Prioritized single-cell proteomics reveals molecular and functional polarization across primary macrophages

R Gray Huffman,[1] Andrew Leduc,[1] Christoph Wichmann,[2] Marco di Gioia,[3] Francesco Borriello,[3] Harrison Specht,[1] Jason Derks,[1] Saad Khan,[1] Edward Emmott,[1,4] Aleksandra A. Petelski,[1] David H Perlman,[5] Jürgen Cox,[2] Ivan Zanoni,[3] and Nikolai Slavov[1]✉

[1]Departments of Bioengineering, Biology, Chemistry and Chemical Biology, Single Cell Center, and Barnett Institute, Northeastern University, Boston, MA 02115, USA

[2]Computational Systems Biochemistry Research Group, Max Planck Institute of Biochemistry, Germany

[3]Boston Children's Hospital and Harvard Medical School, Boston, MA 02115, USA

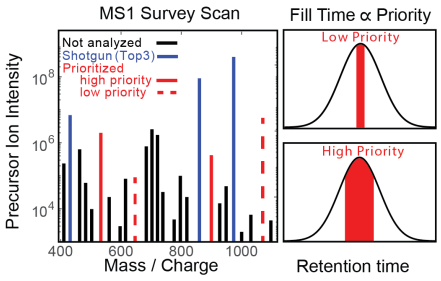
[4]Centre for Proteome Research, Department of Biochemistry & Systems Biology, Institute for Systems, Molecular & Integrative Biology, University of Liverpool, Liverpool, L69 7ZB, UK

[5]Merck Exploratory Sciences Center, Merck Sharp & Dohme Corp., 320 Bent St. Cambridge, MA 02141

✉ Correspondence: nslavov@northeastern.edu

∈ Data, code & protocols: scp.slavovlab.net/pSCoPE

Major aims of single-cell proteomics include increasing the consistency, sensitivity, and depth of protein quantification, especially for proteins and modifications of biological interest. To simultaneously advance all of these aims, we developed prioritized Single Cell ProtEomics (pSCoPE). pSCoPE ensures duty-cycle time for analyzing prioritized peptides across all single cells (thus increasing data consistency) while analyzing identifiable peptides at full duty-cycle, thus increasing proteome depth. These strategies increased the quantified data points for challenging peptides and the overall proteome coverage about 2-fold. pSCoPE enabled quantifying proteome polarization in primary mouse macrophages and linking it to phenotypic variability in endocytic activity. Proteins annotated to phagosome maturation and proton transport showed concerted variation for both untreated and lipopolysaccharide-treated macrophages, indicating a conserved axis of polarization. pSCoPE further quantified proteolytic products, suggesting a gradient of cathepsin activities within a treatment condition. pSCoPE is easily accessible and likely to benefit many applications, especially mechanistic analysis seeking to focus on proteins of interest without sacrificing proteome coverage.



Macrophages are innate immune myeloid cells performing diverse functions in development, tissue homeostasis and immune responses. Despite this diversity, macrophages are traditionally described in terms of dichotomous states (M1, pro-inflammatory; M2, anti-inflammatory). Single-cell measurements, though, have revealed a more complex and continuous spectrum of macrophage polarization in terms of molecular and functional phenotypes[1–3]. Thus, we sought to explore this continuum of polarized states in primary macrophages using single-cell mass-spectrometry (MS). Shotgun MS methods can analyze hundreds of single cells per day and quantify thousands of proteins but remain biased towards abundant proteins[3–11]. This bias reflects an intentionally programmed 'topN' heuristic for selecting the N most abundant peptide precursors for sequence identification and quantification, as illustrated in Fig. 1a[12].

Peptide selection by the topN heuristic is limited by three challenges: (i) abundance bias, (ii) stochasticity, which results in missing data across single cells, and (iii) precursors that cannot be confidently sequenced waste much of the instrument time. Such inefficient use of time is particularly limiting for single-cell proteomics due to the long ion accumulation times needed for sensitive MS[13]. While no existing method resolves all 3 challenges, the challenges can be partially mitigated. For example, targeted MS can alleviate challenges (i) and (ii) but has remained limited to analyzing hundreds of peptides or fewer[14–20]. Real-time database searching can increase the fraction of sequenced peptide features and alleviate challenge (iii), but it has not allowed for selecting peptides of interest[21,22]. We sought to simultaneously address all three challenges by prioritized selection of peptide precursors, Fig. 1a.

Results

Prioritized analysis aims to simultaneously maximize the consistency of peptide analysis, proteome coverage, and instrument time utilization. To achieve these aims, we built upon the real-time retention-time alignment of MaxQuant.Live[23] and introduced prioritization tiers that define priority of peptide analysis when the duty-cycle time is insufficient for analyzing all peptides, Fig. 1a.

