Submitted to *Molecular & Cellular Proteomics*

**An improved Boosting to Amplify Signal with Isobaric Labeling (iBASIL) strategy for precise quantitative single-cell proteomics**

Chia-Feng Tsai\(^1\), Rui Zhao\(^2\), Sarah M. Williams\(^2\), Ronald J. Moore\(^1\), Kendall Schultz\(^1\), William B. Chrisler\(^1\), Ljiljana Pasa-Tolic\(^2\), Karin D. Rodland\(^1\), Richard D. Smith\(^1\), Tujin Shi\(^1\), Ying Zhu\(^2\)\(*\), and Tao Liu\(^1\)\(**

\(^1\)Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99354
\(^2\)Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354

**Running title:** iBASIL for precise quantitative single-cell proteomics

**Keywords:** boosting ratio, automatic gain control (AGC), iBASIL, ion injection time (IT), single-cell proteomics

\(*\)Corresponding authors:

Dr. Tao Liu  
Integrative Omics Group  
Biological Sciences Division  
Pacific Northwest National Laboratory  
Richland, WA 99354  
**Tel:** (509) 371-6346  
**Email:** tao.liu@pnnl.gov

Dr. Ying Zhu  
Environmental Molecular Sciences Laboratory  
Pacific Northwest National Laboratory  
Richland, WA 99354  
**Tel:** (509) 375-4523  
**Email:** ying.zhu@pnnl.gov
Abstract

Mass spectrometry (MS)-based proteomics has great potential for overcoming the limitations of antibody-based immunoassays for antibody-independent, comprehensive, and quantitative proteomic analysis of single cells. Indeed, recent advances in nanoscale sample preparation have enabled effective processing of single cells. In particular, the concept of using boosting/carrier channels in isobaric labeling to increase the sensitivity in MS detection has also been increasingly used for quantitative proteomic analysis of small-sized samples including single cells. However, the full potential of such boosting/carrier approaches has not been significantly explored, nor has the resulting quantitation quality been carefully evaluated. Herein, we have further evaluated and optimized our recent boosting to amplify signal with isobaric labeling (BASIL) approach, originally developed for quantifying phosphorylation in small number of cells, for highly effective analysis of proteins in single cells. This improved BASIL (iBASIL) approach enables reliable quantitative single-cell proteomics analysis with greater proteome coverage by carefully controlling the boosting-to-sample ratio (e.g., in general <100x) and optimizing MS automatic gain control (AGC) and ion injection time settings in MS/MS analysis (e.g., 5E5 and 300 ms, respectively, which is significantly higher than that used in typical bulk analysis). By coupling with a nanodroplet-based single cell preparation (nanoPOTS) platform, iBASIL enabled identification of ~2,500 proteins and precise quantification of ~1,500 proteins in the analysis of 104 FACS-isolated single cells, with the resulting protein profiles robustly clustering the cells from three different acute myeloid leukemia cell lines. This study highlights the importance of carefully evaluating and optimizing the boosting ratios and MS data acquisition conditions for achieving robust, comprehensive proteomic analysis of single cells.
Abbreviations

SP3-CTP: SP3-clinical tissue proteomics

DDM: n-Dodecyl β-D-Maltoside

NanoPOTS: nanodroplet processing in one pot for trace samples

TMT: tandem mass tag

SCoPE-MS: single cell proteomics by mass spectrometry

BASIL: boosting to amplify signal with isobaric labeling

iBASIL: improved boosting to amplify signal with isobaric labeling

S/N: signal-to-noise

ITs: ion injection times

AML: acute myeloid leukemia

CVs: coefficient of variations

AGC: automatic gain control

FACS: fluorescence activated cell sorting

B/S ratios: boosting-to-sample ratios
Introduction

Cell and tissue heterogeneity is an important fundamental issue in many research areas (e.g., developmental and cancer biology\(^1\)), but the resulting variation is lost in conventional “bulk” omics analysis. While recent advances in DNA and RNA sequencing technologies are enabling routine single-cell genomics and transcriptomics analysis\(^2\)-\(^3\), the ability to measure protein expression in single cells still lags far behind in, e.g., proteome coverage and sample throughput. Antibody-based immunoassays\(^4\)-\(^5\) have been used for single-cell proteomics analysis but they have inherent limitations (e.g., low multiplexing ability and lack of high-quality antibodies). Mass spectrometry (MS)-based proteomics has great potential to overcome these limitations for antibody-free, comprehensive, and quantitative proteomic analysis of single cells. However, such potential has not been fully explored primarily due to inefficient sample processing of single cells, as well as limited MS sensitivity.

To tackle this issue, a first step is to significantly reduce sample loss during sample processing, such as cell lysis and protein digestion. Recent significant advances in sample preparation are enabling effective processing of smaller samples with the potential of moving towards single cells. Hughes et al. introduced a paramagnetic bead-based protocol, termed Single-Pot Solid-Phase-enhanced Sample Preparation (SP3), for rapid and unbiased sample preparation in a single tube\(^6\). The SP3 protocol was further optimized as a SP3-Clinical Tissue Proteomics (SP3-CTP) platform for in-depth proteome profiling of small clinical tumor specimens\(^7\). Myers et al. developed a microreactor-tip-in-a-Stage-tip device for performing all sample processing steps in single microreactor for proteomic analysis using low protein input (~2 μg)\(^8\). Our group recently introduced a carrier-assisted single-tube processing approach for ultrasensitive targeted proteomics analysis of small numbers of cells\(^9\). This approach was
demonstrated to enable targeted quantification of the majority of epidermal growth factor receptor pathway proteins in 10-100 mammalian cells. We have also demonstrated that the addition of a MS-compatible detergent, n-Dodecyl β-D-Maltoside (DDM), can significantly reduce surface adsorption for improving sample recovery. Most importantly, we have recently developed a nanoPOTS (nanodroplet Processing in One Pot for Trace Samples) platform to dramatically improve sample processing efficiency for small number of cells down to single cells. The nanoPOTS not only efficiently reduces adsorptive protein/peptide loss because of the use of nanowells, but also significantly enhanced tryptic digestion kinetics due to increased protein and trypsin concentrations in nanoliter volumes. NanoPOTS integration with a state-of-the-art MS platform has provided reliable identification of ~670 and ~3,000 protein groups from single cells and 10-14 cells, respectively.

