carrier proteome identification and single cell quantification scans enabling the carrier proteome scan to be used for post-run peptide identification, maintaining a key advantage of the carrier proteome. Through analysis of equimolar and mixed-ratio bulk-dilution samples, we compared the SCoPE-MS approach to toma-scpMS and demonstrated that toma-scpMS was more robust to the utilization of higher carrier proteome levels as the number of quantified proteins and the quantification accuracy did not decrease as carrier levels increase. Lastly, we utilize toma-scpMS to analyze single cells at varying carrier levels and with FTMS2 or synchronous precursor selection with Orbitrap MS3 scan (SPS-MS3) quantitation to demonstrate the ability to separate cell types and identify cell type-specific proteins.

EXPERIMENTAL PROCEDURES

Diluted Bulk Sample Preparation

Experiments analyzing HeLa and K562 samples used commercial tryptic HeLa (Thermo Scientific Pierce) and K562 digests (Promega). Peptide solutions were dried and then reconstituted in 200 mM Hepes (pH 8.5), before being divided into nine (SCoPE-MS) or eleven (toma-scpMS) aliquots. Aliquots were then labeled with TMT (200 mM Hepes pH 8.5, 28% acetonitrile, 2:1 ratio of label:peptide, reaction time of 1 h), with one carrier proteome aliquot being labeled with channel TMT126 (SCoPE-MS) or TMT-SH (toma-scpMS). For SCoPE-MS, quantitative channels were 127n, 128n, 128c, 129n, 129c, 130n, 130c, 131n, and channel TMT127c was left empty to circumvent the 13C-isotope signal of the carrier proteome (channel TMT126). For toma-scpMS, equimolar quantitative channels were 126, 127n, 127c, 128n, 128c, 129n, 129c, 130n, 130c, 131n, and mixed ratio quantitative channels were 126, 127n, 127c, 128n, 128c, 129n, 129c, 130n. For equimolar mixtures, all quantitative channels were mixed to achieve 400 pg on column per channel per injection. For mixed ratio samples, the final four quantitative channels were mixed to achieve 800 pg on column per channel per injection. After labeling, all samples were dried, desalted, and then dried again before reconstituting in 0.1% formic acid.

To create samples with varying amounts of carrier proteome, each sample was mixed with the appropriate carrier proteome, such that the carrier proteome was at 5× compared to a single quantitative channel. This sample was then serially diluted in a background of carrier proteome, resulting in samples with carrier proteome levels of 20×, 50×, 100×, 250×, 500×, and 1,000×. Performing the dilution in this manner ensured the relative amount of quantitative signals was constant in each experiment.

Single Cell Sample Preparation

HEK293 and HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 1% MEM nonessential amino acid, 2 mM glutamine. The incubator was operated at 37 °C and 5% CO₂. Cells were harvested at 70 to 80% confluence.

HEK293 and HeLa cells were sorted into two types of nanoPOT chips. Single cells were sorted on nested nanoPOT N2 chips (4 × 4 well cluster in 3 × 11 array with a well diameter of 500 μm), and carrier

samples were sorted on nanoPOTS chip (4 × 12 well array with a well diameter of 1200 μm). CellenOne system (Cellenion) was used for cell sorting and sample preparation. Both HEK293 and HeLa cells were stained with 10 μM Calcein AM (Thermo Fisher Scientific) prior to sorting.

For the single cell sample, seven single HEK293 cells and seven single HeLa cells were sorted in the nested nanowells. Ten nanoliters of lysis buffer containing 0.1% n-Dodecyl-B-D-maltoside (DDM), 2 mM TCEP, and 5 mM CAA in 100 mM Hepes, pH 8.0 was dispensed into each nanowell. The cell sorted N2 chip was placed in a preheated humidity box and incubated at 60 °C for 30 min, then cooled back to room temperature. Proteins were digested by adding a mixture of 0.25 ng trypsin and 0.075 ng of lysC in 4 nl of 100 mM Hepes, pH 8.0 and incubated at 37 °C overnight in humidity condition. For TMTPro labeling, we added 10 nl (100 ng) of each of the TMTPro tags dissolved in dimethyl sulfoxide into each of the corresponding nanowells and incubated for 1 h at room temperature. The TMTPro labeling was quenched with 2 nl of 5% hydroxylamine to each nanowell and incubated at room temperature for 15 min.

For the carrier sample, we sorted an equal mixture of HEK293 and HeLa cells on the nanoPOTS chip to reach 20 (10 HeLa & 10 HEK293), 50 (25 HeLa & 25 HEK293), and 100 (50 HeLa & 50 HEK293) cells. Hundred nanoliters of lysis buffer containing 0.1% DDM, 2 mM TCEP, and 5 mM CAA in 100 mM Hepes, pH 8.0 was dispensed into each nanowell. The cell-sorted nanoPOT chip was placed in a preheated humidity box and incubated at 60 °C for 30 min, then cooled back to room temperature. Proteins were digested by adding a mixture of 0.50 ng trypsin and 0.15 ng of lysC in 8 nl of 100 mM Hepes, pH 8.0 and incubated at 37 °C overnight in humidity condition. For TMTPro-SH labeling, we added 50 nl (500 ng) of the TMTPro-SH tag dissolved in dimethyl sulfoxide into each of the corresponding nanowells and incubated for 1 h at room temperature. The TMTPro labeling was quenched with 8 nl of 5% hydroxylamine to each nanowell and incubated at room temperature for 15 min. The single cell and its corresponding carrier samples were pooled together manually with 5 μL of 0.02 % DDM, 0.1 % formic acid and placed in an auto sampler vial prior to LC-MS.

Mass Spectrometry Acquisition

For bulk diluted samples, peptides were separated using a Thermo UltiMate 3000 RSLCnano ProFlow system (Thermo Fisher Scientific) with a gradient of 4% buffer A (98% H₂O, 2% acetonitrile with 0.1% formic acid) to 30% at a flow rate of 300 nl/min over 96 min. Samples were separated over a 25-cm IonOpticks Aurora Series column (IonOpticks). Samples were analyzed by a nanoscale liquid chromatography (nLC)-MS/MS analysis of 125 min using an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific). The native instrument method was used for MS1 and initial ion-trap MS2 analysis. MS1 survey scans were collected in the Orbitrap (200,000 resolution, automated gain control (AGC) = 1×10^6 , maximum injection time (max IT) = 50 ms, 350–1350 m/z range) and used to select the top precursors such that an MS1 was taken every 1 s (Top Speed selection, dynamic exclusion: select one time per 10 s, mass tolerance = 10 ppm) using the MIPS filter, considering charge state 2 to 6 and using a precursor fit filter (fit error = 70% and fit window = 0.5). Precursors were isolated with a 0.5 m/z window and analyzed in the ion-trap (scan rate = Rapid, AGC = 2×10^4 , max IT = 35 ms) following fragmentation with collision induced dissociation (CID) at 35 normalized collision energy (NCE). Resulting ion-trap scans were analyzed further by inSeqAPI (described below).

For single cell samples, peptides were separated using a gradient of 5% buffer A to 11% at 15 min (300 nl/min), 12% at 17 min (300 nl/min to 150 nl/min) to 18.5% at 40 min (150 nl/min), 32% at 114 min

(150 nl/min), 45% at 126 min (150 nl/min) before column washing and equilibration. Samples were loaded onto a precolumn (Acclaim Pep-Map 100, 75 μm × 2 cm, C18, 3 μm and 100 Å particles, Thermo Fisher Scientific) before being separated over a 25-cm IonOpticks Aurora Series column (IonOpticks). Samples were analyzed by an nLC-MS/MS analysis of 140 min using an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific). The native instrument method was used for MS1 and initial ion-trap MS2 analysis. MS1 survey scans were collected in the Orbitrap (200,000 resolution, AGC = 1×10^6 , max IT = 50 ms, 350–1350 m/z range) and used to select the top 15 precursors (dynamic exclusion: single charge state per precursor, select one time per 35 s, mass tolerance = 10 ppm) using the MIPS filter, considering charge state 2 to 6, and using a precursor fit filter (fit error = 70% and fit window = 0.5). Precursors were isolated with a 0.5 m/z window and analyzed in the ion-trap (scan rate = Turbo, analysis mass range 200–1200 m/z, AGC = 1.5×10^4 , max IT = 254 ms) following fragmentation with CID at 35 NCE. Resulting ion-trap scans were analyzed further by inSeqAPI (described below).

inSeqAPI Acquisition Methods

For SCoPE-MS FTMS2, ion-trap scans triggered by the instrument method were analyzed within inSeqAPI using a “Real-time Search Filter” comprising a comet search that was identical to the off-line comet search, except that the database did not contain decoy entries. Following the real-time database search, a “PPM Filter” was utilized to adjust the systematic PPM error. Then an “LDA Filter” was applied using linear-discriminant analysis (LDA) with predetermined LDA coefficients to determine the confidence of the identification. If peptides passed this LDA filter, inSeqAPI sent instructions to the mass spectrometer to perform an FTMS2 scan (isolation width - 0.5 m/z, HCD at 37.5 NCE, Orbitrap, 50,000 resolution) with variable ion sampling parameters (high - 5×10^5 AGC and 750 ms max IT; medium - 3.5×10^5 AGC and 500 ms max IT; low - 3×10^5 AGC and 250 ms max IT), depending on the nLC-MS/MS experiment.

For SCoPE SPS-MS3, ion-trap scans triggered by the instrument were analyzed identically to SCoPE FTMS2, except that additional filters were in place following the LDA filter. Following the LDA filter, potential synchronous precursor selection (SPS) ions were filtered by a “Mass Range Filter” (400–2000 m/z), “Precursor Range Exclusion Filter” (–15 m/z below, 5 m/z above precursor), “Tag Loss Exclusion Filter” (tag = TMT), “SPS Filter” (quant modification = ‘TMT’, tolerance = 0.35 Da), and “TopNFilter” (topN = 6). If all filters were passed, an SPS-MS3 scan was sent to the instrument (CID at 35 followed by HCD at NCE 55, Orbitrap, 50,000 resolution, isolation width 1.2 m/z followed by 2 m/z, 5×10^5 AGC, and 750 ms max IT).