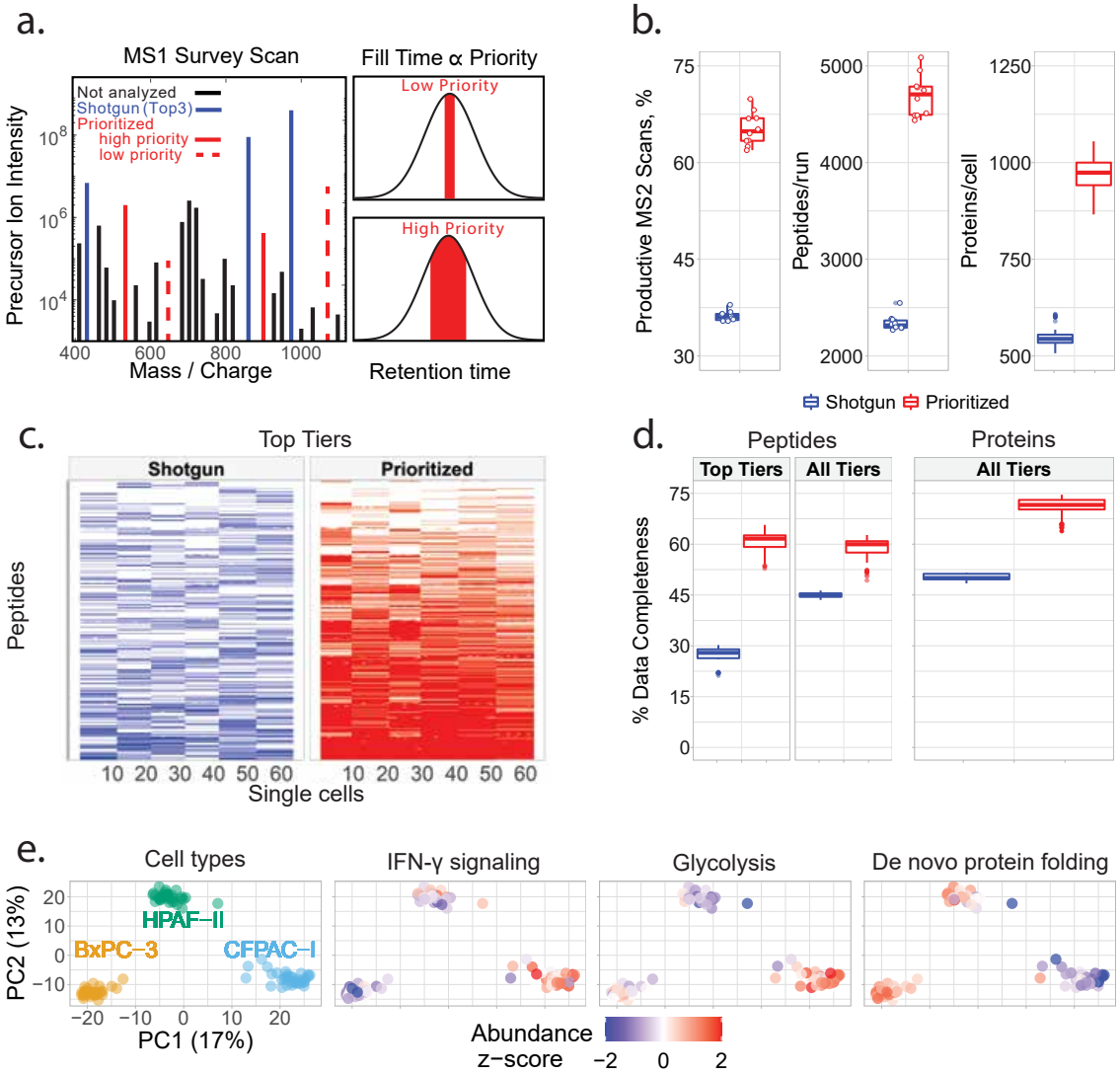


Figure 1 | Prioritized analysis increases increases proteome coverage and data completeness. (**a**) Shotgun TopN analysis selects the N most abundant precursors for isolation and fragmentation. Prioritized analysis selects the precursors with highest priority (shown in solid red) and then from lower priority tiers (shown with dashed red lines). It also affords longer accumulation times and thus increased sensitivity for high priority peptides. (**b**) Relative to shotgun analysis, prioritization increased the fraction of MS2 scans assigned to peptide sequences, the number of peptides per run (60-min active gradient) and the number of quantified proteins per single cell. (**c**) A heatmap showing quantification consistency across single cells (columns) for 857 peptides (rows) from the top two priority tiers. (**d**) Prioritized analysis reduces missing data at both peptide and protein levels across all priority tiers. (**e**) Principal component analysis of the single cells associated with (**b**) cluster cluster by time. Protein sets enriched in the PCs are visualized by color-coding the single cells by the median protein abundance of the set in each cell.All experiments used 60 min active gradients per run. All peptide and protein identifications were filtered at 1% FDR with additional filtration metrics indicated in Methods.

This prioritization approach maximizes productive instrument usage by supplying identifiable peptide precursors via the low-priority (bottom tier) portion of the inclusion list, allowing a larger

fraction of MS2 scans to be confidently matched to amino acid sequences. Yet, this full duty-cycle regime does not reduce the probability of MS2 analysis (and thus consistency of quantification) for peptides from the high-priority (top tier) portion of the inclusion list, because these precursors are preferentially isolated for MS2 scans, Fig. S1. The prioritization scheme also allows for selectively increased accumulation times for peptides of biological interest (Fig. 1a), which increases the number of ion copies sampled from these peptides.

Increasing proteome coverage and quantification consistency

We applied prioritized Single-Cell ProtEomics (pSCoPE) to pancreatic ductal adenocarcinoma (PDAC) cells to evaluate the depth and the consistency of proteome coverage, Fig. 1b-d. Single-cell samples were prepared by nano-ProteOmic sample Preparation (nPOP)[24]. They were analyzed by either shotgun or prioritized methods using 60-min active chromatographic gradients and narrow isolation windows (0.5 Th), which resulted in good quantitative agreement between different peptides originating from the same protein, Fig. S2a,b. Relative to shotgun analysis, pSCoPE increased the fraction of MS2 spectra assigned to a confident peptide sequence by 2-fold, reaching 65%. The remaining 35% of MS2 spectra correspond to sequences generating few peptide fragments and having low confidence of identification in previous experiments used for generating the inclusion list, Fig. S3. The increase in productive MS2 scans doubled the number of unique peptides per run (increased by 103%) and increase the number of quantified proteins per single cell by 75%, Fig. 1b.

Next, we evaluated pSCoPE's consistency of quantifying peptides that were identified with less than 50% probability in shotgun SCoPE sets. The results from six representative sets indicate that pSCoPE significantly increased the quantification consistency for these challenging peptides, Fig. 1c. By using pSCoPE, the peptide-level data completeness per single cell was increased by 121% for the challenging peptides (top tier) and by about 34% for peptides from all priority tiers. At the protein level, pSCoPE also increased data completeness to about 72% for all proteins, which represents a 43% gain over the performance of shotgun analysis, Fig. 1d. The tiered approach im-

plemented here allows for high-probability analysis of thousands of peptides. Specifically, pSCoPE sent for MS2 scans 99% of the peptides from a 970-peptide list and 96% of the peptides from a 3,028-peptide list Fig. S1b, and 81% of the peptides from a 6,245-peptide list, Fig. S1a.

These performance benchmarking sets included single cells from three PDAC cell lines with epithelial origin and thus allowed us to examine protein variation between single cells from the sample cell type. These epithelial cells were clearly separated by PCA, Fig. 1e. Protein Set Enrichment Analysis (PSEA) performed on the principal components (PCs) identified enrichment for multiple functional sets of proteins, including glycolysis, protein folding and the IFN-γ signaling pathway (represented by proteins such as HLA-A, ICAM1, HLA-DRB1, HLA-A, SUMO1, PRKCD, PML), Fig. 1e. These results demonstrate the ability of pSCoPE to identify biological differences between closely related cell lines.

Polarized proteome states

Next we used pSCoPE to explore the molecular and functional heterogeneity of murine bone-marrow-derived macrophages (BMDMs) responding to inflammatory stimuli, such as lipopolysaccharide (LPS), the major component of gram-negative bacteria's outer membrane. The macrophages were differentiated using M-CSF and either treated with LPS for 24 hours or untreated. Applying pSCoPE, we quantified 1,123 proteins across 373 single cells, achieving 71% data completeness for all proteins (Fig. S4) and good quantitative agreement between peptides originating from the same protein (Fig. S2c). The PCA projection of the data results in 2 clusters corresponding to the treatment conditions, Fig. 2a. Projected bulk samples cluster with the corresponding treatment groups, indicating that the cluster separation reflects treatment response. This treatment-specific clustering is also reflected in the abundance of proteins that vary across treatments but not within a treatment, as exemplified by proteins functioning in the type-1 interferon-mediated signaling pathway, Fig. 2a.