Another strategy to enhance MS detection sensitivity is the use of isobaric tags such as the tandem mass tag (TMT) for sample multiplexing, especially when one or several TMT channels are labeled with a large amount of relevant “boosting” (or “carrier”) so as to enhance protein detection and minimize sample surface losses of the much smaller amounts of labeled samples labeled in the other channels. This design significantly enhances the detectability of the MS1 signal for triggering MS/MS sequencing; the reporter ion intensities from study sample channels are then used for reliable quantification of each individual sample. Using this concept, Russell et al. developed TMTcalibrator™, in which cell lines or tissue-derived references were used as TMT boosting channels for sensitive detection of low abundance proteins in body fluids (e.g., cerebrospinal fluid and plasma) and Budnik et al. developed a SCoPE-MS (Single Cell ProtEomics by Mass Spectrometry) approach for quantitative single-cell analysis. We have recently developed a BASIL (Boosting to Amplify Signal with Isobaric Labeling) strategy for
enabling comprehensive phosphoproteomic analysis of smaller samples\textsuperscript{16} (e.g., quantification of >20,000 phosphosites from human pancreatic islet). More recently, we have also incorporated isobaric TMT labeling into our nanoPOTS workflow for enabling reliable clustering of 61 single cells from three different cell lines\textsuperscript{17}.

All the above TMT-boosting approaches have demonstrated the potential of using isobaric TMT labeling for high-throughput, sensitive, and quantitative nanoscale and single-cell proteomics analysis. The isobaric labeling-based boosting strategy, coupling to highly effective sample processing (e.g., nanoPOTS), provides attractive promise for quantitative single-cell proteomics. However, many technical details such as the quantitation quality and optimal experimental conditions, have not been thoroughly investigated. For example, the boosting-to-sample ratio varied from 30\textsuperscript{16} to 200\textsuperscript{15} or even 500\textsuperscript{18} in different studies. In this study, we systematically evaluated and optimized the BASIL conditions for reliable identification and quantitation performance with nanoPOTS. We found that excessively high boosting ratios degrade both signal stabilities and signal-to-noise ratios (S/N) of lower abundance proteins, due mainly to the limited charge capacity of the Orbitrap; significantly increased automatic gain control (AGC) and ion injection time settings that those used in typical bulk analysis, on the other hand, help improve the signal in the sample channels, resulting in precise protein quantification in the single cells while achieving improved proteome coverage. This improved BASIL (iBASIL) strategy was demonstrated by the reliable precise quantification of \(~\text{1,500}~\) proteins in 104 FACS-isolated single cells, which led to robust separation and clustering of these cells from 3 different acute myeloid leukemia (AML) cell lines.
Experimental Procedures

Reagents

Urea, dithiothreitol (DTT), iodoacetamide (IAA), triethylammonium bicarbonate, trifluoroacetic acid (TFA), ethylenediaminetetraacetic acid (EDTA), ammonia phosphate \( (\text{NH}_4\text{H}_2\text{PO}_4) \), trifluoroacetic acid (TFA) and formic acid (FA) were obtained from Sigma (St. Louis, MO). Empore\textsuperscript{TM} extraction disk C18 were from 3M (St. Paul, MN). TMT-10 reagents were purchased from Thermo Fisher Scientific (Waltham, MA). Water was processed using a Millipore Milli-Q system (Bedford, MA).

Cell Culture and bulk-scale protein digestion

The MCF-7 and MCF10A breast cell line were obtained from the American Type Culture Collection and was prepared as previously described\textsuperscript{19}.

For the AML cells, MOLM-14 and K562 cells were maintained in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS), and CMK cells were maintained in RPMI-1640 medium supplemented with 20% FBS.

Cells were washed with ice-cold phosphate-buffered saline (PBS), lysed in a lysis buffer containing 50 mM \text{NH}_4\text{HCO}_3, pH 8.0, 8 M urea, and a 1% phosphatase inhibitor, and sonicated in an ice bath for 3 min. The protein concentration was determined using the BCA protein assay (Thermo Fisher Scientific). The proteins solutions were denatured with 5 mM DTT for 1 h at 37 °C and alkylated with 20 mM iodoacetamide in the dark for 1 h at room temperature. The resulting samples were diluted 8-fold with 50 mM \text{NH}_4\text{HCO}_3 and digested with lysyl
endopeptidase at 37 °C (protein:enzyme, 50:1, w/w) for 3 h followed by trypsin (protein: enzyme, 50:1, w/w) at 37 °C overnight. The digested tryptic peptides were acidified by TFA with a final TFA concentration of 0.5%, and then desalted by C18 SPE extraction and concentrated for BCA assay analysis.

TMT labeling of bulk samples

The tryptic peptides from bulk samples were dissolved with 200 mM HEPES (pH 8.5) and then mixed with a TMT-10 reagent in 100% ACN. An optimized ratio of TMT to peptide amount of 1:1 (w/w) recently reported by Jana et al.20 was used (i.e., 100 μg of peptides labeled by 100 μg of TMT reagent). After incubation for 1 h at room temperature, the reaction was terminated by adding 5% hydroxylamine for 15 min. The TMT-labeled peptides were then acidified with 0.5% FA and then diluted to 4% ACN, before desalting using C18 SPE tips. The BCA assay was used to estimate the peptides amount after TMT labeling. The different amounts of peptides (0.1 ng, 0.5 ng, or 10 ng) were directly diluted from the bulk TMT-labeled peptides.

Single cell isolation by FACS

The Influx II cell sorter (BD Biosciences, San Jose, CA) was used to isolate single cells directly into the nanowells11. The procedures for microchip fabrication and single cell isolation were described previously17.

Proteomic sample preparation and TMT labeling in nanoPOTS
The nanoPOTS-based proteomic sample preparation and TMT labeling procedures were described previously\textsuperscript{17}, with slight modification. In brief, isolated single cells were lysed in 100 nL lysis buffer (0.1% (m/v) DDM in 50 mM TEAB and 0.5 × PBS buffer) and incubated at 70 °C for 60 mins to achieve efficient protein extraction. The extracted proteins were directly digested by adding 50 nL, 20 ng/μL trypsin solution (1 ng) and then incubated at 37 °C for 8 h. The tryptic peptides were incubated with 100 nL, 10 μg/μL TMT reagent (1 μg) for 1 h at room temperature. Next, 50 nL, 1% (v/v) hydroxylamine was used for quenching unreacted TMT at room temperature for 15 min, and 50 nL, 5% FA was used to acidify the sample. Finally, the TMT-labeled peptides of single cells from multiple smaller nanowells were collected and combined into the larger “boosting” well which contained 10 ng TMT-labeled peptides prepared in bulk scale. The samples in nanoPOTS chip were dried out in a vacuum desiccator and stored in -80°C freezer.