For toma-scpMS FTMS2, ion-trap scans triggered by the instrument were analyzed identically to SCoPE FTMS2, except that an additional filter was in place following the LDA filter. A “Modification Filter” was added that only triggered subsequent scans when the real-time identification was of an offset carrier proteome peptide (exclude modification [–6.0138]). If all filtered were passed, inSeqAPI sent instructions to the mass spectrometer to perform an offset FTMS2 scan (isolation width - 0.5 m/z, HCD at 37.5 NCE, Orbitrap, 50,000 resolution) with variable ion sampling parameters (high - 5×10^5 AGC and 750 ms max IT; medium - 3.5×10^5 AGC and 500 ms max IT; low - 3×10^5 AGC and 250 ms max IT), depending on the nLC-MS/MS experiment. The offset mass FTMS2 isolation m/z was determined by converting any TMT-SH modifications to TMT within the identified peptide in real-time. For single cell samples, the method was identical except for the inSeqAPI-triggered offset FTMS2 AGC target (3.5×10^5), the modification filter (exclude modification [–9.0239]), and that the mass offsets were determined by converting TMTpro-SH to TMTpro modifications within identified peptides in real-time.

For toma-scpMS SPS-MS3, ion-trap scans triggered by the instrument were analyzed identically to toma-scpMS FTMS2. If all filtered were passed, inSeqAPI sent instructions to the mass spectrometer to perform an offset ion-trap MS2 scan (CID at 34.5 NCE, 200–1200 m/z range, isolation width = 0.5 m/z, 1.5×10^4 AGC, 300 ms max IT). The offset mass ion-trap isolation m/z was determined by converting any TMT-SH modifications to TMT within identified peptides in real-time. The offset ion-trap MS2 sent by inSeqAPI was then analyzed by inSeqAPI by a “Percent TIC Filter” (percentTIC = 2), “Mass Range Filter” (400–2000 m/z), “Precursor Range Exclusion Filter” (–15 m/z below and 5 m/z above), “Tag Loss Exclusion Filter” (tag = TMT), “SPS Filter” (quant modification = ‘TMT’, tolerance = 0.35 Da), and “TopNFilter” (topN = 6). If all filters were passed, an SPS-MS3 scan was sent to the instrument (CID at 35 followed by HCD at NCE 55, Orbitrap, 50,000 resolution, isolation width 1.2 m/z followed by 2 m/z, 5×10^5 AGC, and 750 ms max IT). For single cell samples, an identical method was used except that for select parameters in the inSeqAPI-triggered ion-trap MS2 scan (max IT = 254 ms), the modification filter (exclude modification [-9.0239]), the precursor range exclusion filter (455–2000 m/z), the tag loss exclusion filter (tag = TMTpro), SPS-MS3 scan (CID at 34.5 NCE followed by HCD at 45 NCE, isolation width 1.2 m/z followed by 3 m/z, 7.5×10^5 AGC), and mass offsets determined by converting TMTpro-SH to TMTpro modifications within identified peptides in real-time.

Data Analysis

Raw MS data were converted to mzXML using msconvert within ProteoWizard version 3.0.21042 (24). The resulting mzXML files were edited prior to search to ensure that only ion-trap scans triggered by the instrument method were searched. For SCoPE-MS RTS-SPS-MS3, mzXML files were searched without modification. For SCoPE-MS RTS-FTMS2 (RETICLE) and toma-scpMS FTMS2, FTMS2 scans were edited to have a scan order of 3 within the mzXML. For toma-scpMS SPS-MS3, inSeqAPI-triggered offset ion-trap MS2 scans were removed from the mzXML and the parent scan number for the offset SPS-MS3 scan was set to the instrument triggered ion-trap MS2 scan number. For the search of offset FTMS2 scans within toma-scpMS FTMS2 analyses, ion trap scans triggered by the instrument were removed from the mzXML and the previous MS1 for each FTMS2 scan was assigned as the parent scan. For the search of offset ion trap MS2 (ITMS2) scans within toma-scpMS SPS-MS3 analyses, ion trap scans triggered by the instrument were removed from the mzXML and the previous MS1 for each ITMS2 scan was assigned as the parent scan.

Edited mzXML files were searched with Comet version 2020013 (25) against a Human database containing contaminants (UniprotKB Human database downloaded August 2017, 364,690 entries including forward and reverse proteins) (26). For the database search, the precursor mass tolerance was 20 ppm around the monoisotopic mass and a fragment bin tolerance of 1.005 Th with a fragment bin offset of 0.4 Th. An exception was the searches of offset FTMS2 scans within toma-scpMS FTMS2; for these, the fragment bin tolerance was 0.02 Th and the fragment bin offset was 0.0 Th. Peptides were considered following a full *in silico* trypsin (with proline rule) digest allowing for one missed cleavage. For bulk diluted samples, static modifications were TMT-SH on lysine and the peptide n-terminus (235.1767) as well as carbamidomethyl on cysteine (57.0215), and variable modifications were oxidized methionine (15.9949) and TMT on lysine and the peptide n-terminus (–6.0138). For single cell experiments, the static modifications were TMTpro-SH on lysine and the n-terminus (313.2310) as well as carbamidomethyl on cysteine (57.0215) and variable modifications were oxidized methionine (15.9949) and TMTpro on lysine and the peptide n-terminus (–9.0239). For all searches, a maximum of three variable mods per peptide were considered for the database

search. Resulting peptide spectral matches were filtered to 2% false discovery rate (FDR) using Percolator version 3.5 (27). A 2% threshold was chosen for offline peptide FDR because identified peptides were further filtered for cases in which an online identification occurred; the resulting FDR of quantified peptides was typically less than 1%. Results were then filtered to a protein FDR of 1% via a picked-protein approach (28). Quantitative signal-to-noise (SN) values were extracted from raw data using SCPCompanion (9) and combined with resulting identification data within R (29).

For quantitative analysis of equimolar and mixed ratio bulk diluted samples, analysis was performed in R (29). For toma-scpMS equimolar ratios, interference from the carrier within the 131n channel led us to disregard this channel and continue with the analysis as a 9-plex sample. Data within quantitative channels was normalized to account for pipetting errors prior to quantitative analysis. The average signal within quantitative “single cell” channels (QuantSN) was calculated by summing single cell channels and dividing by the number of noncarrier samples within the experiment. The carrier signal was calculated by summing signals within 126 and 127c for SCoPE-MS or 131c and 132 for toma-scpMS. Carrier ratios were calculated at the peptide spectral match (PSM) level by dividing the carrier signal by the average QuantSN. Measurement co-efficient of variation (CV)s were calculated by dividing the SD of quantitative channels by the average QuantSN for each PSM. Ion injection times were also calculated at the PSM level. Quantitative cutoffs were calculated by SCPCompanion and were 126 and 142 QuantSN for SCoPE-MS and toma-scpMS equimolar mixtures, respectively as well as 63.3 QuantSN for mixed ratio samples (based on four 400 pg samples). When SCoPE-MS data was analyzed without the 128c channel, the QuantSN cutoff was set to 47.25 QuantSN (based on three 400 pg samples). The number of quantified proteins reported is the number of nondecoy proteins quantified following the application of protein FDR. For equimolar SCoPE-MS, the 50x carrier displayed a low number of quantified proteins that was not indicative of method performance for both FTMS2 and SPS-MS3 and these points were not plotted.

For HeLa and HEK293 single cell experiments, data were subsequently analyzed in R (29). Carrier ratio was determined by the summed signal of carrier channels divided by the average signal within the single cell channels (average QuantSN). Peptide data were summed to the protein level before application of a summed QuantSN (total signal within single cell channels) cutoff of 220. Reported quantified proteins represent nondecoy proteins following the application of protein FDR. Data from multiple plexes were normalized per protein across all channels before being combined. Principal component analysis was performed within R on normalized and combined data. Log2 ratios were calculated using all HeLa and HEK293 samples across all experiments. Subsequently, *p*-values were calculated using a Student's *t* test (two tails, unequal variance) and adjusted by Benjamini–Hochberg. Protein changes were called significant if the adjusted *p*-value was ≤ 0.05 .