The spread of the macrophage clusters along PC1 suggests that proteins varying across treatment groups may also vary within a treatment group. Indeed, PSEA using the PCA loadings

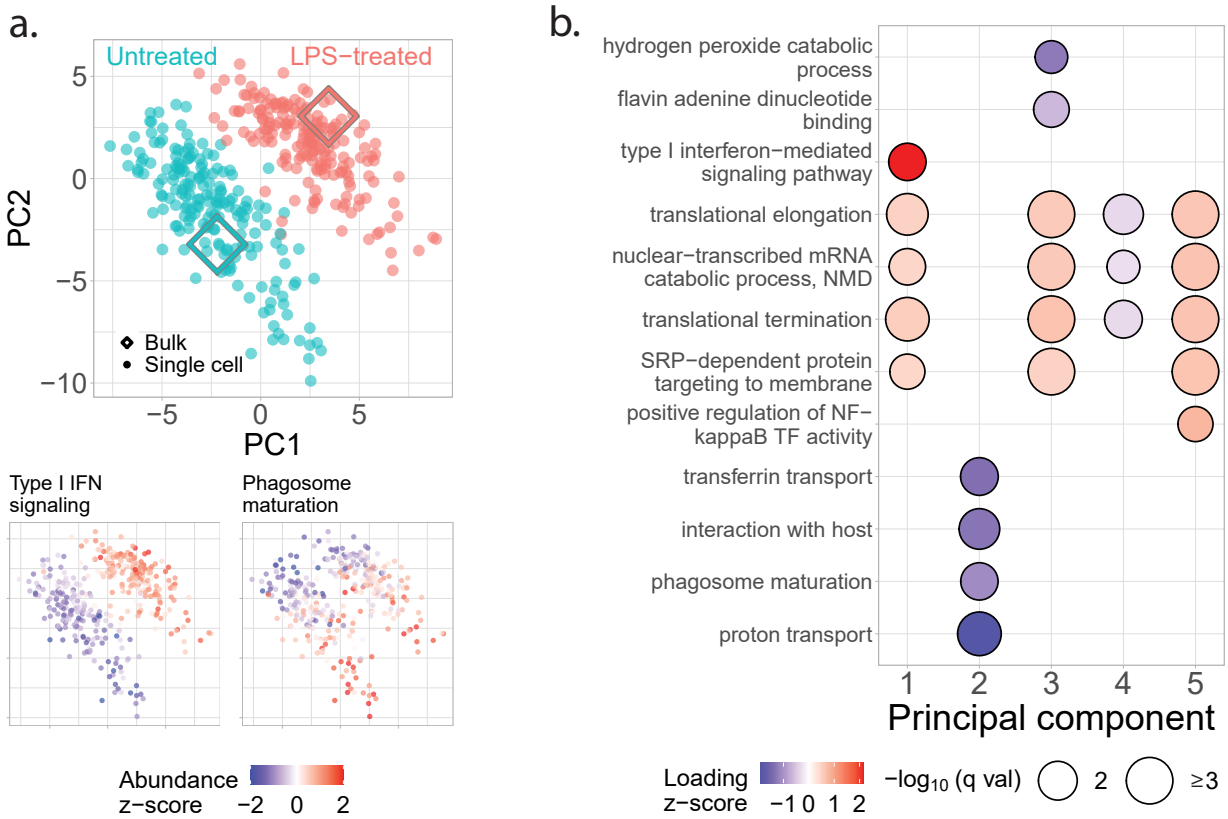


Figure 2 | Prioritized analysis of primary macrophages identifies protein variation within and across treatment conditions. (**a**) Principal component analysis of 373 BMDM and 1,123 proteins color-coded by treatment condition. The diamond markers indicate bulk samples projected in the same low-dimensional space as the single cells. The adjoining PCA plots are color-coded by the z-scored median relative abundance of proteins corresponding to type I interferon-mediated signaling and phagosome maturation. (**b**) Protein groups identified by protein set enrichment analysis (PSEA) performed using the PC vectors with protein weights from the PCA shown in panel a.

identified such protein sets, as exemplified with the phagosome maturation pathway, Fig. 2a-b. This protein heterogeneity within a treatment group is reminiscent of previously reported RNA heterogeneity within dendritic cells[25]. These finding are recapitulated when the PCA is performed without data imputation (Fig. S5), suggesting robustness to choices of data analysis. Additionally, color-coding the original PCA by per-sample data completeness indicates that the cross-condition and intra-condition sample separation is not driven by missing data, Fig. S6. To systematically investigate proteome variations within a condition, we performed PCA of each treatment group separately and PSEA on the associated PCA protein loading. Remarkably, the first PCs of the treated and untreated macrophages correlate strongly ($r = 0.8$, $p < 10^{-15}$), suggesting that the

within-condition protein variability is similar across the two conditions. This observation is naturally reflected in very similar functional enrichment results for treated and untreated macrophages, Fig. 3a.