**LC-MS/MS Analysis for evaluating the boosting ratios, ITs, and AGC settings**

Lyophilized peptides were reconstituted in 0.1% FA with 2% ACN containing 0.01% DDM and injected using a PAL autosampler (CTC Analytics AG, Switzerland). The sample was concentrated into an online SPE column (150 μm i.d., 360 μm o.d., 4 cm long) and separated using a 50 μm i.d., 360 μm o.d., 50 cm long column packed with 3-μm C18 particles (300-Å pore size; Phenomenex, Terrence, CA). The nanoLC separation used a Dionex UlitMate NCP-3200RS system (Thermo Scientific, Waltham, MA) with mobile phases of water with 0.1% FA (buffer A) and ACN with 0.1% FA (buffer B). Peptides were separated through a linear gradient from 8% to 35% buffer B over 100 min at a flow rate of 150 nL/min. The separated peptides were analyzed using a Thermo Scientific Q Exactive Plus for method optimization with iBASIL.
Data were acquired in a data-dependent mode with MS scans from m/z 300-1800 at a resolution of 35,000 at m/z 400. Top 10 precursor ions were selected for MS/MS sequencing at a higher-energy collision dissociation (HCD) energy of 35%. The MS/MS scan resolution was set at 70,000. Different AGC settings and maximum ITs at MS/MS level were tested for optimization (see Table S1).

**LC-MS/MS analysis comparing MS2 and MS3 methods for iBASIL analysis**

Lyophilized, previously TMT-labeled and mixed tryptic peptides (0.1 ng sample per channel and 100 ng boosting sample) from the three AML cell lines (MOLM14, K562 and CMK) were reconstituted in 12 μL of 0.1% FA with 2% ACN and 5 μL of the resulting sample was analyzed by LC-MS/MS using an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Scientific) connected to a nanoACQUITY UPLC system (Waters Corp., Milford, MA). Peptides were separated on analytical column (75 μm i.d. × 20 cm) packed using 1.9-μm ReproSil C18 and with a column heater set at 50 °C. The mobile phases consisted of 0.1% FA with 3% ACN (buffer A) and 0.1% FA in 90% ACN (buffer B). The gradient setting for the peptide separation was: 2-6% buffer B in 1 min, 6-30% buffer B in 84 min, 30-60% buffer B in 9 min, 60-90% buffer B in 1 min, and finally 90% buffer B for 5 min. The flow rate was set at 200 nL/min.

Full-scan spectra were acquired at a target value of $4 \times 10^5$ with a resolution of 60,000 with an AGC setting of $4 \times 10^5$ and maximum ion injection time of 50 ms. Data were acquired in data-dependent acquisition mode. For the synchronous precursor selection (SPS)-MS3 method, the peptide precursors were first isolated by quadrupole with an isolation window of 0.7 Da and interrogated by MS2 in the ion trap using collision-induced dissociation (CID) at a collision
energy of 35. The MS/MS spectra were recorded at a target value of $1 \times 10^4$ with 50 ms max ion injection time. The MS3 analysis was performed for each MS2 scan with multiple MS2 fragment ions using multinotch isolation waveform and a normalized collision energy of 65% for HCD. The fragment ions were detected by Orbitrap (60,000 resolution) with an AGC target of $5 \times 10^5$ and the maximum ion injection time of 300 ms. For the MS2 method, The MS/MS isolation window was also set at 0.7 Da, and HCD fragmentation was performed at a normalized collision energy of 35% with an AGC of $5 \times 10^5$ and a maximum ion injection time of 300 ms. The MS/MS spectra were acquired at a resolution of 50,000.

LC-MS/MS for nanoPOTS-based single-cell analysis

The LC-MS/MS method used for single-cell analysis were described previously\textsuperscript{17} with slightly modification. The SPE precolumn (100 μm i.d., 360 μm o.d., 4 cm long) was packed with 3-μm C18 packing material (300-Å pore size, Phenomenex, Torrence, CA) and LC column (50 μm i.d., 360 μm o.d., 30 cm long) were packed with 1.7 μm C18 particles (Bridged Ethylene-Hybrid C18, Waters, Milford, MA). The LC column was heated at 50 °C during separation. The flow rate of the LC separation was 100 nL/min using a nanoUPLC pump (Dionex UltiMate NCP-3200RS, Thermo Scientific, Waltham, MA). A linear 100 min gradient of 8–30% buffer B (0.1% formic acid in acetonitrile) was used for the LC separation. Then the LC column was washed by ramping buffer B to 45% in 20 min and then keeping at 90% in 5 min, and finally re-equilibrated with 2% buffer B for another 10 min. A Thermo Scientific Q Exactive Plus was used for the MCF10A cells and an Orbitrap Fusion Lumos Tribrid mass spectrometer was employed for AML cells. The parameters for AGC and ion injection times were listed in the Table S1.
Data Analysis

The MaxQuant (version 1.6.2.10)\textsuperscript{21-22} was used to process the raw files and MS/MS spectra were searched against a human UniProt database (fasta file dated April 12, 2017 with 20,198 sequences). The search type was set to “Reporter ion MS2” or “Reporter ion MS3” for isobaric label measurements. The mass tolerance for precursor ions and fragment ions are using default value in MaxQuant. A peptide search was performed with full tryptic digestion (Trypsin/P) and allowed a maximum of two missed cleavages. Carbamidomethyl (C) was set as a fixed modification; acetylation (protein N-term) and oxidation (M) were set as variable modifications. Proteins were considered as identified when the false discovery rate (FDR) at both peptide and protein levels was lower than 1%; no additional filtering was performed. The intensities of all ten TMT reporter ions were extracted from MaxQuant outputs and analyzed by Perseus\textsuperscript{23} for statistical analyses. The number of quantifiable peptides and proteins was determined by removing identified peptides and proteins with missing values.

For analysis of the data from the 104 AML single cells, the relative abundances from 13 TMT plexes were log2 transformed and normalized to the reference channel. The data matrices were then combined after separate row-and-column-centering by median values, and further normalized via width adjustment. The normalized TMT signals were further analyzed by Perseus for statistical analysis.

Results and Discussion

Systematic optimization of iBASIL for comprehensive and precise quantitative analysis
The orbitrap mass analyzer provides both high-resolution and high-sensitivity measurements for bottom-up proteomics, but the total number of trapped ions is limited by the space charge effect and is in the range of 1 million elementary charges\textsuperscript{24}. Due to this limitation, ions from the single-cell samples cannot be effectively detected when high boosting-to-sample ratios (B/S ratios) are used. To evaluate the effects of the B/S ratios on protein identification and quantitation, a fixed amount of tryptic peptides from MCF7 cells labeled with TMT126, TMT127N and TMT127C were mixed with varying amounts of TMT131N-labeled boosting peptides to generate the B/S ratios of 10×, 50×, 100× and 200× (Figure 1a). As expected, the number of identified peptides increased (from 5,055 to 6,714) as the B/S ratios increased (Figure 1b). However, unlike the increased TMT signals in the boosting channel (Figures S1a), the TMT signals in the sample channels decreased as the B/S ratios increase (Figure 1c), resulting in higher standard deviations (S.D.) (Figure S2), decreased Pearson correlation coefficients (from 0.991 to 0.967; Figure 1d), and increased median coefficient of variations (CVs) (from 7.7% to 11.9%; Figure 1e) for the TMT signals from sample channels. Note that the TMT intensities of the empty channels were not affected as the B/S ratios increase (Figure S3) (TMT130N intensity was not plotted because it was affected by the isotopic impurities of TMT131N). A similar trend was also observed for 10 ng samples of peptides with varying B/S ratios (Figure S4 and S1b). These results demonstrated that high B/S ratios provide increased peptide/protein identifications at the expense of quantitation quality.