RESULTS

Offset Mass Carrier Proteome for Single Cell Proteomics

Multiplexed scpMS approaches (e.g., SCoPE-MS or improved boosting to amplify signal with isobaric labeling [iBASIL]) utilize a 20 to 500x carrier proteome to enable the detection of precursor peptides within MS1 survey scans and improve the identification of peptides by contributing abundant fragment ions within identification scans (Fig. 1A). While a carrier proteome improves the number of peptides identified within a scpMS experiment, the presence of highly abundant

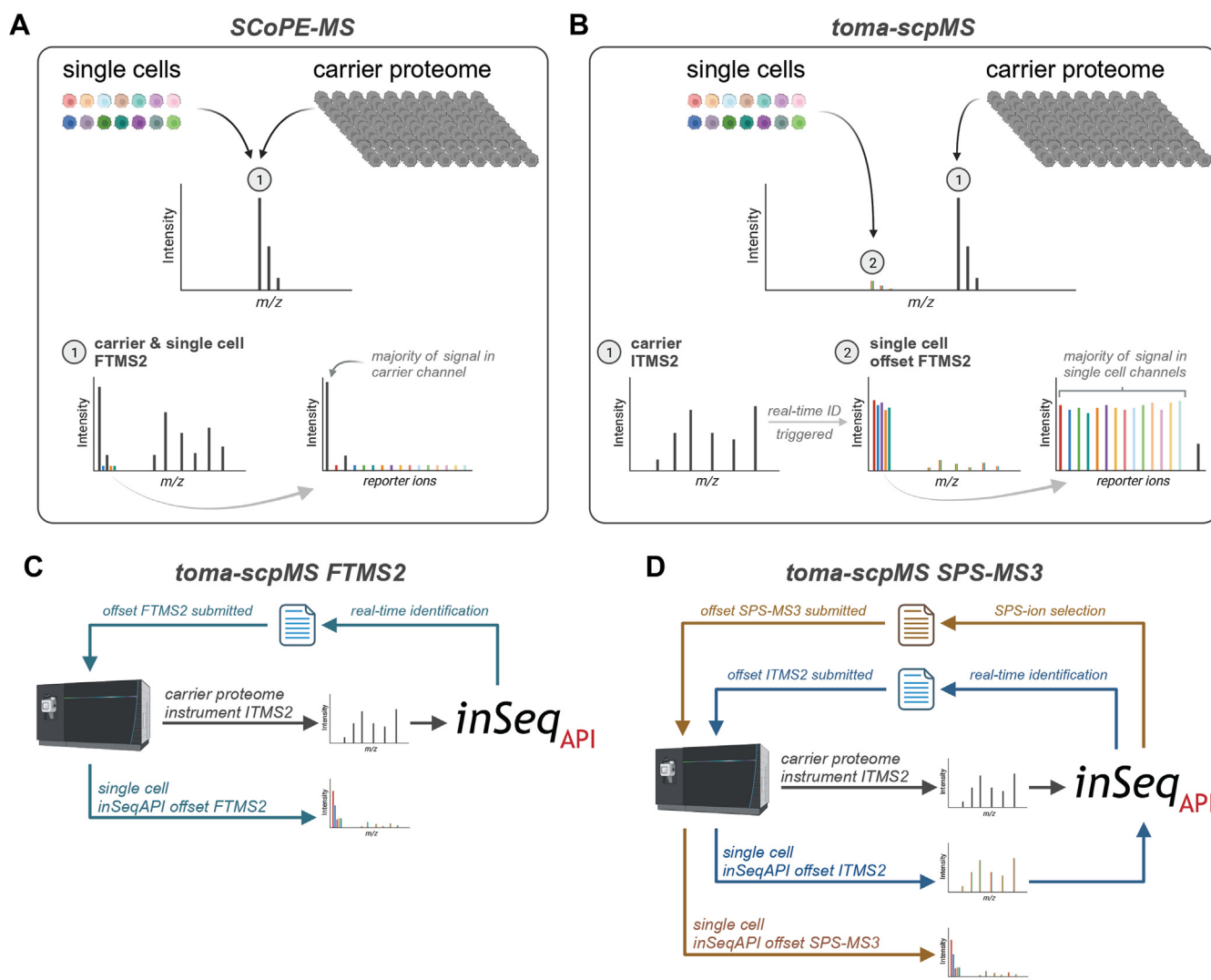


FIG. 1. **SCoPE-MS and toma-scpMS schemes and toma-scpMS inSeqAPI workflow.** A, diagram of single cell proteomics by mass spectrometry (SCoPE-MS) workflow. B, diagram of the targeted by offset mass acquisition single cell proteomics (toma-scpMS) workflow. C, workflow for toma-scpMS FTMS2 inSeqAPI method. D, workflow for toma-scpMS SPS-MS3 inSeqAPI method.

carrier proteome reporter ions may reduce the quantitative accuracy of single cell proteome measurements (9). This effect can be somewhat overcome by increasing ion sampling; however both the carrier and single cell precursor ions are isolated together such that sampling more single cell proteome ions requires additional sampling of carrier proteome ions (Fig. 1A).

We surmised that separating the carrier and single cell proteomes would maintain the benefits of the carrier proteome for peptide identification while also enabling increased sampling of single cell peptides to improve quantification (Fig. 1B). In principle, this approach is similar to the targeted peptide and sample multiplexing approach TOMAHAQ (30, 31); however, we sought to employ this method in an untargeted, discovery workflow. To accomplish this, we utilized the iAPI to

create an IDA program called inSeqAPI that enables a real-time comet (18, 25) database search and subsequent instrument control. Our program, inSeqAPI, is similar to other iAPI-based IDA implementations (17–19) but enables creation of instrument methods similar to the native instrument method editor with methods consisting of one or more “Nodes” which comprise filter groups, “filters,” and “scans” (supplemental Fig. S1, A and B). This structure enables the recapitulation of methods that are built within the native method editor as well as the construction of complex acquisition methods that cannot be typically performed.

Here, we constructed a toma-scpMS that utilized the native instrument method to perform ITMS2 scans and subsequently captured the resulting data for analysis within inSeqAPI (Fig. 1, C and D). MS2 spectra were searched in real-time by comet

against a database comprised of both carrier and single cell proteome peptides, and identification of a carrier proteome peptide prompted inSeqAPI to calculate the m/z of the partner single cell peptide and instruct the mass spectrometer to perform a quantitative scan of the corresponding single cell peptide. Quantitative scans followed either an FTMS2 (Fig. 1C) or SPS-MS3 workflow (Fig. 1D). For FTMS2 (toma-scpMS FTMS2), inSeqAPI instructed the instrument to collect an offset FTMS2 to quantify the offset single cell peptide (Figs. 1C and supplemental Fig. S1A). For SPS-MS3 (toma-scpMS SPS-MS3), inSeqAPI prompted the collection of an offset ITMS2 scan on the single cell peptide; this scan was then analyzed by inSeqAPI to select SPS ions for a subsequent single cell SPS-MS3 scan sent by inSeqAPI (Figs. 1D and supplemental Fig. S1b).

Evaluation of an Offset Mass Carrier Proteome for Equimolar Mixtures

Due to the separate isolation of the carrier and single cell peptides, an offset mass carrier proteome should enable the utilization of higher carrier proteome levels. Due to this, we sought to compare FTMS2 and SPS-MS quantification for toma-scpMS and SCoPE-MS at carrier proteome levels up to 1000 \times . To analyze these samples, four instrument methods were implemented within inSeqAPI. Two of these methods were used to analyze SCoPE-MS samples: 1) “RTS-FTMS2” (RETICLE) (22), an ion trap scan followed by a real-time peptide identification-triggered FTMS2 (supplemental Fig. S2A) and 2) “RTS-SPS-MS3,” an ion trap scan followed by a real-time identification-triggered SPS-MS3 (supplemental Fig. S2B). The two methods described in Figure 1 were used to analyze samples containing an offset mass carrier proteome: 1) “toma-scpMS FTMS2,” an ion trap scan of a carrier peptide followed by a real-time identification-triggered FTMS2 on the single cell peptide (supplemental Fig. S2C) and 2) “toma-scpMS SPS-MS3,” an ion trap scan of a carrier peptide followed by a real-time identification-triggered ion trap scan of a single cell peptide which was used to select SPS ions for a subsequent SPS-MS3 scan of the single cell peptide (supplemental Fig. S2D).

We utilized these methods to analyze separate mixtures of bulk diluted peptides simulating single cell SCoPE-MS and toma-scpMS samples with carrier levels of 5 \times , 20 \times , 50 \times , 100 \times , 250 \times , 500 \times , and 1000 \times (Fig. 2A and supplemental Table S1). For SCoPE-MS, we utilized TMT to create an 8-plex sample with the carrier in channel 126, an empty 127c channel, and the remaining eight channels filled with 400 pg of a tryptic K562 digest (Fig. 2A). For toma-scpMS, we utilized TMT and TMT-SH to create a 10-plex sample with 400 pg of tryptic K562 digest in channels 126-131n and the carrier labeled with TMT-SH (Fig. 2A). Note, the TMT-SH label produces a 132 reporter ion, theoretically enabling more channels to be utilized in the toma-scpMS approach. However, subsequent analysis of the reporter ion intensities demonstrated

that within toma-scpMS FTMS2 data, the 131n channel experienced noticeable interference from the -1 Da impurity of the TMT-SH reporter ion beginning at 100 \times carrier proteome (supplemental Fig. S3, A–D). This was due to the 132 reporter ion containing both the “n” and “c” heavy atoms and producing interference into both the 131n and 131c channels. As a result, the 131n channel was disregarded and toma-scpMS data were analyzed further as a 9-plex experiment.

As described above, the offset mass carrier proteome enables the isolation of single cell peptides separate from the matching peptide within the carrier proteome. However, interfering peptides, with different sequences than the single cell peptide, may be co-isolated during quantitative analysis and result in carrier proteome reporter ion within the quantitative scan of toma-scpMS methods. Previously, we have described how SCPCCompanion measures the observed carrier proteome ratio by dividing the sum of the signals for the carrier reporter ions (including impurities) by the average signal within single cell reporter channels (9). To characterize the level of carrier proteome reporter ions measured from co-isolated carrier proteome (SCoPE-MS) or interfering peptides (toma-scpMS) interference, we calculated the carrier proteome ratio within each quantitative scan for all analyses (Fig. 2B). We found the measured carrier proteome ratio was dramatically lower for toma-scpMS as compared to SCoPE-MS. For example, within the 1000 \times carrier sample, SCoPE-MS exhibited carrier ratios of $\sim 1400\times$ and $\sim 1900\times$ for FTMS2 and SPS-MS3, while toma-scpMS data exhibited carrier ratios of $\sim 100\times$ and $\sim 10\times$ for FTMS2 and SPS-MS3, respectively (Fig. 2B). These data demonstrated that the offset carrier proteome approach was able to decrease the amount of carrier proteome ion by 14 \times for FTMS2 and 190 \times for SPS-MS3 analysis.