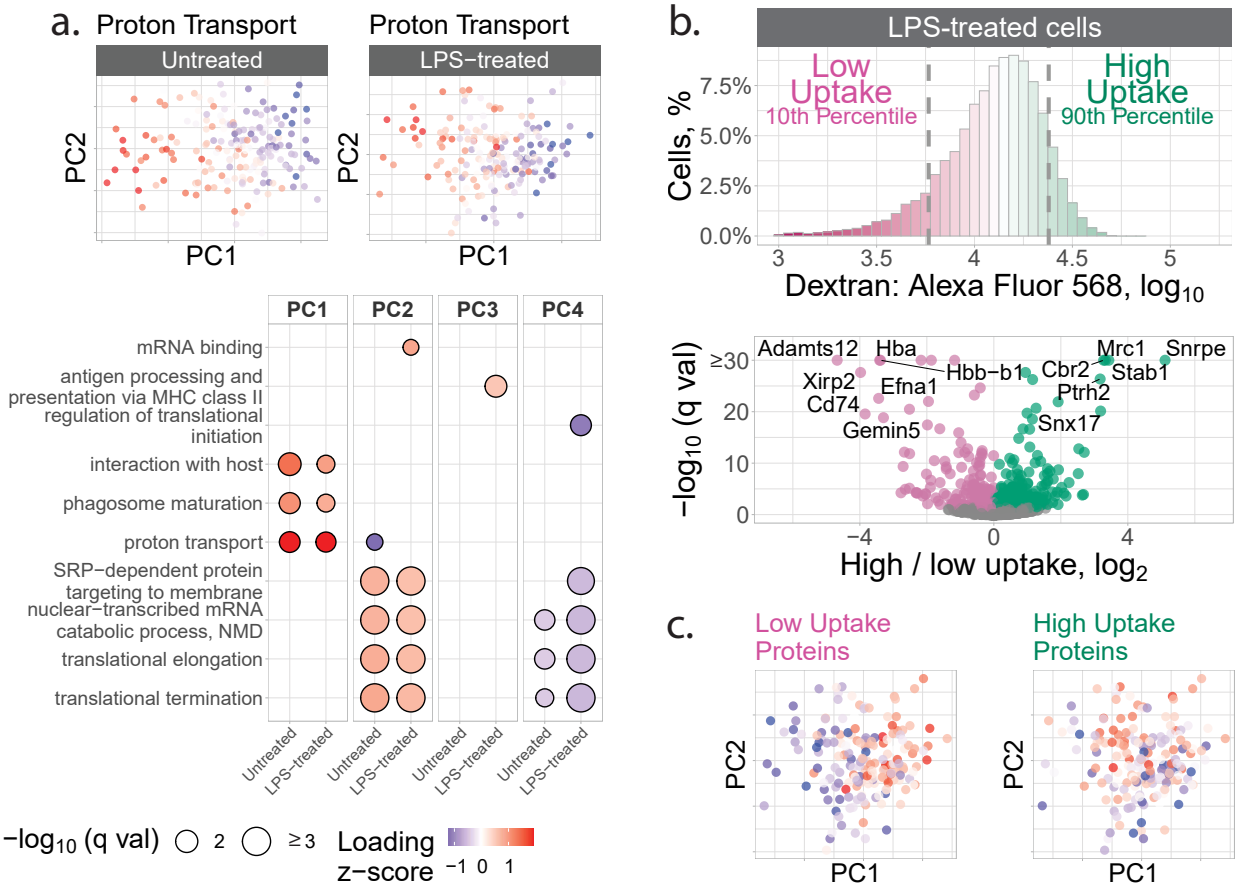


Figure 3 | Axes of proteome polarization are similar between untreated and LPS-treated macrophages and correlated to Dextran update (**a**) Untreated and LPS-treated macrophages were analyzed separately by PCA and PSEA performed on the corresponding PCs. The PCA plots are color-coded by the median abundance of proteins annotated to proton transport. (**b**) The uptake of fluorescent dextran by LPS-treated macrophages was measured by FACS, and the cells with the lowest and highest uptake were isolated for protein analysis. The volcano plot displays the fold changes for deferentially abundant proteins and the associated statistical significance. (**c**) The LPS-stimulated macrophages were displayed in the space of their PCs and color-coded by the median abundance of the low-uptake or the high-uptake proteins. The low-uptake proteins correlate to PC1 (Spearman $r = 0.55$, $q \leq 3 \times 10^{-15}$) and the high-uptake proteins correlate to PC2 (Spearman $r = 0.33$, $q \leq 2 \times 10^{-5}$).

These results suggest that a 24-hr LPS treatment does not fundamentally alter axes of protein variation for murine BMDMs, such as phagosome maturation, proton transport, and protein targeting to the membrane, Fig. 3a. In addition to these shared functional groups, some protein sets vary

only within the LPS-stimulated cells, as illustrated by proteins annotated to antigen processing and presentation via MHC class II, mRNA binding, and regulation of translational initiation, Fig. 3a. The coherence of protein variability within functional groups suggests it is functionally relevant, but it does not prove it.

Connecting protein variation to functional variation

To examine whether the observed protein heterogeneity has functional consequences, we sought to directly measure the endocytic activity of macrophages and its relationship to such protein heterogeneity. To this end, we measured the uptake of fluorescently labeled dextran and FACS sorted macrophages from the top and bottom deciles of the fluorescence distribution, Fig. S7. Both the LP-treated (Fig. 3b) and the untreated macrophages (Fig. S8) exhibit large variance in dextran uptake per cell, with the median uptake being higher for the LPS-treated cells. The proteomes of the sorted subpopulations were analyzed by data independent acquisition, which allowed us to identify proteins whose abundance is significantly different between the most and least endocytically active cells, Fig. 3b and Fig. S8. Then for each cell, we estimated the median abundances of these proteins associated with endocytic activity and correlated them to the PCs for each treatment condition. For the LPS-treated samples, the proteins associated with high dextran uptake (such as Mrc1, Stab1, and Snx17) were found to be significantly correlated to PC1, while the proteins annotated to low dextran uptake were inversely correlated to PC1 and significantly correlated to PC2, Fig. 3c. Notably, some proteins (such as Mrc1 and Stab1, and Cd74) exhibit similar association with dextran uptake both in the untreated and LPS-treated macrophages, Fig. 3b and Fig. S8.

To more directly measure regulatory mechanisms, we sought to quantify proteolysis, which plays major functional roles in macrophage activation[27–29]. To avoid products of proteolysis that may occur during sample preparation, we focused only on proteolytic products present in the macrophages prior to trypsin digestion. These products were identified in discovery bulk samples in which amine groups were covalently labeled prior to trypsin digestion as commonly performed[30,31]. The proteolytic products were matched to annotated proteolytic products in the MEROPS

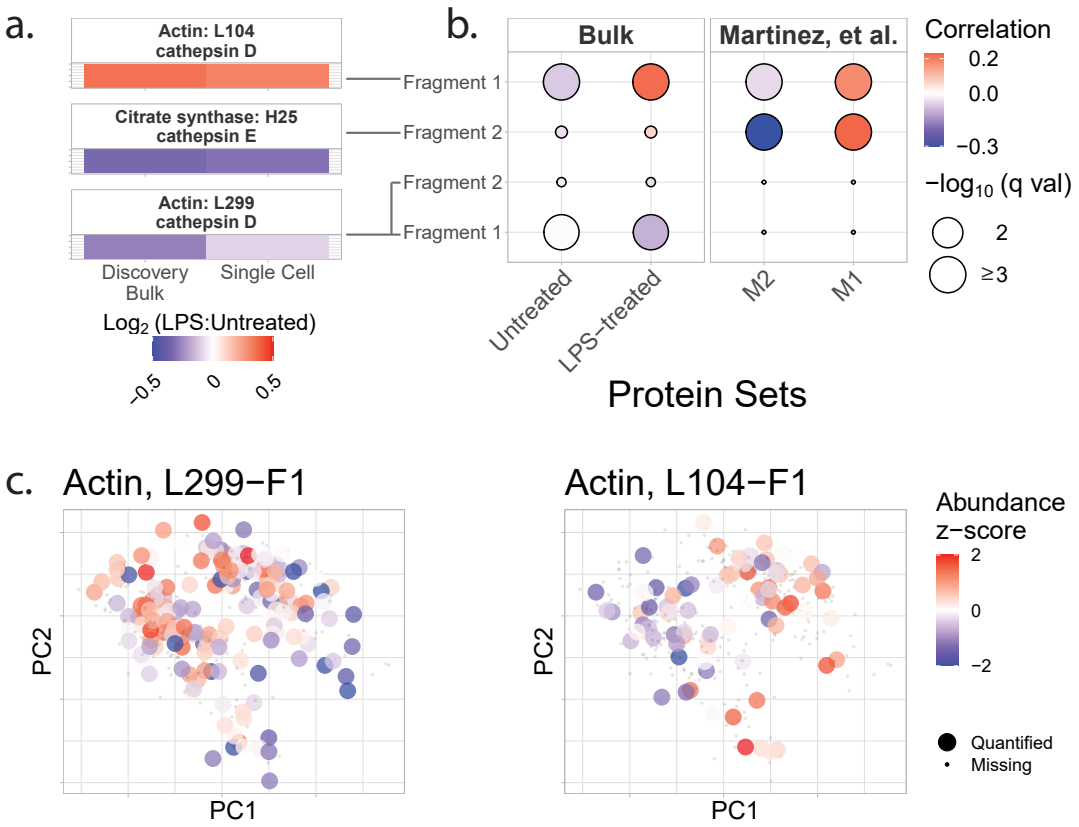


Figure 4 | Proteolytic products in individual macrophages correlate to inflammatory markers and vary within treatment groups (**a**) A comparisons between untreated and LPS-treated ratios of proteolytic products quantified in discovery bulk experiments and in single cells. The annotations are derived from the MEROPS database[26]. (**b**) Correlation analysis of proteolyic products to treatment-group-specific and macrophage-polarization-specific protein panels. (**c**) The untreated and LPS-treated cells were projected by PCA and color-coded by the relative abundance of the indicated actin fragments.

database[26] and analyzed by pSCoPE in single cells. To evaluate the single-cell quantification, the fold-changes between LPS-treated and untreated cells were compared to the corresponding bulk estimates, Fig. 4a. The good agreement between the measurements from established bulk methods and pSCoPE support the accuracy of the single-cell quantification. To infer the functional association of the validated proteolytic products, we correlated their single-cell abundances to the abundances of pro- and anti-inflammatory protein panels, Fig. 4b: (i) proteins that we identified as differentially abundant between bulk samples of untreated or LPS-stimulated macrophages, and (ii) previously reported markers for M1 or M2 macrophages.[32] The cathepsin-D-cleaved actin peptide (L104) and the cathepsin-E-cleaved citrate synthase peptide (H25) were found to be significantly

positively correlated to inflammatory markers. Both peptide fragments annotated to cathepsin D cleavage at L299 were inversely correlated to the set of proteins which were more abundant in LPS-treated BMDMs.