To understand this phenomenon, we extracted the MS2 ITs at the different B/S ratios. As shown in Figures 1f and S4f, the MS2 injection times (ITs) were found to be significantly affected by the B/S ratios. When the B/S ratios increased from 10× to 200×, the median MS2 ITs
decreased from 300 ms to 168 ms, suggesting that ions from the boosting channel at the high B/S ratios filled most of ion space in Orbitrap quickly. Therefore, systematic optimization of MS data acquisition parameters is necessary for optimized detection and quantification of single-cell samples prepared using the BASIL strategy.

(b) The effect of automatic gain control setting

During MS/MS analysis on Orbitrap mass analyzer, the total number of precursor ions is controlled by AGC and the maximum ITs. For typical “bulk” global proteomics analysis, the AGC value is set at a range from 5E4 to 1E5 to increase MS2 scan rate and improve the overall proteome coverage. However, for size-limited study samples prepared using BASIL with a much larger amount of boosting sample present, such AGC settings may not be sufficiently high to allow sufficient ions from the study samples to be accumulated for optimal identification and quantitation performance.

To test this, tryptic peptides (0.5 ng) from MCF7 cell digests labeled individually with TMT126, TMT127N and TMT127C were mixed with TMT131-labeled boosting peptides at a B/S ratio of 100 and analyzed by LC-MS/MS using three different AGC settings, 5E4 (regular AGC setting for bulk global proteomics analysis), 5E5 and 5E6, while the maximum IT was fixed at 300 ms. As expected, higher AGC allowed for accumulation of more ions from both study and boosting channels (Figure 2a) providing significantly increased TMT signals from the study channels (Figure 2b), which can be attributed to the dramatically increased ITs. As shown in Figure 2c, the median ITs increased from 32 ms to 300 ms when AGC increased from 5E4 to 5E6. The enhanced TMT signal in the study sample channels led to significantly improved
quantitation quality. For example, the Pearson correlation coefficients increased from 0.933 to 0.994 (Figure 2d), the median CVs decreased from 14.6% to 8.8% (Figure 2e) and the missing value decreased from 2% to 0.2% (Figure 2f). However, the number of quantifiable peptides decreased from 7,208 to 5,993 when the AGC increased from 5E4 to 5E6 (Figure 2f), due to the significantly increased duty cycle times which resulted in reduced MS2 scan rate. A similar trend was also observed for 0.5 ng input samples at a B/S ratio of 1000 with varying AGC settings (Figure S5). Because of the higher B/S ratio, the insufficient AGC (5E4) resulted in reduced TMT intensities of study channels in comparison with higher AGC (Figure S5a). The higher boosting ratio (1,000x) also resulted worse Pearson correlation (0.61, Figure S5c), CVs distribution (37.4%, Figure S5d) and higher missing value (73%, Figure S5e) for quantitation when the lower AGC used (5E4). However, the higher AGC (5E6) can greatly improve the quantitation performance indicated by the significantly improved Pearson correlation (0.969, Figure S5c), CVs distribution (9%, Figure S5d) and lower missing value (1%, Figure S5e) due the increasing of TMT signals (Figure S5a). These results also demonstrated the higher AGC was the most essential MS setting when high boosting ratios were used.

(c) The effect of maximum ion injection time setting

Besides AGC, the maximum IT is another key parameter to control the total number of peptide ions for MS2 analysis. To similarly test the IT effect tryptic peptides labeled individually with TMT126, 127N and TMT127C were mixed with the TMT131N-labeled boosting peptides at the B/S ratios of 10×, 50×, 100×, and 200×, and analyzed with different ITs ranging from 300 ms to 1,000 ms while AGC was fixed at 5E5. Consistent with a previous report25 that longer ITs
can increase MS detection sensitivity for low abundance proteins, the median TMT signals from
the sample channels and the Pearson correlation coefficients were improved using longer ITs for
all the B/S ratios (Figures S6a and S6b, respectively). Longer ITs also led to increased number of
quantifiable peptides at the lowest B/S ratio (10×). However, at larger B/S ratios (e.g., ≥100), the
numbers of quantifiable peptides rapidly decreased due to higher duty cycle times (i.e., fewer
measurements; Figure S6b).

The results from the evaluation of B/S ratios and AGC and maximum IT settings clearly
show a trade-off between achievable proteome coverage and quality of quantification. Our
measurements suggest a B/S ratio of 100 and relatively high AGC (5E5) and IT (300 ms) settings
are a reasonable starting point for implementation of iBASIL for single cell analysis, with the
latter two being specific to the Q Exactive Plus MS instrument.

Feasibility of single-cell proteomics analysis using iBASIL and large boosting ratios

To additionally evaluate the combined effects of the iBASIL settings for quantitative
single-cell proteomics analysis using boosting (as compared to the typical bulk analysis MS
settings), single cell equivalents (0.1 ng of tryptic peptides) from three AML cell lines (MOLM-
14, K562 and CMK) were labeled individually with TMT126-130C and then combined with the
TMT131N-labeled boosting sample (as in Figure 4a) at a B/S ratio of 1,000. The B/S ratio of
1,000 was used to represent the worst-case scenario for boosting (larger than those reported in
the literature). TMT130N channel was left empty to avoid skewed measurements caused by
isotopic impurity of the TMT131N reagent (i.e., the boosting channel). The samples were
analyzed by both the optimized iBASIL MS settings (AGC=5E6; max IT=300 ms) and normal
MS settings typically used for conventional bulk analysis (AGC=5E4; max IT=100ms). When compared to the results obtained using the normal MS setting, iBASIL significantly increased the number of quantifiable peptides from 900 to 3,541 and proteins from 585 to 1,131 (Figure 3a). More importantly, the median values of TMT signals in the study sample channels increased by >4 fold when the improved iBASIL settings were implemented (Figure 3b), resulting in better quantification as evidenced by the improved separation and clustering of the AML cells in the PCA analysis (Figure 3c).