Based on the reduction of the co-isolated carrier proteome, we concluded that the majority of the peptide ions isolated within toma-scpMS were single cell ions. This was demonstrated by calculating the average SN per single cell channel (QuantSN) by summing the single cell reporter ion signals and dividing the result by the number of multiplexed samples (Fig. 2C). We found that average QuantSN generally decreased as the carrier level increased, but toma-scpMS methods generally produced $\sim 5\times$ more QuantSN at a given carrier level (Fig. 2C). Furthermore, for toma-scpMS FTMS2, the lower bound of the interquartile range remained above the QuantSN cutoff of 16 SN/channel (red line, Fig. 2C). For both SPS-MS3 methods, the average QuantSN was consistently lower than the matching FTMS2 method, and for SCoPE-MS, only the 5 \times and 10 \times carrier samples exhibited a median QuantSN value at or above the quantification threshold (Fig. 2C). Conversely, toma-scpMS methods maintained a median QuantSN at or above the threshold until the carrier level reached 250 \times (Fig. 2C).

When all quantitative data were considered, increased carrier proteome levels led to a larger number of protein

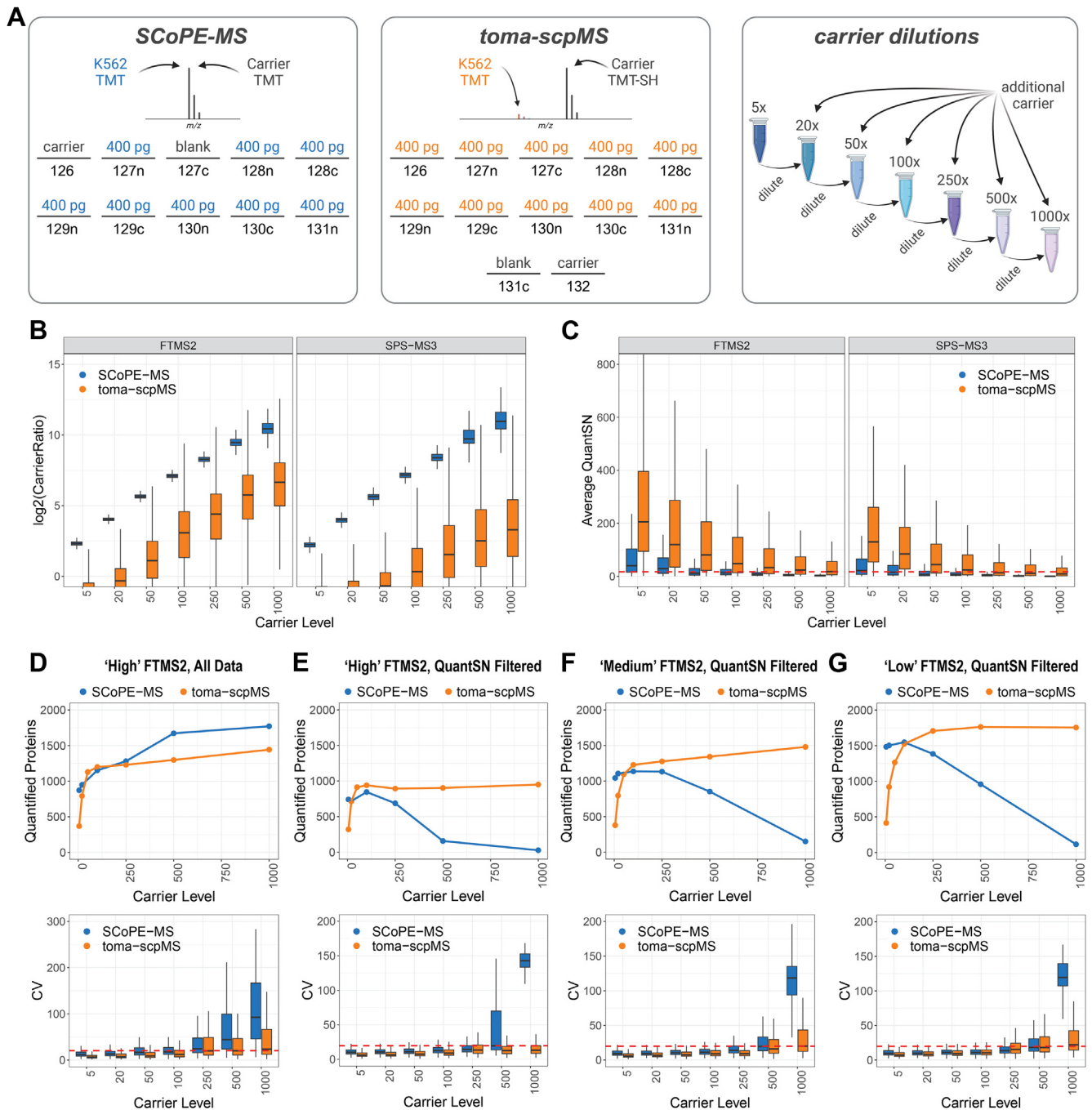


FIG. 2. Comparison of SCoPE-MS and toma-scpMS for the analysis of equimolar mixtures. **A**, schematics describing the sample composition and dilution scheme for the creation of equimolar bulk-diluted "single cell" mixtures. **B**, Log2 transformation of the measured carrier ratio for SCoPE-MS and toma-scpMS as calculated by the sum of carrier reporter ion channels divided by the average of "single cell" channels. **C**, average QuantSN (summed signal-to-noise ratio across "single cell" channels divided by the number of single cell channels) for SCoPE-MS and toma-scpMS. The dotted line represents a QuantSN threshold of 16, the approximate threshold to be considered quantified. **D**, the number of quantified proteins and coefficient of variation (CV) when all data is considered for toma-scpMS FTMS2 using "High" ion sampling parameters. **E–G**, CV and quantified proteins for QuantSN filtered data collected using high, medium, or low ion sampling parameters, respectively. For all box plots, the line denotes the median while the box denotes the interquartile range and the whiskers the highest and lowest value, excluding outliers (not shown).

identifications (Figs. 2D and supplemental Fig. S4A - 'All Data'), particularly for SCoPE-MS. This was due to a dramatic reduction in SCoPE-MS ion injection times for quantification scans within 100-100 \times carrier proteome samples (supplemental Fig. S4B). The decreased injection time was due to the ion injection time calculation being dependent on the precursor peptide peak which is a combination of both the carrier and single cell proteomes within SCoPE-MS. For toma-scpMS, injection times are dependent on the level of the multiplexed single cell proteomes and appeared to be stable until 500 \times carrier, at which point they declined (supplemental Fig. S4B). This decline was attributed to precursor interference within the offset single cell proteome isolation window causing a decrease in the predicted injection time required to reach an ion sampling target.

Previously, we have demonstrated that postanalysis filtering based on SN was a key component of controlling measurement variability (9). Using SCPCompanion, we determined a QuantSN cutoff of 126 for the 8-plex SCoPE-MS samples and 142 for the 9-plex toma-scpMS samples (Figs. 2E and supplemental Fig. S4A - 'QuantSN Filtered') (9). After applying this cutoff, we found that SCoPE-MS produced more quantified proteins at carrier levels below 100 \times (Figs. 2E and supplemental Fig. S4A). This was attributed to increased spectral complexity of toma-scpMS where each peptide was represented by two precursor clusters (one each from the carrier and single cell proteomes) within MS1 survey scans. However, at carrier levels 100 \times and above, the offset carrier proteome became the dominant species within the survey scans and the number of proteins quantified by toma-scpMS matched and then surpassed SCoPE-MS, reaching a plateau ~900 proteins for FTMS2 (Fig. 2E) and ~620 proteins for SPS-MS3 (supplemental Fig. S4A). Conversely, SCoPE-MS experienced a dramatic decrease in quantified proteins above 100 \times carrier proteome (Figs. 2E and supplemental Fig. S4A). This was a result of the previously described challenge of sampling sufficient single cell channels in the presence of high levels of carrier proteome (9). The measurement variability was then assessed by plotting the CV for the quantified proteins in each analysis method (Figs. 2E and supplemental Fig. S4C). Here we found that toma-scpMS consistently produced measurements with lower CVs as compared to SCoPE-MS (Figs. 2E and supplemental Fig. S4C). This result is a direct result of the higher QuantSN demonstrated by toma-scpMS methods because higher intensity measurements lead to lower variability of quantification (9).

Previous efforts to optimize single cell proteomics data depth and quantitation quality have focused on increased ion sampling through adjustment of the AGC and max IT (4, 9, 12). The data above relates to "high" ion sampling conditions (5×10^5 AGC, 750 ms max IT), but we also tested "medium" (3.5×10^5 AGC, 500 ms max IT) and "low" (3×10^5 AGC, 250 ms max IT) ion sampling conditions (Fig. 2, F and G).

Decreasing ion sampling conditions led to an overall increase in quantified proteins; however, the measurement variability of toma-scpMS and SCoPE-MS remained similar to the high sampling conditions. At low carrier levels, SCoPE-MS quantifies more proteins, and at 100 \times carrier proteome, the methods are equal. However, when the carrier proteome is >100 \times , toma-scpMS quantifies more proteins with ~1500 and ~1700 proteins for medium and low ion sampling, respectively (Fig. 2, F and G). Analysis of measurement variability demonstrated that toma-scpMS utilizing medium ion sampling consistently produced measurements with lower CVs as compared to SCoPE-MS (Fig. 2F), but the differences in CV with low ion sampling parameters were negligible (Fig. 2G). Furthermore, for carrier levels above 500 \times for medium and 250 \times for low ion sampling, toma-scpMS quantification became more variable such that the upper bound the interquartile range is above 20% CV (Fig. 2, F and G). This demonstrates that ion sampling parameters must still be carefully considered, even when employing an offset mass carrier proteome.