Having established the reliability of single-cell quantification of proteolytic products across conditions, we next examined their abundance within a condition, Fig. 4c. The data indicate that the actin proteolytic products exhibit significant variability within each treatment condition. For example, the actin fragment cleaved at L299 correlates significantly to PC1 (Spearman $r = -0.32$, $p < 2 \times 10^{-5}$), Fig. 4c. These results point to the possibility of using pSCoPE for analysing proteolytic activity at single-cell resolution.

Discussion

Our analysis demonstrates the potential of prioritized data acquisition to simultaneously optimize multiple aspects of single-cell proteomics[13], including the consistency and depth of protein quantification. These gains are achieved using multiplexed and widely accessible methods[5] and a new software module that is compatible with any Orbitrap instrument, Fig. S9. The demonstrated ability of these accessible methods to relate protein and functional heterogeneity is likely to find many applications, especially for extending single-cell proteomics to less abundant proteins and post-translational modifications, such as proteolytic cleavage.

Many MS methods allow for analyzing a pre-selected group of peptides. They range from targeted methods that maximize sensitivity and probability of target quantification[14–20] to directed methods that use inclusion lists[18,20]. pSCoPE extends the directed family of methods by introducing a tiered approach that allows the prioritization of thousands of peptides for isolation and fragmentation (thus achieving 96% success rate of sending high-priority peptides for MS2) while working at a full duty cycle and thus achieving high proteome coverage, Fig. 1. This tiered approach may also be implemented with other approaches for performing real-time retention time alignment[33,34].

Prioritization can help increase the throughput of single-cell proteomics by enabling consistent analysis of proteins of interest on short chromatographic gradients[35]. All results presented here used 60 min active gradients, though shorter gradients may increase both the throughput and the sensitivity (via narrower elution peaks) while still affording enough time to analyze thousands of prioritized peptides. Thus, pSCoPE may provide accurate and consistent protein quantification across many single cells to support sufficiently powered biological investigations[36,37].

Data Availability

The raw data and search results are available at MassIVE: MSV000089055

Code Availability

Data, code & protocols are available at scp.slavovlab.net/pSCoPE, github.com/SlavovLab/pSCoPE, and scp.slavovlab.net/Huffman_et_al_2022

Acknowledgments: We thank T. Colombani and M. Jovanovic for discussions and constructive comments. This work was funded by a New Innovator Award from the NIGMS from the National Institutes of Health to N.S. under Award Number DP2GM123497, an Allen Distinguished Investigator award through The Paul G. Allen Frontiers Group to N.S., a Seed Networks Award from CZI CZF2019-002424 to N.S., and through a Merck Exploratory Science Center Fellowship, Merck Sharpe & Dohme Corp. to N.S.

Competing Interests: The authors declare that they have no competing financial interests.

Correspondence: Correspondence and materials requests should be addressed to nslavov@alum.mit.edu and nslavov@northeastern.edu

Author Contributions

Experimental design: R.G.H., N.S.

LC-MS/MS: R.G.H., A.L., H.S., J.D., S.K., A.P, E.E.

Sample preparation: R.G.H., A.L., M.G., F.B.

Cell culture: A.L., M.G., F.B

Raising funding: N.S.

Supervision: N.S.

Data analysis: R.G.H., N.S.

Software development: C.W.

Initial draft: R.G.H., N.S.

Results interpretation: N.S, J.C., I.Z., R.G.H., A.L., H.S., J.D., S.K., M.G., F.B., A.P., E.E., D.H.P.

Writing: All authors contributed to writing the paper and approved the final manuscript.

Methods

Implementation of prioritized analysis

To maximize the probability of analyzing high-priority peptides (i.e., peptides of high experimental importance) when operating at full duty cycle, we developed a new feature of MaxQuant.Live[23]: *multi-tier prioritization*. Multi-tier prioritization uses the real-time instrument control capabilities of MaxQuant.Live and adds a priority feature that determines which precursors are analyzed when duty-cycle time becomes limiting. The initial priority for each peptide is a user-defined integer number which is by default set to zero. By assigning non-zero values, it is possible to prioritize a single set of peptides or to implement a multi-tier approach, depending on the experimental objectives. During data acquisition, the peptides are selected for fragmentation based on their priority. After each fragmentation event, the corresponding peptide priority value is decremented unless fragmentation occurred outside of the retention-time tolerance. The prioritization feature is part of the latest release of MaxQuant.Live (version 2.1), available at: MaxQuant.Live and scp.slavovlab.net/pSCoPE.

The initial user-defined priorities are set via a column in the inclusion list table. This column was added to allow for easy definition of priority for every peptide on the list. The higher the integer number associated with a peptide (and thus its tier), the higher the probability that it will be chosen for fragmentation when duty cycle is limited. MaxQuant.Live was tested on a Q Exactive (as described below), but it was written to be compatible with all Orbitrap instruments.

Prioritization workflow

All prioritized single-cell experiments followed the four stages of the workflow displayed in Fig. S9 and described below.

1. Compilation of proteins of interest from literature or prior LC-MS/MS analyses

2. DIA analysis of a 1x concentrated injection of the combined carrier-reference sample to generate accurate retention times for precursors which will subsequently be prioritized.
    - This step is enabled by using a spectral library generated from prior DIA analysis of a 5-10x concentrated injection of the combined carrier-reference samples.
3. Assignment of precursors identified in step 2 to priority tiers based on proteins of interest defined in step 1.
    - The minimal set of precursor characteristics needed for a prioritized inclusion list are the mass, expected apex retention time, and priority.
4. Acquire data from SCoPE samples using MaxQuant.Live's prioritization feature and the inclusion list generated in step 3.
    - Performing a test run on a 1x injection of the combined carrier and reference samples can be useful for troubleshooting methods before acquiring data from single cells.