We then evaluated the potential impact of a large B/S ratio (1,000) on quantification, even with the iBASIL settings, using the same three AML cell lines at different B/S ratios (no boosting, 100, and 1,000; Figure 4a). As expected, the number of quantifiable peptides and proteins increased with the B/S ratio (Figure 4b). For all the B/S ratios, the three cell types were reasonably well clustered and separated in the PCA analysis; however, the quality of the separation and reproducibility did slightly decline as the B/S ratio increased (Figure 4c). Interestingly, only a moderate increase (14.4%) in the quantifiable proteins was observed when the B/S ratio was increased from 100 to 1,000, suggesting that extremely high B/S ratios may not be needed for achieving deep proteome coverage (Figure 4b). Furthermore, at the large B/S ratio of 1,000, when TMT131N was used as the boosting channel, in addition to TMT130N, two additional TMT sample channels (TMT129N and TMT130C) were also significantly affected, most likely due to isotopic impurities of TMT131N (Figure S7). Therefore, when higher B/S ratios are used, the experimental design should be modified to leave the affected channels empty (we also note these observations suggest the utility of future modifications to the mass tagging reagents to reduce/eliminate the reduced multiplexing or distorted quantification resulting from reagent isotopic impurity).
To further evaluate the quantification at different B/S ratios, we also calculated the fold-change between any two cell types and performed pairwise correlation for the data with and without boosting. As shown in Figure 4d, the fold changes in the data with 100× boosting and without boosting exhibited good correlation with Pearson correlation coefficients of 0.74, 0.83 and 0.87 for the three pairwise comparisons, respectively (Figure 4d), suggesting an overall good quantification with 100x boosting. The slopes of linear regression for the same pairwise comparisons, however, were 0.59, 0.79 and 0.84, respectively, which indicates that the dynamic range in quantitation was slightly reduced after applying the boosting strategy; this was further confirmed by the poorer correlation and lower slopes when comparing the data with 1,000x boosting and without boosting (Figure 4d). These results showed that although the boosting strategy can significantly increase the proteome coverage for single-cell analysis, the higher boosting ratio would reduce the dynamic range in quantitation due to the limited dynamic range of MS detectors, e.g. the linear dynamic range of Orbitrap (intra-scan) range26 from 1,000 to 1,0000.

Collectively, these data showed that the B/S ratios should be carefully selected for quantitative analysis of small-sized samples with matched AGC and IT settings for well-balanced results in both proteome coverage and quantification quality.

Comparison of iBASIL-MS2 to iBASIL-MS3 methods

In TMT-based quantitative global proteomic analysis, the greater selectivity provided by SPS-MS3 has been shown to mitigate the isobaric labeling ratio “compression” issue associated with MS2 approaches27. We therefore evaluated if SPS-MS3 could further improve quantitation
quantification using iBASIL. The same single-cell equivalents from three AML cell lines with the B/S ratio of 100 were analyzed by SPS-MS3 on an Orbitrap Lumos Tribrid MS. Unlike the Q Exactive Plus instrument used for the above iBASIL optimization experiments that has an effective maximum AGC of 5E6, the maximum effective AGC setting on the Lumos instrument seems to be 1E6 (Figure S8). Therefore, the AGC value was set at 5E5 for the comparison of MS2 and SPS-MS3 analyses on the Lumos instrument. As shown in Figure S9a, a slight decrease of TMT reporter ion intensities for SPS-MS3 was observed, however the three AML cell lines were readily separated in both analyses (Figure S9b). While SPS-MS3 indeed provided a greater quantitation dynamic range than MS2 (Figure S9c), the number of quantifiable protein groups with SPS-MS3 was significantly lower than that with MS2 (Figure S9d) due to the longer duty cycle times (i.e., lower number of measurements) in SPS-MS3 (Figure S9e). This situation, however, may be significantly improved by using the newest Real Time Search-MS3 method (RTS-MS3) that became available very recently.

**Single-cell proteomics analysis using nanoPOTS and iBASIL**

Finally, the optimized iBASIL strategy was applied for precise quantitative proteomic analysis of FACS-sorted single MCF10A cells. Single cells were sorted and processed with our recently developed nanoPOTS platform (Figure S10a). Ten nanograms of tryptic peptides from a bulk MCF10A cell digest was used as the boosting channel (Figure S10b). When compared to single-cell proteomics results without boosting, ~2-fold more proteins (from 434 to 853; 390 and 664 have two or more peptides hits, respectively) were quantified by iBASIL (Figure S10c). Although a 50 μm-i.d. LC column and an older generation Q Exactive MS were employed in this
study, the result is comparable to our previous single-cell proteomics work\textsuperscript{17}, where a 30 μm-i.d. LC column and an Orbitrap Lumos MS were used. The quantification performance of the iBASIL MS settings (AGC: 5E6; IT: 300 ms) and other reported MS settings for single-cell proteomics analysis (AGC: 5E4; IT: 300 ms)\textsuperscript{29} was also compared. Although the number of quantifiable proteins was slightly lower (Figure S10d), the single-cell signals using the iBASIL settings were significantly increased by ~1.8-fold (Figure S10e), resulting in a large improvement in the quantification (Figure S10f). These results further demonstrate the importance of selecting appropriate MS parameters for comprehensive and quantitative single-cell proteomics analysis.

We then initially demonstrated the feasibility of the iBASIL strategy for large-scale quantitative single-cell proteomic analysis using 104 FACS-sorted single AML cells (39 MOLM-14, 39 K561 and 26 CMK cells) labeled in 13 TMT plexes on nanoPOTS platform (Figure 5a)\textsuperscript{11}. The 10-ng and 0.2-ng mixed peptides prepared from the 3 cell lines were labeled with TMT126 and TMT127N and used as the boosting and reference samples, respectively (Figure 5b). The iBASIL analysis of these AML single cells identified an average of 11,572 peptides from 1,926 proteins in each TMT plex (Figure 5c). Among all the identified proteins (2,622), 2,009 proteins have two or more peptide identifications; 1,534 proteins (59%) were quantified in more than 70% of the single-cell TMT channels (Figure 5d), among which 1,462 (95%) have two or more peptide identifications. The single cells from three different AML cell lines clearly clustered by respective cell types which were well separated from one another (Figure 5e). Furthermore, even at the coverage available with as few as ~100 AML cells, functionally distinct differences were evident in the three AML cell lines that correlate well with their known driving mutations. For example, the CMK cell line, which harbors a JAK mutation,
shows increased protein abundance associated with JAK/STAT signaling and also BRAF-associated signaling (Cluster 3 in Figure 5f; Figure S11), which suggests the RAF inhibitors might be a useful adjunct to JAK inhibitors in some cases. The MOLM-14 cell line, which harbors the microenvironment-responsive FLT3-ITD mutation, shows an increase in proteins associated with immune function (Cluster 1 in Figure 5f; Figure S11), while the K562 cell line, which harbors the prototypical BCR-ABL fusion, is characterized by increase in proteins associated with proliferation and metabolism (Cluster 2 in Figure 5f; Figure S11). Together, these results showed the potential of the iBASIL strategy for deep, robust quantitative single-cell proteomics analysis.