Lastly, inSeqAPI methods were designed such that 90% of identification scans have associated quantification and 80% of inSeqAPI quantitative scans have an associated identification. From the high ion sampling data, we found that 98% of identification scans had associated quantification for toma-scpMS FTMS2 and both SCoPE-MS approaches (supplemental Table S1). For toma-scpMS SPS-MS3, we found that 81% of the identifications had associated quantification. The lower percentage was likely due to the additional requirement of matching a fragment ion within the offset ITMS2 scan to trigger the quantitative offset SPS-MS3 (supplemental Table S1). We also found 70 to 80% of inSeqAPI triggered quantification scans had an associated identification (supplemental Table S1), which is in the range of the target percentage and demonstrates limited collection of quantitative data that is not used in the final analysis.

Evaluation of an Offset Mass Carrier Proteome for Mixed Ratios

Equimolar mixtures enable the characterization of measurement variability, but accuracy is best measured with a mixed ratio sample. To understand the utilization of an offset mass carrier proteome for the quantification of mixed ratios, we created 8-plex 1:2 mixtures for both SCoPE-MS and toma-scpMS where 1 \times and 2 \times channels contained 400 pg and 800 pg on column, respectively (Figs. 3 and supplemental Table S2). We combined this mixture with carrier proteome to produce samples with 5 \times , 20 \times , 50 \times , 100 \times , 250 \times , 500 \times , and 1000 \times carrier proteome (Fig. 3A). These samples were analyzed with the previously described methods (supplemental Fig. S2) and the resulting data were filtered with

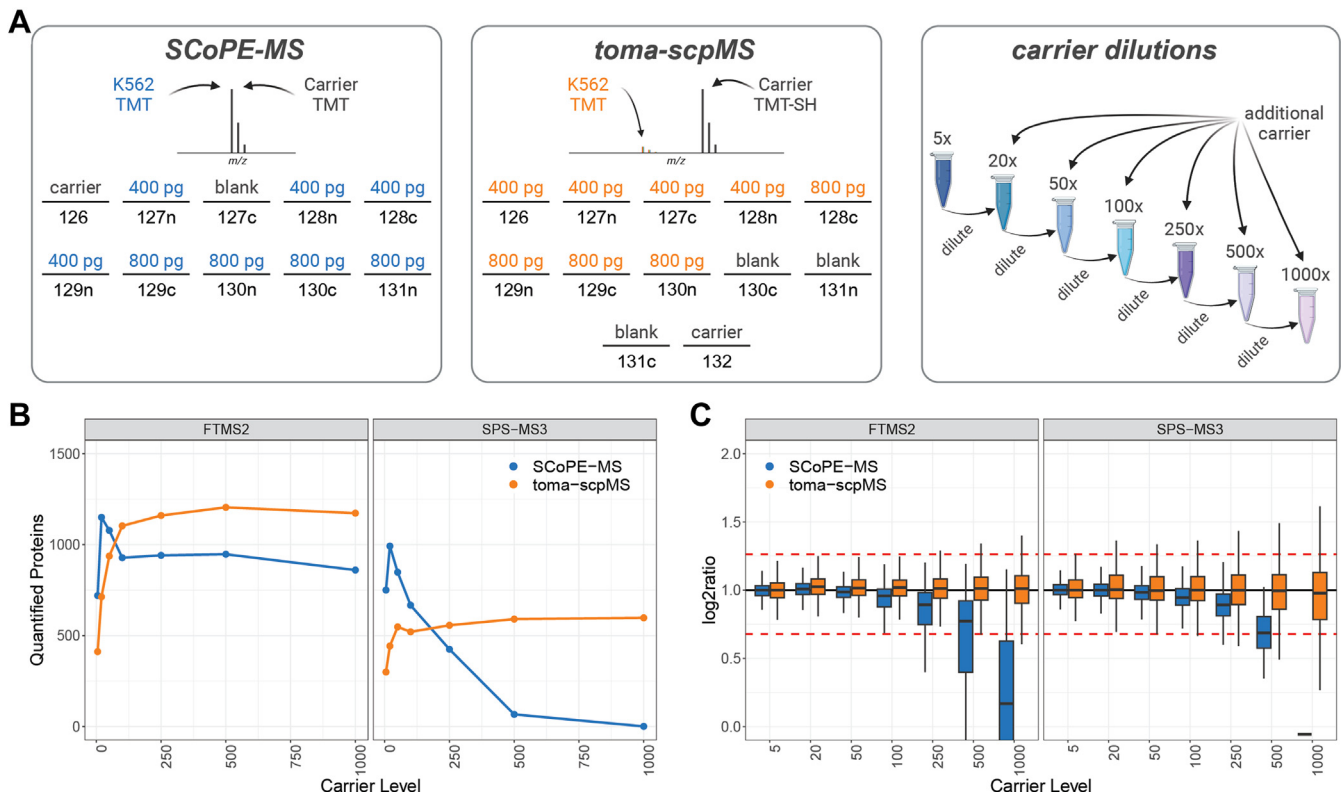


FIG. 3. Comparison of SCoPE-MS and toma-scpMS for the analysis of mixed ratio samples. A, schematics describing the sample composition and dilution scheme for the creation of mixed ratio bulk-diluted “single cell” mixtures. B, the number of quantified proteins after applying a QuantSN cutoff at the peptide spectral match (PSM) level. C, distribution of measured protein ratios for the mixed ratio. The target ratio was 2 ($\log_2\text{ratio} = 1$, black horizontal line) and the red dotted lines represent a 20% measurement error. For box plots, the line denotes the median while the box denotes the interquartile range and the whiskers the highest and lowest value, excluding outliers (not shown).

a QuantSN ≥ 126 SN. Similar to 1:1 mixtures, SCoPE-MS outperforms toma-scpMS for low carrier levels (5x–50x); however at high carrier proteome levels (250x–1000x), toma-scpMS quantifies more proteins as compared to SCoPE-MS (Fig. 3B).

We then calculated the fold change for each quantified protein to determine the quantitative accuracy of SCoPE-MS and toma-scpMS (Fig. 3C). Curiously, we found that SCoPE-MS demonstrated noticeable ratio compression with as little as 100x carrier proteome (Fig. 3C). This was particularly interesting because these samples did not contain a background interference proteome, demonstrating that ratio compression will occur at high carrier proteome levels independent of background interference. Ratio compression was not detected for toma-scpMS FTMS at any carrier proteome level, but a slight compression was present for the 1000x carrier for toma-scpMS SPS-MS3 (Fig. 3C). There was a slight increase in the median ratio for toma-scpMS FTMS2 that may have been due to interference from the -2 Da impurity of the carrier proteome label within the 130n and 130c reporter ion channels (Fig. 3C). This is supported by the elimination of this ratio increase for toma-scpMS SPS-MS3, where the amount of carrier proteome measured in the quantification scan is

lower than toma-scpMS FTMS2 (supplemental Fig. S5). Together these data demonstrate that toma-scpMS provides deeper proteome coverage and more accurate quantification data at higher carrier levels as compared to SCoPE-MS.

Single Cell Analysis Utilizing an Offset Mass Carrier Proteome

Having demonstrated the advantages of the toma-scpMS approach on diluted bulk samples, we next employed toma-scpMS FTMS2 and SPS-MS3 for the analysis of single cells (Fig. 4 and supplemental Table S3). HeLa and HEK293 cells were analyzed by multiplexed experiments containing 14 single cells (seven HeLa and seven HEK293) and either 20x, 50x, or 100x offset carrier proteome (Fig. 4A). Sample preparation was performed using a CellenOne system to sort single cells into either an N2 (32) or nanoPOTS (33) chip for the multiplexed single cell or the carrier proteome preparation, respectively (Fig. 4A). For the carrier proteome, 20, 50, or 100 cells (equal mixture of HeLa and HEK293) were prepared alongside single cells to better match the single cell and carrier proteomes. Cells were reduced, alkylated, digested with trypsin/lysC, and labeled with either the appropriate TMTpro channel for single cells or TMTpro-SH for carrier