Benchmarking prioritized and shotgun analysis using PDAC cells

These experimental sets, the results of which were presented in Fig. 1, were designed to assess the relative performance of shotgun and prioritized methods with respect to sequence coverage and consistency of quantification across single-cell samples. The experiments presented in Fig. 1b/e are a matched set of ten shotgun analyses and ten prioritized analyses; the experiments presented in Fig. 1c/d are a matched set of six shotgun analyses and six prioritized analyses. The parameters for experiments that directly compared shotgun and prioritized analysis were identical, including LC gradients and data acquisition parameters with the only exception of increasing fill times for selected prioritized precursors as explicitly described in the main text. Additional information regarding sample preparation, instrument parameters, MaxQuant.Live parameters, prioritized inclusion list design, analysis of raw data, single-cell data processing, and figure generation can be found in their respective sections. **The active gradient in all experiments was 60 min.**

Bone-Marrow-Derived Macrophage samples prepared by nPOP

These experiments were designed to present a use case for prioritized LC-MS/MS methods. Twenty shotgun and forty prioritized single-cell experiments containing samples from both treatment conditions (untreated or treated for 24hrs with LPS) were conducted as part of this module. A side-by-side comparison of the twenty shotgun experiments and the first twenty prioritized experiments can be found in Fig. S4. Only the results of the forty prioritized analyses were included in Fig. 2, Fig. 3, and Fig. 4. Additional information regarding sample preparation, instrument parameters, MaxQuant.Live parameters, prioritized inclusion list design, analysis of raw data, single-cell data processing, and figure generation can be found in their respective sections. **The active gradient in all experiments was 60 min.**

Endocytosis experiments, BMDM samples

In order to identify protein sets associated with endocytosis that were specific to murine BMDMs, bulk samples from each treatment condition (untreated or treated for 24hrs with LPS) were incubated with fluorescently labeled dextran, and samples from the top and bottom deciles of dextran uptake were isolated by FACS for downstream LC-MS/MS analysis. Protein sets found to be differential between dextran uptake deciles were then added to the top priority tier in subsequent prioritized analyses of single-cell BMDM samples. Additional information regarding sample preparation, instrument parameters, raw data analysis, and differential protein detection can be found in their respective sections.

MEROPS experiments, BMDM samples

Bulk BMDM samples from each treatment condition (untreated or treated for 24hrs with LPS) were lysed, cysteine residues were reduced and alkylated, and samples were incubated with TMTPro so that all pre-digestion n-termini would be distinguishable from neo-n-termini produced by a subsequent tryptic digestion. The raw LC-MS/MS data was then searched with a FASTA database con-

taining all murine SwissProt reviewed sequences, as well as semitryptic peptides consistent with MEROPS-annotated proteolytic cleavage sites. These experiments were used to validate semitryptic MEROPS-annotated peptides observed in the prioritized single-cell samples. Additional information regarding sample preparation, MEROPS database integration, instrument parameters, and data analysis can be found in their respective sections.

Bulk BMDM sample analyses by DDA and DIA

Bulk BMDM samples from each treatment condition (untreated or treated for 24hrs with LPS) were lysed, digested, and labeled with TMTPro for DDA analysis as a duplex sample or sequentially analyzed as labeled single-condition samples via DIA. These experiments were used to identify differentially abundant proteins between the treatment conditions which were then added to the top priority tier in subsequent prioritized analyses of single-cell BMDM samples. Additional information regarding sample preparation, instrument parameters, raw data analysis, and differential protein detection can be found in their respective sections.

BMDM samples prepared via mPOP methods

This set of experiments represents an early troubleshooting investigation to both assess the sizes of the BMDMs from each treatment condition by using the cellenONE's optical system (Scienion) and contrast against data generated in a prior set of single-cell BMDM samples isolated via FACS that may have experienced sorting issues. The results from this set were not used to generate any of the publication figures and are included merely for completeness, as a subset of identifications from these experiments informed the inclusion-list construction of the nPOP-prepared pSCoPE sets. Additional information regarding sample preparation, MEROPS database integration, instrument parameters, and raw data analysis can be found in their respective sections.

LC-MS platform

The LC-MS/MS equipment and setup used for all analyses are detailed in the SCoPE2 protocol[5]. Briefly, samples were separated via online nLC on a Dionex UltiMate 3000 UHPLC; $1\mu l$ of sample was loaded onto a 25cm x $75\mu m$ IonOpticks Aurora Series UHPLC column (AUR2-25075C18A); mass spectrometry analyses were performed via a Thermo Scientific Q Exactive mass spectrometer; an Active Background Ion Reduction Device (ABIRD, by ESI Source Solutions, LLC, Woburn MA, USA) was used at the ion source to remove background contaminants. In the LC separations, buffer A was 0.1% formic acid in LC-MS-grade water, and buffer B was 80% Acetonitrile, 0.1% formic acid in LC-MS grade water; all buffer B percentages described in the subsequent instrument methods are relative to this concentration.

Cell Culture

Culturing and harvesting Pancreatic Ductal Adenocarcinoma (PDAC) cells

HPAF-II cells (ATCC CRL-1997) were cultured in EMEM (ATCC 30-2003), CFPAC-I cells (ATCC CRL-1918) were cultured in IMDM (ATCC 30-2005), and BxPC-3 cells (ATCC CRL-1687) were cultured in RPMI 1640 (ATCC 30-2001). All media were supplemented with 10% fetal bovine serum (FBS, Millipore Sigma F4135) and 1% penicillin-streptomycin (pen/strep, Thermo Fisher 15140122). Cells were passaged at 70% confluence.

Prior to harvesting, media was removed from cell cultures, which were then rinsed with 0.05 % trypsin-EDTA (Gibco, Thermo Fisher 25300054) at 4 °C. After rinsing, adherent cultures were incubated with 4 °C 0.05% trypsin-EDTA for 15 minutes, until cells were detached from the culture vessel. Cold 1x PBS was added to PDAC culture vessel, and the resulting suspension was pelleted via centrifugation at 250g, before being washed with 1x PBS and repelleted at 250g. The washed cell pellets were diluted in 1x PBS at 4 ^oC and their density estimated by counting at least 1000 cells using a hemocytometer. Cells which were harvested for carrier and reference samples were

resuspended in water (Optima LC/MS Grade, Fisher Scientific W6500) and frozen at -80 °C. Cells which were harvested for single-cell sorting on the cellenONE system were diluted in 1x PBS to a concentration of 300 cells/μl and placed on ice.

Culturing and harvesting Bone-Marrow-Derived Macrophages (BMDMs)

C57BL/6J (Jax 000664) mice were purchased from The Jackson Laboratory. Bone-marrow-derived macrophages (BMDMs) were differentiated from bone marrow in Dulbeccos modified Eagle medium (DMEM; Thermo Fisher Scientific), 30% L929-M-CSF supernatant and 10% fetal bovine serum (FBS). After 7 days, BMDMs were replated at 1x10^6 cells/ml in DMEM supplemented with 10% FBS, and each plate was either stimulated for 24 hours with LPS (Serotype O55:B5, Enzo Life Sciences) at $1\mu g$/ml or allowed to rest. Prior to harvesting, cells were washed twice with 1x PBS and incubated with PBS-2mM EDTA to detach from the plate. Cells were then spun down at 300g for 5 minutes and washed with 1x PBS before being resuspended. The washed cell pellets were diluted in 1x PBS at 4 ^oC and their density estimated by counting at least 1000 cells using a hemocytometer. Cells which were harvested for carrier and reference samples were resuspended in water (Optima LC/MS Grade, Fisher Scientific W6500) and frozen at -80 °C. Cells which were harvested for single-cell sorting on the cellenONE system were diluted in 1x PBS to a concentration of 300 cells/μl and placed on ice.