Potential limitations and perspective of boosting/carrier strategy for single-cell proteomics analysis

Our results presented herein demonstrated that the iBASIL strategy allowed for precise quantification of over 1,500 proteins that reveal functionally distinct differences in the single cells, representing a significant step forward towards comprehensive, high-fidelity single-cell proteomics analysis. One fundamental limitation of boosting-like approaches\(^{13,15-17}\) is that the signal increase is only at MS1 level, but not at MS2 level (of the reporter ions); this is partially addressed by iBASIL by using much higher AGC settings, however it will eventually hit a ceiling where duty cycle becomes impractically long, thus drastically reducing the achievable proteome coverage. Another limitation of the boosting strategy lies in the fact that although single-cell proteome coverage generally increases with increased amount of the boosting sample, the proportion of signal from the sample channels would decrease, and the quantitation dynamic
range would be compressed (Figure 4d). Therefore, there is a fine balance between proteome coverage and quantitation quality, and as a result, at present the proteome coverage of single-cell analysis is still lag far behind compared to that of the bulk analysis. Other issues associated with the isobaric labeling-based analysis (e.g., the boosting strategy) include potential contaminating signals from other co-isolated and co-fragmented peptides or even potentially metabolites in MS2 spectrum that skew the quantitation accuracy. The innovations on MS instrumentation, labeling methodologies and sample preparation strategies are expected to overcome some of the abovementioned limitations, and further improve detection sensitivity in single-cell proteome analysis. For example, the combination of iBASIL with nanoscale fractionation\(^{10}\) can further improve the proteome coverage while reducing the precursor co-isolation; the newly available Real Time Search-MS3 method (RTS-MS3)\(^{28}\) or ion mobility technique\(^{30,31}\) maybe a solution for precise and accurate quantitation without sacrificing the proteome coverage.

**Conclusion**

Taking advantage of the multiplexed detection enabled by isobaric labeling, the boosting/carrier strategies are increasingly being used in nanoproteomics or single-cell proteomics analysis for enhanced sensitivity. However, quite different boosting settings have been used in the previous reports, and the quantitation quality has not been carefully evaluated. In this study, we have systematically evaluated and optimized the boosting settings on the most common Orbitrap MS platforms and proposed an iBASIL method that can be easily implemented for broad applications. Using BASIL and “BASIL-like” approaches, there is always a balance between the achievable proteome coverage and quantification quality for each
individual parameter. We suggest the B/S ratio of 100 be used, in conjunction with high AGC and IT values as starting points for iBASIL implementation and optimization. When extremely high B/S ratios (e.g., ≥1,000) are used to increase the proteome coverage, higher AGC and/or IT setting are required to achieve reliable quantitation, but with the trade-off of a decreased number of quantifiable proteins. In this work we demonstrated that the optimized iBASIL strategy enabled precise quantification of ~1,500 proteins in 104 FACS-isolated single AML cells by coupling with nanoPOTS-based sample preparation platform. These results allowed for robust clustering and separation of single cells from 3 different AML cell lines, and revealed functionally distinct differences in their proteomes. In conclusion, we believe that iBASIL will have broad utility in systems biology and biomedical research for precise quantitative single-cell proteomics and nanophosphoproteomics, as well as for analysis of precious mass-limited clinical specimens not readily accessed by current proteomics platforms.

Data Availability

Supplementary figures and Supplementary tables are available. The experimental conditions and raw file index are descripted in Table S1. The raw datasets and identified peptides and proteins lists (output files including Evidence.txt, Peptides.txt and ProteinGroups.txt from Maxquant) have been deposited in Japan ProteOme STandard Repository32 (jPOST; https://repository.jpostdb.org/). The accession number is PXD016057 and PXD017626 for ProteomeXchange (http://www.proteomexchange.org/) and JPST000690 and JPST000723 for jPOST. The preview link is https://repository.jpostdb.org/preview/16206540165db720294d19b (the access key is 1810) and https://repository.jpostdb.org/preview/3618873295e4dc5546abd1 (The access key is 3649)
The authors declare no competing financial interest.

Acknowledgements

The authors thank Anupriya Agarwal and Karina Thiel-Klare at the Oregon Health & Science University for providing the cell stocks for the AML cells, and Matthew Geffrey at the Pacific Northwest National Laboratory (PNNL) for expanding and culturing the AML cells. Portions of the research were supported by a U24CA210955 grant (to TL and RDS) and U01CA214116 grant (to KDR) from the NCI Clinical Proteomic Tumor Analysis Consortium (CPTAC), NCI Early Detection Research Network (EDRN) Interagency Agreement ACN15006-001 (to TL and KDR), a P41 GM103493 grant from the National Institute of General Medical Sciences (to RDS), a Laboratory Directed Research and Development award from PNNL (to YZ), and a R21CA223715 grant from NCI Innovative Molecular Analysis Technologies (IMAT) (to TS).

The research was performed using EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research.
Figure Legend

Figure 1. The effects of boosting ratio on BASIL analysis. The TMT channels for the study samples and boosting sample are shown (a). The number of quantifiable peptides (b) and TMT reporter ion intensities for the sample channels (c) are shown at 4 different TMT boosting ratios (10x, 50x, 100x, and 200x). The Pearson correlation coefficients of peptides (d), CVs of peptides (e), and the distribution of actual IT times (f) are also shown for the samples prepared with these 4 different boosting ratios. The Pearson correlation coefficients were calculated among three TMT channels (TMT126, TMT127N and TMT127C). The quantifiable peptides are those that have TMT signals detected in all the sample channels.

Figure 2. The effects of AGC on BASIL analysis. Higher AGC settings allow for accumulation of more ions from the study sample channels with the presence of the boosting sample (a). The TMT reporter ion intensities of the sample channels (b), actual ion injection times (c), Pearson correlation coefficients (d), CV (e), and number of quantifiable peptides (those that have TMT signals detected in all the sample channels) (f) are shown at three different AGC settings: 5E4, 5E5 and 5E6.