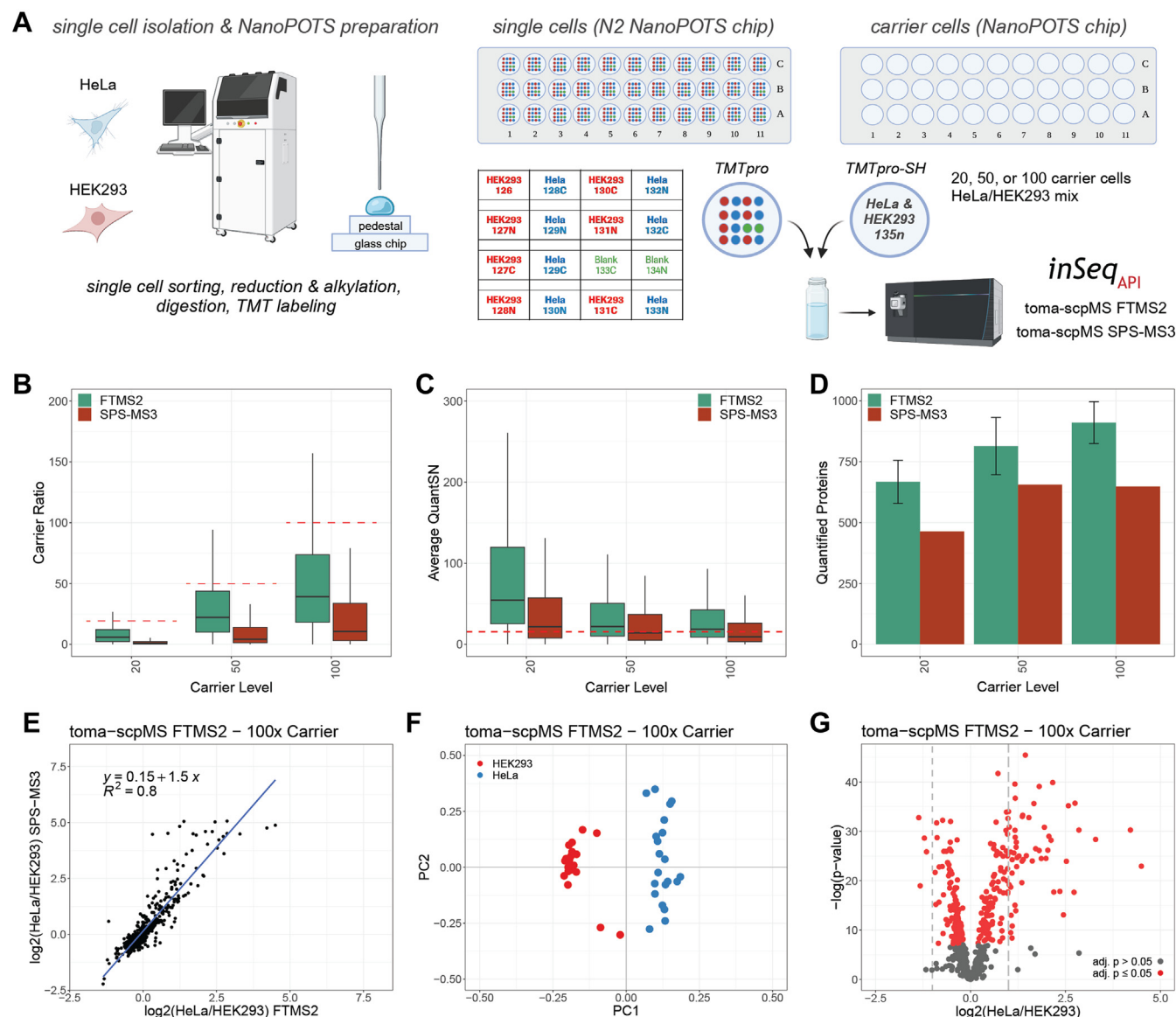


FIG. 4. Analysis of single cells from two cell types with toma-scpMS. A, schematic describing the preparation of single cells for analysis by toma-scpMS. Cells were isolated and prepared by the CellenOne instrument using either an N2 NanoPOTS chip (single cells) or a NanoPOTS chip (carrier proteome). The single cells and carrier cells were labeled as described for all multiplexed experiments. Three plexes were analyzed with toma-scpMS FTMS2 for the 20x, 50x, and 100x carrier proteome levels, while one plex was analyzed for each carrier level for toma-scpMS SPS-MS3. B, the measured carrier ratio as calculated by the sum of carrier reporter ion channels divided by the average of single cell channels. C, average QuantSN (summed signal-to-noise ratio across single cell channels divided by the number of single cell channels). The dotted line represents a QuantSN threshold of 16, the approximate threshold to be considered quantified. For all box plots in (C and D), the black line denotes the median while the box denotes the interquartile range and the whiskers the highest and lowest value, excluding outliers (not shown). D, the number of quantified proteins for each carrier level analyzed by toma-scpMS FTMS2 or SPS-MS3. For FTMS2, bars represent the average number of proteins quantified across three plexes and the bars represent \pm SD. E, correlation of toma-scpMS FTMS2 and SPS-MS3 measurements for proteins quantified in both methods. F, principal component analysis based on proteins quantified across all three plexes differentiates cell types. G, volcano plot displaying the differential analysis of protein abundance across the two cell types.

proteomes (Fig. 4A). Following sample preparation, multiplexed single cells were combined with carrier proteome and analyzed with either a toma-scpMS FTMS2 (three plexes totaling 42 single cells per carrier level) or SPS-MS3 method (one plex totaling 14 single cells per carrier level) (Fig. 4A).

Data were collected with 3.0×10^5 AGC, 750 ms max IT ion sampling parameters for quantitative scans.

As described above, an offset carrier proteome will lead to the co-isolation of single cell and carrier proteome peptides of different sequences. To determine the level of co-isolated

interference, we calculated the carrier proteome ratio for each sample and toma-scpMS method (Fig. 4B). For toma-scpMS FTMS2, we found that the measured carrier proteome level was generally $\leq 50\%$ of the target carrier proteome level, measuring 6 \times (20 \times), 22 \times (50 \times), and 39 \times (100 \times) (Fig. 4B). For toma-scpMS SPS-MS3, the estimated carrier proteome ratio was decreased as much as 10-fold, measuring 1 \times (20 \times), 4 \times (50 \times), and 10 \times (100 \times) (Fig. 4B). By decreasing the amount of carrier proteome reporter ions, we increase the dynamic range of single cell ion measurements and enable more accurate quantification, as supported by the diluted bulk mixed ratio samples above.

While the dramatic reduction of the carrier proteome reporter ions appears to favor quantitation by toma-scpMS SPS-MS3, the gas phase purification leads to a reduction of QuantSN. To characterize the level of QuantSN for both methods, the summed QuantSN was extracted for all PSMs within each sample and data collection method. We found a general decrease in QuantSN as the level of carrier proteome increases, matching the results of our diluted bulk samples above (Fig. 4C). For toma-scpMS FTMS2, a large proportion of PSM measurements were above the average QuantSN cutoff of ~ 16 SN calculated by SCPCompainon for all carrier levels (red line, Fig. 4C). This was only the case for the 20 \times carrier sample analyzed by toma-scpMS SPS-MS3, as $\sim 50\%$ of PSMs failed to acquire more than 220 SN for the 50 \times and 100 \times carrier samples (Fig. 4C). After applying the QuantSN filter, we found that toma-scpMS FTMS on average quantified 667, 814, and 911 proteins for the 20 \times , 50 \times , and 100 \times carrier proteome samples, respectively (Fig. 4D). The number of quantified proteins per plex was lower with toma-scpMS SPS-MS3 as 464, 656, and 648 proteins were quantified in the 20 \times , 50 \times , and 100 \times carrier proteome samples, respectively (Fig. 4D). Interestingly, it appears that the number of proteins quantified by toma-scpMS FTMS2 increased from 50 \times to 100 \times carrier proteome, while the number of quantified proteins remained roughly the same for toma-scpMS SPS-MS3 (Fig. 4D). This was likely due to the lower QuantSN observed for toma-scpMS SPS-MS3 scans leading to more data being filtered out due to falling below the quantitative threshold.

While toma-scpMS SPS-MS3 quantified fewer proteins, it was possible that the gas phase purification improved quantification accuracy such that the dynamic range of measurements was larger. To understand the potential impact of ratio compression within toma-scpMS, we compared the \log_2 (HeLa/HEK293) ratio for proteins in toma-scpMS FTMS2 as compared to SPS-MS3 for each carrier proteome level (Figs. 4E and supplemental Fig. S6, A and B). The slope of a calculated regression line can be used to characterize the level of ratio compression as values larger than one relate to higher ratios measured with the SPS-MS3 approach. Surprisingly, we found a moderate impact of ratio compression as the slope of the regression lines were 1.3, 1, and 1.5 for the 20 \times , 50 \times , and 100 \times carrier levels, respectively (Figs. 4E and supplemental

Fig. S6, A and B). Note, the low level of ratio compression may have been due to these data representing high abundance proteins that are less impacted by precursor interference. Nonetheless, these results are encouraging because SPS-MS3 analysis dramatically decreases QuantSN for single cell measurements, making toma-scpMS FTMS2 the preferred method.

Lastly, we sought to demonstrate that toma-scpMS FTMS2 data is sufficient to separate cell types and determine cell-type-specific proteins. To demonstrate a separation of cell types, we performed principal component analysis and found that each toma-scpMS approach easily separated cell types (Figs. 4F and supplemental Fig. S6, C–G). Interestingly, HeLa cells demonstrated higher variability as compared to HEK293 cells across all experiments; we attribute this to variations in HeLa cell sizes as there were no clear associations with TMT-plex or channel (Figs. 4F and supplemental Fig. S6, C–G). We calculated the \log_2 (HeLa/HEK293) ratios and determined significantly changing proteins ($p < 0.05$, student's t test, two tails, unequal variance, Benjamini–Hochberg correction) (Figs. 4G and supplemental Fig. S6, H and I). For FTMS2, we found a similar number of significantly changing proteins (20 \times : 212, 50 \times : 290, 100 \times : 260) as well as significantly changing proteins with levels altered more than 2-fold (20 \times : 68, 50 \times : 68, and 100 \times : 64) (Fig. 4G and supplemental Fig. S6, H and I). These results demonstrate that a similar number of significantly changing proteins can be detected irrespective of carrier proteome level when employing toma-scpMS FTMS2.

DISCUSSION

scpMS was greatly accelerated by SCoPE-MS, the first single cell analysis of cancer cell lines (10). An enabling feature of the initial SCoPE-MS approach was the carrier proteome, a sample containing 200 cells that was added to improve peptide selection and identification (10). Since this description, the carrier proteome has been extensively studied and the field has reached a consensus on the importance of optimizing ion sampling (AGC and max IT) for producing low-variability, high accuracy data. However, the level of carrier proteome used within scpMS remains variable with as little as 20 \times or as much as 500 \times carrier proteome being employed. The continued reliance on higher carrier proteome levels suggests that an approach that is more robust to high carrier proteome levels would be important for scpMS analyses.