Sample preparation

Single-cell samples

PDAC single-cell sample preparation

A ∼200-cell carrier and ∼5-cell reference composed of HPAF-II, CPAC-I, and BxPC-3 cell lines were prepared following the method outlined in in the SCoPE2 protocol[5]. In addition to serving as the carrier and reference for all single-cell sets analyzed in the technical section, the combined carrier and reference sample was used in all spectral-library-generation and retention-time-calibration experiments for the coverage and consistency experiments shown in Fig. 1.

All single-cell samples were prepared using the droplet nano-ProteOmic sample Preparation (nPOP)[24] as detailed in Leduc *et al.* (2021). In addition to sorted single cells, the SCoPE sets contained negative control samples to be used for downstream quality control purposes. These negative control samples received all reagents and proceeded through all sample handling steps, but no single cells were dispensed into these droplets[5]. The distribution of protein-level CVs (i.e. quantification variability) associated with the single cell and control samples for these experiments can be found in Fig. S2a/b.

BMDM single-cell nPOP sample prep

Carrier and reference samples composed of equivalent amounts of untreated and LPS-stimulated murine BMDMs were prepared following the SCoPE2 protocol[5,38], such that the carrier was composed of ∼200 cells and the reference was composed of ∼5 cells. This sample design was then used in the preparation of single-cell sets by nPOP, as well as in the generation of spectral libraries and retention-time-calibration experiments for the experiments shown in Fig. 2 and Fig. 3, as well as Fig. S5 and Fig. S4.[24]

Single-cell samples were prepared using nPOP[24], as detailed in Leduc *et al.* (2021). Single-cell samples from the 24-hr LPS-treated group and from the untreated group were combined within each nPOP set. The majority (87%) of the labeled sets also contained negative control samples for

quality-control purposes. These control samples received all reagents and proceeded through all sample handling steps, but no single cells were dispensed into these droplets. The distribution of protein-level CVs (i.e. quantification variability) associated with the single cell and control samples for these experiments can be found in Fig. S2c.

BMDM single-cell mPOP sample prep

Carrier and reference samples composed of equivalent amounts of untreated and LPS-stimulated murine BMDMs were prepared following the SCoPE2 protocol[5,38], such that the carrier was composed of ∼200 cells and the reference was composed of ∼5 cells. This sample design was then used in the preparation of single-cell sets by mPOP[39], in which single-cells from each condition (untreated and 24-hr LPS treated) were sorted into a 384-well plate (Thermo AB1384) via the cellenONE liquid handling system (Scienion). The mixed carrier and reference sample was also used in the generation of retention-time estimate runs for the set of 10 samples analyzed by pSCoPE.

Endocytosis assay samples

BMDM endocytosis assay

Murine BMDMs were differentiated and divided into treatment groups, as indicated previously, and incubated with dextran conjugated to Alexa Fluor 568 (Thermo, D22912) at a final concentration of 0.5 mg/ml for 45 minutes at 37 °C. After the incubation period, cells were washed twice with 1x PBS and incubated with PBS-2mM EDTA to detach from the plate. Prior to FACS analysis, cells were spun down at 300g for 5 minutes and washed with 1x PBS before being resuspended. Using a Sony MA900, Dextran-AF568 fluorescence in the PE-Texas Red channel was then analyzed for cells from each treatment condition, and a minimum of 70,000 cells from the top and bottom ∼10% of the PE-TR fluorescence distribution were then sorted for downstream sample preparation and mass spectrometry analysis.

Preparation of endocytosis-assay samples for LC-MS/MS analysis

Each FACS-isolated sample was lysed using a freeze/thaw cycle[39]. Post-lysis, approximately 70,000 cells worth of lysate was digested for 12 hours at 37 °C using 11ng/μl of trypsin gold and 150 mM TEAB in 65μl. Samples were then stage-tipped[40], and $\sim$10,000 cells worth of digest was injected in 0.1% formic acid for analysis by mass spectrometry using DIA instrument method 6, detailed below.

MEROPS Bulk validation experiments, BMDMs

Murine BMDMs were differentiated and divided into treatment groups, as indicated previously. Samples initially contained 125,000 BMDMs in 62.5 μl of LC-MS water (Optima LC/MS Grade, Fisher Scientific W6500). SDS (Sigma, L3771-100G) and HEPES (Thermo, Fisher Scientific AAJ63218AE) were added to final concentrations of 1% and 0.1M, respectively. cOmplete Protease inhibitor (Roche, Sigma Aldrich 05892791001) was then added to a 2x final concentration. The samples were then heated to 95 °C for 5 minutes and subsequently chilled at -80 °C for 10 minutes. 1U of benzonase (Millepore, Sigma Aldrich E1014-25KU) was added and allowed to incubate at room temperature for 30 minutes. 500mM DTT (Pierce, Thermo Fisher A39255) was added to a final concentration of 15mM and allowed to incubate for 30 minutes. Iodoacetamide (Pierce, Thermo Fisher A39271) was added to 15mM final concentration and incubated at room temperature in the dark for 30 minutes. DTT was then added a second time to a 15mM final concentration and incubated for 1 hr. SP3 Beads (Cytiva, Fisher Scientific 09-981-123 ; Cytiva, Fisher Scientific 09-981-121) were prepared and mixed as per manufacturer recommendations.

2.5μl of prepared SP3 beads (100 $\mu g/\mu l$) were added to each of the four samples. 17.3μl of LC-MS grade water was added to each tube resulting in a total volume of 141μl. 564μl of ethanol (200 proof, HPLC/spectrophotometric grade, Sigma 459828-1L), was added to each sample and incubated for 18 minutes. Samples were then incubated for 5 minutes on a magnetic stand, the supernatant was removed, and the beads were washed twice with 400μl of 90% ethanol, after which the remaining supernatant was removed.

Each sample was resuspended in 22.5μl 6M GuCl (Sigma, G-3272), 30μl of 0.5M HEPES pH8, and TCEP (10mM final concentration) (Supelco, Millepore Sigma 646547). Samples were then incubated for 30 minutes at room temperature. 57μl of TMTPro (Thermo, A44520) at 8ng/μl was then incubated in each sample for 1.5 hours, with the untreated condition being labeled with 127C and the LPS-treated condition being labeled with 128N. Samples were then quenched with 6μl of 1M TRIS (Thermo Fisher, AM9855G) for 45 minutes. Following quenching, 1.2μl of SP3 beads (100μg/μl) were added to each TMT-labeled sample. 484.4μl of 100% ethanol was added to each sample and allowed to incubate for 15 minutes. Samples were then placed on magnetic stand for 5 minutes, the supernatant was removed, and the beads were washed twice with 600μl of 90% ethanol. The samples were then centrifuged and the remaining liquid was removed.

Samples were resuspended in 100μl to a final concentration of 200mM HEPES and 12 ng/μl of trypsin gold (Promega, V5280). Samples were then placed in a bioshaker (Bulldog Bio, VWR 102407-834) and digested at 37 °C, 200 rpm for 18 hours. After digestion, samples were removed from the bioshaker, briefly sonicated, spun down, vortexed, spun down again, and incubated on a magnetic stand for 5 minutes. The supernatant was then removed and stored at -80 °C. Before analysis by LC-MS/MS, the samples were stage-tipped[40]. Samples were resuspended in 0.1% formic acid at approximately 1 μg worth of digest per μl in glass HPLC inserts (Thermo Fisher C4010-630) prior to analysis, then injected and analyzed via DIA instrument method 4, detailed below (raw files: eGH692-eGH694).