Figure 3. Evaluation of iBASIL performance. The numbers of quantified peptides and proteins (those that have TMT signals detected without missing value at least in one cell type) (a) and TMT reporter ion intensity for the sample channels (b) for the same samples analyzed with regular BASIL and iBASIL are shown. PCA analysis using the quantifiable proteins showed the separation of three AML cell lines using the normal BASIL (left) and iBASIL (right) (c).
Figure 4. The effects of large boosting ratio on iBASIL analysis. The TMT channels for the study samples (K562, MOLM-14 and CMK cells) and boosting sample are shown (a). The numbers of quantified peptides and proteins (b) and the respective PCA results (c) are shown for the quantitative single-cell proteomics analysis using three different boosting ratios (no boosting, 100x, and 1,000x). Both the correlation and slope of fold changes of the quantified proteins between the different AML cell lines with and without using the boosting sample decrease as the boosting ratio goes from 100x (d; top panels) to 1,000x (d; bottom panels).

Figure 5. Quantitative analysis of 104 FACS-sorted AML single cells by coupling nanoPOTS with iBASIL. A schematic and the TMT experiment design for the nanoPOTS-iBASIL analysis of FACS-isolated AML single cells are given in (a) and (b), respectively. The numbers of identified peptides and protein in each TMT experiment are shown in (c). The numbers of quantifiable proteins using different filtering thresholds for valid values are shown in (d). PCA analysis shows the clustering of single cells from the same cell lines and the separation of cells from different cell lines (e). Heatmap of significantly changed proteins shows clear differences in the proteome profiles for the single cells from the 3 different cell lines (f); Clusters 1-3 highlighted enriched cellular functions in the different AML cells (see Figure S11 for additional information).
REFERENCES

1. Heath, J. R.; Ribas, A.; Mischel, P. S., Single-cell analysis tools for drug discovery and development. *Nat Rev Drug Discov* **2016**, *15*(3), 204-16.

2. Klein, A. M.; Mazutis, L.; Akartuna, I.; Tallapragada, N.; Veres, A.; Li, V.; Peshkin, L.; Weitz, D. A.; Kirschner, M. W., Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **2015**, *161*(5), 1187-1201.

3. Macosko, E. Z.; Basu, A.; Satija, R.; Nemesh, J.; Shekhar, K.; Goldman, M.; Tirosh, I.; Bialas, A. R.; Kamitaki, N.; Martersteck, E. M.; Trombetta, J. J.; Weitz, D. A.; Sanes, J. R.; Shalek, A. K.; Regev, A.; McCarroll, S. A., Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **2015**, *161*(5), 1202-1214.

4. Baumgarth, N.; Roederer, M., A practical approach to multicolor flow cytometry for immunophenotyping. *J Immunol Methods* **2000**, *243*(1-2), 77-97.

5. Bendall, S. C.; Simonds, E. F.; Qiu, P.; Amir el, A. D.; Krutzik, P. O.; Finck, R.; Bruggner, R. V.; Melamed, R.; Trejo, A.; Ornatsky, O. I.; Balderas, R. S.; Plevritis, S. K.; Sachs, K.; Pe'er, D.; Tanner, S. D.; Nolan, G. P., Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **2011**, *332*(6030), 687-96.

6. Hughes, C. S.; Foehr, S.; Garfield, D. A.; Furlong, E. E.; Steinmetz, L. M.; Krijgsveeld, J., Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* **2014**, *10*, 757.

7. Hughes, C. S.; McConnechy, M. K.; Cochrane, D. R.; Nazeran, T.; Karnezis, A. N.; Huntsman, D. G.; Morin, G. B., Quantitative Profiling of Single Formalin Fixed Tumour Sections: proteomics for translational research. *Sci Rep* **2016**, *6*, 34949.
8. Myers, S. A.; Rhoads, A.; Cocco, A. R.; Peckner, R.; Haber, A. L.; Schweitzer, L. D.; Krug, K.; Mani, D. R.; Clauser, K. R.; Rozenblatt-Rosen, O.; Hacohen, N.; Regev, A.; Carr, S. A., Streamlined Protocol for Deep Proteomic Profiling of FAC-sorted Cells and Its Application to Freshly Isolated Murine Immune Cells. *Mol Cell Proteomics* 2019, 18 (5), 995-1009.

9. Zhang, P.; Gaffrey, M. J.; Zhu, Y.; Chrisler, W. B.; Fillmore, T. L.; Yi, L.; Nicora, C. D.; Zhang, T.; Wu, H.; Jacobs, J.; Tang, K.; Kagan, J.; Srivastava, S.; Rodland, K. D.; Qian, W. J.; Smith, R. D.; Liu, T.; Wiley, H. S.; Shi, T., Carrier-Assisted Single-Tube Processing Approach for Targeted Proteomics Analysis of Low Numbers of Mammalian Cells. *Anal Chem* 2019, 91 (2), 1441-1451.

10. Dou, M.; Tsai, C. F.; Piehowski, P. D.; Wang, Y.; Fillmore, T. L.; Zhao, R.; Moore, R. J.; Zhang, P.; Qian, W. J.; Smith, R. D.; Liu, T.; Kelly, R. T.; Shi, T.; Zhu, Y., Automated Nanoflow Two-Dimensional Reversed-Phase Liquid Chromatography System Enables In-Depth Proteome and Phosphoproteome Profiling of Nanoscale Samples. *Anal Chem* 2019, 91 (15), 9707-9715.

11. Zhu, Y.; Piehowski, P. D.; Zhao, R.; Chen, J.; Shen, Y.; Moore, R. J.; Shukla, A. K.; Petyuk, V. A.; Campbell-Thompson, M.; Mathews, C. E.; Smith, R. D.; Qian, W. J.; Kelly, R. T., Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat Commun* 2018, 9 (1), 882.

12. Thompson, A.; Schafer, J.; Kuhn, K.; Kienie, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C., Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* 2003, 75 (8), 1895-904.
13. Russell, C. L.; Heslegrave, A.; Mitra, V.; Zetterberg, H.; Pocock, J. M.; Ward, M. A.; Pike, I., Combined tissue and fluid proteomics with Tandem Mass Tags to identify low-abundance protein biomarkers of disease in peripheral body fluid: An Alzheimer’s Disease case study. *Rapid Commun Mass Spectrom* **2017**, *31* (2), 153-159.

14. Zubiri, I.; Lombardi, V.; Bremang, M.; Mitra, V.; Nardo, G.; Adiutori, R.; Lu, C. H.; Leoni, E.; Yip, P.; Yildiz, O.; Ward, M.; Greensmith, L.; Bendotti, C.; Pike, I.; Malaspina, A., Tissue-enhanced plasma proteomic analysis for disease stratification in amyotrophic lateral sclerosis. *Mol Neurodegener* **2018**, *13* (1), 60.

15. Budnik, B.; Levy, E.; Harmange, G.; Slavov, N., SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol* **2018**, *19* (1), 161.