Here we introduce a toma-scpMS that utilizes an offset carrier proteome that is more robust to the utilization of higher carrier proteome levels as compared to SCoPE-MS. To enable data acquisition of samples utilizing an offset mass carrier proteome, we constructed an IDA method within inSeqAPI, an iAPI program that enables real-time analysis of data and instrument control. Four inSeqAPI methods were constructed to analyze toma-scpMS or SCoPE-MS samples by FTMS2 or

SPS-MS3. While the SCoPE-MS methods implemented within inSeqAPI can be reproduced through the native method editor, the toma-scpMS methods are only possible when utilizing the iAPI. This is because implementing toma-scpMS required the conversion of an identified carrier proteome peptide to the equivalent single cell proteome peptide. To accomplish this, identified peptides were reduced to their unmodified form, and modifications corresponding to the single cell proteome were added to the base peptide sequence. This ensured the appropriate precursor, fragment, and potential SPS ions were calculated for toma-scpMS FTMS or SPS-MS3 analysis. Currently this logic is not available within the native method editor.

Through the analysis of equimolar diluted bulk samples, we compared measurement variability for SCoPE-MS and toma-scpMS in the presence of carrier levels ranging from 5 \times to 1000 \times . These data demonstrated that toma-scpMS produced higher levels of reporter ion signal within single cell channels (QuantSN), leading to lower variability measurements across all carrier proteome levels. Interestingly, we noted that the number of quantified proteins was higher for SCoPE-MS at the 5 \times and 20 \times carrier proteome levels. This was likely due to a greater spectral complexity for the toma-scpMS samples as both the single cell and carrier proteome peptides are readily detectable within an MS1 survey scan at low carrier levels. The two methods appeared to quantify an equal number of proteins at 100 \times carrier proteome level, before the number of quantified proteins dropped precipitously for SCoPE-MS. Conversely, toma-scpMS maintained the number of quantified proteins as the carrier level increased, demonstrating a greater robustness to the level of carrier proteome employed.

By analyzing a diluted bulk 1:2 mixed ratio sample, we were able to compare the accuracy of toma-scpMS and SCoPE-MS. While the trend in quantified proteins was similar to that detected in the equimolar mixtures, the analysis of the ratios revealed ratio compression within SCoPE-MS beginning at 50 \times carrier proteome. Interestingly, these samples did not contain an equimolar interference proteome, so the ratio compression could not be due to precursor co-isolation. We attribute the ratio compression within SCoPE-MS to a truncation of the single cell reporter ion channels due to a highly abundant carrier proteome reporter ion channel. We previously reported this effect and demonstrated how the removal of the carrier proteome reporter ion can increase the dynamic range of measured single cell signals (9). This appears to be a fundamental limitation of utilizing high levels of carrier proteomes within SCoPE-MS. Unlike SCoPE-MS, toma-scpMS did not experience ratio compression within the mixed ratio sample, rather we observed ratio expansion with high levels of carrier proteome and toma-scpMS FTMS analysis. We surmised this was due to -2 Da interference from co-isolated carrier proteome reporter ion interfering with the single cell signals in the 130n and 130c channels. This was supported by the elimination of this effect for toma-scpMS SPS-MS3 where

interfering co-isolated peptide signal was drastically reduced. Together these data strongly support the utilization of toma-scpMS when utilizing carrier levels $\geq 100\times$.

Through the analysis of single cells, we demonstrated that toma-scpMS can produce high quality data using either an FTMS2 or SPS-MS3 quantification strategy. Within these single cell samples, we demonstrate that the measured carrier proteome ratio is less than 50% the amount within the sample for toma-scpMS FTMS and less than 10% for toma-scpMS SPS-MS3, decreasing the negative impact of highly abundant carrier reporter ions. These data also demonstrated that increased carrier proteome levels reduced QuantSN within single cell channels such that the number of proteins plateaued for SPS-MS3 quantification. The number of quantified proteins with FTMS2 increased from 50 \times to 100 \times carrier proteome; however, the percentage of measurements above the quantitative threshold fell. This suggests that increasing the carrier proteome further would require a concomitant increase in ion sampling which would likely reduce quantified proteins due to longer duty cycle. Encouragingly, when comparing log₂(HeLa/HEK293) ratios for toma-scpMS FTMS2 and SPS-MS3 analyses, we found only minor ratio compression. From these data, we conclude that FTMS2 data acquisition remains superior to SPS-MS3 due to higher reporter ion signal and that 100 \times carrier proteome would be the recommended limit for offset carrier proteome within toma-scpMS. Lastly, we demonstrate that toma-scpMS FTMS2 acquisition produces data that enables separation of cell types and the identification of cell-type-specific proteins. These results demonstrated larger variability in HeLa cells as compared to HEK293 cells; however, this variation could not be associated with TMT-plex or label and was likely due to differences in HeLa cell size. This was supported by observations of cell size variance during cell sorting.

Previous exploration of using an offset carrier proteome utilized the offset scan for both identification and quantification but found that the number of identified peptides and proteins decreased substantially as carrier levels increased (14). An important consideration for toma-scpMS is that identifications are determined by carrier proteome scans, and single cell reporter ion signals are extracted from offset quantitative scans. One concern with this approach is that single cell reporter ions measured in an offset scan are not derived from the identified carrier proteome peptide. We note here that this is also a concern for SCoPE-MS, as reporter ion signals within single cell channels may be derived from co-isolated interference. For SCoPE-MS, the carrier proteome fragment ions mask detection of single cell peptide fragment ions that would demonstrate the detectable presence of single cell ions. From our data, the correlation of toma-scpMS FTMS2 and SPS-MS3 ratios demonstrates that quantification within toma-scpMS is faithful to the identified peptides. This is due to SPS-MS3 data being highly specific as a result of isolating the single cell precursor and then subsequently single cell fragment ions before quantification.

To address this concern more directly, offset FTMS2 scans (toma-scpMS FTMS2) and offset ion trap MS2 scans (toms-scpMS SPS-MS3) collected for the single cell experiments were searched to understand if peptides could be identified from offset scans. For toma-scpMS FTMS2, ~2500 peptides were identified per run without a clear dependence on the carrier proteome level (supplemental Table S4). This is in contrast to the carrier proteome where the IDs ranged from ~2500 (20×) to ~4200 (100×) depending on the carrier proteome level (supplemental Table S4). For spectra that had a peptide identification in both analyses, we identified the same peptide in 97 to 99% of cases, in line with the FDR used for analysis (supplemental Table S4). These data demonstrate that the offset FTMS2 scan is representative of the peptide identified from the carrier proteome and can be confidently used for quantification.

For toma-scpMS SPS-MS3, ~2000 peptides were identified per run without a clear dependence on the carrier proteome level (supplemental Table S4). The number of peptide IDs ranged from ~2000 (20×) to ~4000 (100×) depending on the carrier proteome level (supplemental Table S4). Here, for spectra that had a peptide identification in both analyses, the same peptide was identified in 92 to 93% of the cases (supplemental Table S4). We believe the matching percentage is lower due to the offset scan being low resolution leading to more ambiguity for spectra with less signal. However, for a quantitative SPS-MS3 to be triggered, we would need to have an online identification of the carrier proteome peptide to trigger an offset ITMS2 and then a fragment ion from the single cells would need to be detected in the offset ITMS2 to trigger an SPS-MS3. Due to this, even though the matching percentage of the peptide identified from a search of the offset ITMS2 is lower, SPS-MS3 scans should still be highly specific as a result of the numerous steps required to trigger a scan.

Lastly, a limitation of the toma-scpMS approach was found to be the identification percentage. The percentage identification scans confidently assigned to a carrier proteome peptide sequence was ~9% even though a maximum injection time of 250 ms was used due to the concern of accumulating enough ions in a single cell experiment (supplemental Table S4). Further analysis demonstrated that the median injection time was 70 ms across all single cell analysis, demonstrating that there are some improvements to be made in the data acquisition method, but the low identification rate was likely not related to ion injection times (supplemental Table S4). We then calculated the identification percentage for the standard mixture samples (Figs. 2 and 3) and found that toma-scpMS typically identifies peptides in ~9% of the scans while SCoPE-MS identified peptides in ~14% of the scans (supplemental Table S4). This disparity could be due to the size of the database as toma-scpMS contains additional variable modifications which will increase the number of

potential matches and potentially lead to more low quality spectra being assigned to decoy peptides. Utilizing machine learning-based rescoring would likely improve toma-scpMS identification percentage, similar to the increases reported for immunopeptidomics datasets.

Overall, the data presented support the implementation of a toma-scpMS experiment utilizing an offset mass carrier proteome and FTMS2 quantification. This approach will quantify a similar number of proteins as SCoPE-MS but produce higher quality measurements that are less impacted by ratio compression caused by the presence of high amounts of carrier reporter ions. Furthermore, toma-scpMS is more robust to high levels of carrier proteome and could be valuable in cases where the level of carrier proteome added to a scpMS experiment cannot be carefully controlled.

EXPERIMENTAL DESIGN AND STATISTICAL RATIONAL

For standard mixture experiments (Figs. 2 and 3), data collected were $N = 1$ for each carrier proteome dilution. For the mixed ratio, K562 standard mixture (Fig. 3) ratios were determined by $N = 4$ for the low and high channels. The exception was the analysis of SCoPE-MS omitting channel 128c; this comparison had only $N = 3$ for the high channels. For this analysis, the difference between the expected and measured ratio was calculated, but there was not a statistical significance assigned to the measurement. For the analysis of single cells, each TMT plex contained $N = 7$ for HEK293 and HeLa cells (Fig. 4). For toma-scpMS FTMS2 analysis, three TMT plexes per carrier level were analyzed; for SPS-MS3 analysis, one TMT plex per carrier level was analyzed. For statistical analysis of toma-scpMS FTMS2 single cell data, results from all three TMT plexes were combined ($N = 21$ for each cell line) before calculating ratios and significance estimates (p -value: Student's t test, two tailed, unequal variance; adjusted p -value: Benjamini-Hochberg correction).

DATA AVAILABILITY

Raw data have been deposited in MASSIVE (34) with the accession MSV000095629. The inSeqAPI program is available upon request and requires an instrument API agreement with ThermoFisher.