16. Yi, L.; Tsai, C. F.; Dirice, E.; Swensen, A. C.; Chen, J.; Shi, T.; Gritsenko, M. A.; Chu, R. K.; Piehowski, P. D.; Smith, R. D.; Rodland, K. D.; Atkinson, M. A.; Mathews, C. E.; Kulkarni, R. N.; Liu, T.; Qian, W. J., Boosting to Amplify Signal with Isobaric Labeling (BASIL) Strategy for Comprehensive Quantitative Phosphoproteomic Characterization of Small Populations of Cells. *Anal Chem* **2019**, *91* (9), 5794-5801.

17. Dou, M.; Clair, G.; Tsai, C. F.; Xu, K.; Chrisler, W. B.; Sontag, R. L.; Zhao, R.; Moore, R. J.; Liu, T.; Pasa-Tolic, L.; Smith, R. D.; Shi, T.; Adkins, J. N.; Qian, W. J.; Kelly, R. T.; Ansong, C.; Zhu, Y., High-Throughput Single Cell Proteomics Enabled by Multiplex Isobaric Labeling in a Nanodroplet Sample Preparation Platform. *Anal Chem* **2019**, *91* (20), 13119-13127.

18. Tan, Z.; Yi, X.; Carruthers, N. J.; Stemmer, P. M.; Lubman, D. M., Single Amino Acid Variant Discovery in Small Numbers of Cells. *J Proteome Res* **2019**, *18* (1), 417-425.
19. Shi, T.; Niepel, M.; McDermott, J. E.; Gao, Y.; Nicora, C. D.; Chrisler, W. B.; Markillie, L. M.; Petyuk, V. A.; Smith, R. D.; Rodland, K. D.; Sorger, P. K.; Qian, W. J.; Wiley, H. S., Conservation of protein abundance patterns reveals the regulatory architecture of the EGFR-MAPK pathway. *Sci Signal* **2016**, *9* (436), rs6.

20. Zecha, J.; Satpathy, S.; Kanashova, T.; Avanessian, S. C.; Kane, M. H.; Clauser, K. R.; Mertins, P.; Carr, S. A.; Kuster, B., TMT Labeling for the Masses: A Robust and Cost-efficient, In-solution Labeling Approach. *Mol Cell Proteomics* **2019**, *18* (7), 1468-1478.

21. Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **2008**, *26* (12), 1367-72.

22. Tyanova, S.; Temu, T.; Cox, J., The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **2016**, *11* (12), 2301-2319.

23. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J., The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **2016**, *13* (9), 731-40.

24. Zubarev, R. A.; Makarov, A., Orbitrap mass spectrometry. *Anal Chem* **2013**, *85* (11), 5288-96.

25. Sun, B.; Kovatch, J. R.; Badiong, A.; Merbouh, N., Optimization and Modeling of Quadrupole Orbitrap Parameters for Sensitive Analysis toward Single-Cell Proteomics. *J Proteome Res* **2017**, *16* (10), 3711-3721.

26. Makarov, A.; Denisov, E.; Lange, O.; Horning, S., Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. *J Am Soc Mass Spectrom* **2006**, *17* (7), 977-982.
27. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wuhr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P., MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Anal Chem 2014, 86 (14), 7150-8.

28. Erickson, B. K.; Mintseris, J.; Schwegge, D. K.; Navarrete-Perea, J.; Erickson, A. R.; Nusinow, D. P.; Paulo, J. A.; Gygi, S. P., Active Instrument Engagement Combined with a Real-Time Database Search for Improved Performance of Sample Multiplexing Workflows. J Proteome Res 2019, 18 (3), 1299-1306.

29. Specht, H.; Emmott, E.; Perlman, D. H.; Koller, A.; Slavov, N., High-throughput single-cell proteomics quantifies the emergence of macrophage heterogeneity. bioRxiv 2019, 665307.

30. Bekker-Jensen, D. B.; Martinez-Val, A.; Steigerwald, S.; Ruther, P. L.; Fort, K. L.; Arrey, T. N.; Harder, A.; Makarov, A. A.; Olsen, J. V., A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients. Mol Cell Proteomics 2020. in press, mcp.TIR119.001906

31. Hebert, A. S.; Prasad, S.; Belford, M. W.; Bailey, D. J.; McAlister, G. C.; Abbatiello, S. E.; Huguet, R.; Wouters, E. R.; Dunyach, J. J.; Brademan, D. R.; Westphall, M. S.; Coon, J. J., Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer. Anal Chem 2018, 90 (15), 9529-9537.

32. Okuda, S.; Watanabe, Y.; Moriya, Y.; Kawano, S.; Yamamoto, T.; Matsumoto, M.; Takami, T.; Kobayashi, D.; Araki, N.; Yoshizawa, A. C.; Tabata, T.; Sugiyama, N.; Goto, S.; Ishihama, Y., jPOSTrepo: an international standard data repository for proteomes. Nucleic Acids Res 2017, 45 (D1), D1107-D1111.
Figure 1

(a) TMT channel & Input

| TMT channel | Input  |
|-------------|--------|
| 126         | 0.5ng  |
| 127N        | 0.5ng  |
| 127C        | 0.5ng  |
| 128N        | H₂O    |
| 128C        | H₂O    |
| 129N        | H₂O    |
| 129C        | H₂O    |
| 130N        | H₂O    |
| 130C        | H₂O    |
| 131N        | 10x, 50x, 100x or 200x |

(b) Number of quantified peptides

| Boosting ratio | 3000 | 4000 | 5000 | 6000 | 7000 |
|----------------|------|------|------|------|------|
| 10x            | ![Graph](image1) |
| 50x            | ![Graph](image2) |
| 100x           | ![Graph](image3) |
| 200x           | ![Graph](image4) |

(c) Log₂ TMT intensity

(d) Pearson correlation

(e) No. of peptides vs. CV (%)

(f) Injection time (ms)

(by guest on May 6, 2020)
Figure 3

(a) Bar chart showing the number of quantifiable peptides and proteins for Normal and iBASIL conditions.

(b) Box plot comparing Log2 TMT intensity across different cell types (MOLM-14, K562, CMK) under Normal and iBASIL conditions.

(c) PCA analysis for Normal and iBASIL conditions, with the projection on the first two principal components.
Figure 4

(a) TMT Cell type Input (ng)

| TMT | Cell type | Input (ng) |
|-----|-----------|------------|
| 126 | MOLM-14   | 0.1        |
| 127N| K562      | 0.1        |
| 127C| MOLM-14   | 0.1        |
| 128N| K562      | 0.1        |
| 128C| CMK       | 0.1        |
| 129N| MOLM-14   | 0.1        |
| 129C| CMK       | 0.1        |
| 130N| Empty     | 0          |
| 130C| K562      | 0.1        |
| 131N| Boost 0x, 100x, 1000x |          |

(b) 

(c) 

(d)