Supplemental data—This article contains supplemental data.

Acknowledgments—We thank Sarah Williams for help with the nanoPOTS chip as well as members of the Proteomic and Genomic Technologies department at Genentech for helpful discussions during the writing of the manuscript.

Author contribution—T. K. C., Y. Z., and C. M. R. writing—review and editing; T. K. C. and C. M. R. methodology; T. K.

C., Y. Z., and C. M. R. investigation; T. K. C. and C. M. R. data curation; Y. Z. and C. M. R. conceptualization; C. M. R. writing—original draft; C. M. R. visualization; C. M. R. validation; C. M. R. supervision; C. M. R. software; C. M. R. resources; C. M. R. project administration; C. M. R. funding acquisition; C. M. R. formal analysis.

Conflicts of interest—T. K. C., Y. Z., and C. M. R. are employees of Genentech, Inc. and shareholders of Roche.

Abbreviations—The abbreviations used are: AGC, automated gain control; CID, collision induced dissociation; CV, co-efficient of variation; DDM, n-Dodecyl-B-D-maltoside; DIA, data-independent acquisition; FDR, false discovery rate; iAPI, instrument application programming interface; IDA, intelligent data acquisition; ITMS2, ion trap MS2; LDA, linear discriminant analysis; max IT, maximum injection time; MS, mass spectrometry; NCE, normalized collision energy; nLC, nano-scale liquid chromatography; PSM, peptide spectral match; scpMS, single cell proteomics by mass spectrometry; SCoPE-MS, single cell proteomics mass spectrometry; SN, signal-to-noise; SPS, synchronous precursor selection; SPS-MS3, synchronous precursor selection with Orbitrap MS3 scan; toma-scpMS, triggered by offset mass acquisition method for scpMS.

Received November 9, 2024, and in revised form, March 6, 2025
Published, MCPRO Papers in Press, April 1, 2025, <https://doi.org/10.1016/j.mcpro.2025.100959>

REFERENCES

- Bennett, H. M., Stephenson, W., Rose, C. M., and Darmanis, S. (2023) Single-cell proteomics enabled by next-generation sequencing or mass spectrometry. *Nat. Methods* **20**, 363–374
- Zhu, Y., Clair, G., Chrisler, W. B., Shen, Y., Zhao, R., Shukla, A. K., et al. (2018) Proteomic analysis of single mammalian cells enabled by microfluidic nanodroplet sample preparation and ultrasensitive NanoLC-MS. *Angew. Chem. Int. Ed. Engl.* **57**, 12370–12374
- Brunner, A., Thielert, M., Vasilopoulou, C., Ammar, C., Coscia, F., Mund, A., et al. (2022) Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol. Syst. Biol.* **18**, e10798
- Petrosius, V., Aragon-Fernandez, P., Üresin, N., Kovacs, G., Phlairaharn, T., Furtwängler, B., et al. (2023) Exploration of cell state heterogeneity using single-cell proteomics through sensitivity-tailored data-independent acquisition. *Nat. Commun.* **14**, 5910
- [preprint] Petrosius, V., Aragon-Fernandez, P., Arrey, T. N., Üresin, N., Furtwängler, B., Stewart, H., et al. (2023) Evaluating the capabilities of the Astral mass analyzer for single-cell proteomics. *bioRxiv*. <https://doi.org/10.1101/2023.06.06.543943>
- Thielert, M., Itang, E. C., Ammar, C., Rosenberger, F. A., Bludau, I., Schweizer, L., et al. (2023) Robust dimethyl-based multiplex-DIA doubles single-cell proteome depth via a reference channel. *Mol. Syst. Biol.* **19**, e11503
- Derks, J., Leduc, A., Wallmann, G., Huffman, R. G., Willetts, M., Khan, S., et al. (2022) Increasing the throughput of sensitive proteomics by plexDIA. *Nat. Biotechnol.* **41**, 50–59
- Dwivedi, P., and Rose, C. M. (2022) Understanding the effect of carrier proteomes in single cell proteomic studies - key lessons. *Expert Rev. Proteom.* **19**, 5–15
- Cheung, T. K., Lee, C.-Y., Bayer, F. P., McCoy, A., Kuster, B., and Rose, C. M. (2020) Defining the carrier proteome limit for single-cell proteomics. *Nat. Methods* **18**, 76–83
- Budnik, B., Levy, E., Harmange, G., and Slavov, N. (2018) SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 161
- Specht, H., Emmott, E., Petelski, A. A., Huffman, R. G., Perlman, D. H., Serra, M., et al. (2021) Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biol.* **22**, 50
- Dou, M., Clair, G., Tsai, C.-F., Xu, K., Chrisler, W. B., Sontag, R. L., et al. (2019) High-throughput single cell proteomics enabled by multiplex isobaric labeling in a nanodroplet sample preparation platform. *Anal. Chem.* **91**, 13119–13127
- Schoof, E. M., Furtwängler, B., Üresin, N., Rapin, N., Savickas, S., Gentil, C., et al. (2021) Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat. Commun.* **12**, 3341
- Corteccka, C., Stejskal, K., Krssakova, G., Mendjan, S., and Mechtler, K. (2022) Quantitative accuracy and precision in multiplexed single-cell proteomics. *Anal. Chem.* **94**, 2434–2443
- Graumann, J., Scheltema, R. A., Zhang, Y., Cox, J., and Mann, M. (2012) A framework for intelligent data acquisition and real-time database searching for shotgun proteomics. *Mol. Cell Proteom.* **11**. <https://doi.org/10.1074/mcp.M111.013185>
- Bailey, D. J., Rose, C. M., McAlister, G. C., Brumbaugh, J., Yu, P., Wenger, C. D., et al. (2012) Instant spectral assignment for advanced decision tree-driven mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 8411–8416
- Erickson, B. K., Mintseris, J., Schweppe, D. K., Navarrete-Perea, J., Erickson, A. R., Nusinow, D. P., et al. (2019) Active instrument engagement combined with a real-time database search for improved performance of sample multiplexing workflows. *J. Proteome Res.* **18**, 1299–1306
- Schweppe, D. K., Eng, J. K., Yu, Q., Bailey, D., Rad, R., Navarrete-Perea, J., et al. (2020) Full-Featured, real-time database searching platform enables fast and accurate multiplexed quantitative proteomics. *J. Proteome Res.* **19**, 2026–2034
- Wichmann, C., Meier, F., Winter, S. V., Brunner, A.-D., Cox, J., and Mann, M. (2019) MaxQuant.Live enables global targeting of more than 25,000 peptides. *Mol. Cell Proteom.* **18**, 982–994
- [preprint] Budayeva, H. G., Ma, T. P., Wang, S., Choi, M., and Rose, C. M. (2023) Increasing the throughput and reproducibility of activity-based proteome profiling studies with hyperplexing and intelligent data acquisition. *bioRxiv*. <https://doi.org/10.1101/2023.09.13.557589>
- [preprint] Schweppe, D. K., Eng, J. K., Bailey, D., Rad, R., Yu, Q., Navarrete-Perea, J., et al. (2019) Full-featured, real-time database searching platform enables fast and accurate multiplexed quantitative proteomics. *bioRxiv*. <https://doi.org/10.1101/668533>
- Furtwängler, B., Üresin, N., Motamedchaboki, K., Huguet, R., Lopez-Ferrer, D., Zabrouskov, V., et al. (2022) Real-time search assisted acquisition on a tribrid mass spectrometer improves coverage in multiplexed single-cell proteomics. *Mol. Cell Proteom.* **21**, 100219
- Huffman, R. G., Leduc, A., Wichmann, C., di Gioia, M., Borriello, F., Specht, H., et al. (2023) Prioritized single-cell proteomics reveals molecular and functional polarization across primary macrophages. *Nat. Methods* **20**, 714–722
- Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., et al. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920
- Eng, J. K., Jahan, T. A., and Hoopmann, M. R. (2013) Comet: an open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24
- Consortium, T. U., Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Ahmad, S., et al. (2022) UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531
- The, M., MacCoss, M. J., Noble, W. S., and Käll, L. (2016) Fast and accurate protein false discovery rates on large-scale proteomics data sets with percolator 3.0. *J. Am. Soc. Mass Spectrom.* **27**, 1719–1727
- Savitski, M. M., Wilhelm, M., Hahne, H., Kuster, B., and Bantscheff, M. (2015) A scalable approach for protein false discovery rate estimation in large proteomic data sets[S]. *Mol. Cell Proteom.* **14**, 2394–2404
- R Core Team. (2021) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria
- Erickson, B. K., Rose, C. M., Braun, C. R., Erickson, A. R., Knott, J., McAlister, G. C., et al. (2017) A strategy to combine sample multiplexing with targeted proteomics assays for high-throughput protein signature characterization. *Mol. Cell* **65**, 361–370

31. Rose, C. M., Erickson, B. K., Schweppe, D. K., Viner, R., Choi, J., Rogers, J., *et al.* (2019) TomahqCompanion: a tool for the creation and analysis of isobaric label based multiplexed targeted assays. *J. Proteome Res.* **18**, 594–605
32. Woo, J., Williams, S. M., Markillie, L. M., Feng, S., Tsai, C.-F., Aguilera-Vazquez, V., *et al.* (2021) High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. *Nat. Commun.* **12**, 6246
33. Zhu, Y., Piehowski, P. D., Zhao, R., Chen, J., Shen, Y., Moore, R. J., *et al.* (2018) Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. *Nat. Commun.* **9**, 882
34. Choi, M., Carver, J., Chiva, C., Tzouros, M., Huang, T., Tsai, T.-H., *et al.* (2020) MassIVE.quant: a community resource of quantitative mass spectrometry-based proteomics datasets. *Nat. Methods* **17**, 981–984