Bulk TMTPro-labeled BMDM samples for differential protein analysis

10,000 cells from each treatment condition (24hr LPS-treated and untreated), resuspended in LC-MS water, were frozen at -80 degrees for 20 minutes, before being lysed at 90 °C in a thermal cycler (BioRad T1000) for 10 minutes. After lysis, benzonase was added to a final concentration of 1U and allowed to incubate for 10 minutes. Trypsin Gold (Promega Trypsin Gold, Mass Spectrometry Grade, PRV5280) was added to a final concentration of 16 ng/μl and triethylammonium bicarbonate (TEAB, Millipore Sigma T7408-100ML) was added to a final concentration of

150mM. The samples were then allowed to digest overnight for 16 hours. Post-digestion, samples were allowed to return to room temperature and were labeled with 85mM TMT 128N (untreated sample) or 85mM TMT 127C (LPS-treated sample). The reaction was then quenched with 0.5μl of 0.5% hydroxylamine (Millipore Sigma 467804-10ML) for 1 hr. Samples were centrifuged briefly to collect liquid following all reagent addition. After labeling, 6000-cells worth of labeled material from each treatment condition were combined in a mass spec insert (Thermo Fisher C4010-630)) and dried down in a speed vacuum (Eppendorf, Germany) before being reconstituted in 3.3ul 0.1% formic acid (Thermo Fisher 85178) and analyzed via DDA bulk BMDM analysis instrument methods 1 and 2. below.

Separate samples containing approximately 1000 cells per injection of the 128N-labeled untreated BMDMs or 127C-labeled 24-hr LPS-treated BMDMS were injected and analyzed via DIA bulk BMDM analysis instrument method 1, below. Proteins which were differentially abundant between the two conditions analyzed via DIA were identified using the process outlined in the Differential protein analysis for DIA samples section, below, and these proteins make up set ζ in the description of the top priority tier composition in the Prioritized inclusion list construction section, also found below.

Spectral-library-generating samples

Prior to performing retention-time-calibration, scout, or prioritized experiments for the nPOP-prepared PDAC and BMDM single-cell samples, spectral libraries were generated by analyzing mixed carrier and reference samples. For the PDAC samples, 1μl injections of a 7.5x concentrated aliquot of the mixed carrier and reference sample were analyzed by DIA instrument methods 1 and 2, and a subsequent 2.5x concentrated aliquot of the mixed carrier and reference sample was injected and analyzed by DIA instrument method 1. For the BMDM samples, 1μl injections of a 5x concentrated aliqout of carrier and reference sample and a 1x concentrated aliquot of carrier and reference sample were sequentially analyzed via DIA Instrument Method 1. Additional in-

formation regarding the instrument methods and search-engine parameters can be found in their respective sections.

BMDM Scout experiment

Prior to assembling an inclusion list for the prioritized analyses of the nPOP-prepared BMDM single-cell samples, a prioritized analysis of a 1x concentrated version of the mixed carrier and reference sample was performed to generate a set of additional DDA-identifiable precursors. Information regarding the inclusion-list construction for this scout experiment, MaxQuant.Live parameters, and analysis of raw data can be found in their respective sections.

Retention-time calibration runs

PDAC samples, Figure 1

A $1\mu l$ injection of a 1x concentrated aliquot of the mixed carrier and reference sample was analyzed via DIA instrument method 1 and searched via Spectronaut with the spectral library generated above (`20210925_234228_wAL00103.raw.kit`, 21,838 precursors) to provide accurate retention times for the subsequent MaxQuant.live-enabled prioritized single-cell analyses. A separate 1x retention-time-calibration experiment was acquired for each set of experiments, Fig1b/e and Fig1c/d. The retention-time-calibration experiment corresponding to Fig1c/d was searched with Spectronaut v. 15.4, while the retention-time-calibration experiment corresponding to F1b/e was searched with v. 15.6.

BMDM nPOP samples, Figures 2/3/4

A $1\mu l$ injection of a 1x concentrated aliquot of mixed carrier and reference sample was injected and analyzed via DIA method 1 and searched via Spectronaut (v. 15.1) with the spectral library generated above (`20210809_120040_Priori_comb_080921.kit`) to provide accurate reten-

tion times for the subsequent MaxQuant.live-enabled prioritized single-cell analyses. All search parameters were kept as default, except for: template correlation profiling enabled for profiling strategy, minimum q-value row selection for profiling row selection, and Biognosys' iRT kit was indicated as not being used.

BMDM mPOP samples

A $1\mu l$ injection of a 1x concentrated aliquot of mixed carrier and reference sample was injected and analyzed via DIA method 3 and searched by Spectronaut (v. 15.0) in directDIA mode using a FASTA containing the SwissProt database for mus musculus, as well as MEROPS cleavage fragments generated as indicated in the MEROPS database preparation section, below (`musmusculus_SPonly_MEROPS_012221.fasta`, 27,117 protein entries). Trypsin was specified as the enzyme for in silico digestion, TMTpro (+ 304.2071 Da) was selected as a fixed modification on lysines, and the following variable mods were used: protein n-terminal acetylation (+ 42.01056 Da), methionine oxidation (+ 15.99492 Da), and TMTPro modification of peptide n-termini. The results were then prefiltered in Spectronaut to only contain precursors with at least one TMTPro modification. All other search settings were kept as default.

Prioritized Inclusion List Construction

PDAC samples, Figure 1b/e

The Spectronaut search results for the retention-time-calibration experiment were then used to generate a prioritized inclusion list in the R programming environment via the data-processing pipeline available at github.com/SlavovLab/pSCoPE. Briefly, the DIA search results were filtered at an elution group PEP ≤ 0.02 and an elution group q-value of ≤ 0.05. The mass of each precursor was then calculated by multiplying its m/z by its charge state and subtracting the mass of the charge-state-associated protons, as MaxQuant.live will later account for their masses. This mass

was then rounded to 7 decimal places. The search results were subsequently filtered for TMTPro-labeled peptides. A numeric place-holder value (e.g., 126) was placed in the "Masses" column to trigger log entries related to species sent for MS2 in the subsequent prioritized analyses. The retention time was then rounded to three decimal places, and 25 minutes were subtracted from the observed retention time to accord with the 25-minute delay in acquisition for the single-cell sample analysis method. This is not a necessary adjustment if MS acquisition is started at minute 0. The library of confidently identified precursor species was then segmented into tertiles by precursor intensity and each tertile was assigned to a priority tier, such that the highest priority tier contained precursors in the top tertile of abundance, the middle priority tier contained precursors in the middle tertile of abundance, and the low priority tier contained precursors in the lowest tertile of abundance. The RealTimeCorrection and TargetedMS2 column entries were then set to TRUE for this set of confidently identified precursor species to enable participation in MaxQuant.Live's real-time retention-time alignment algorithm, as well as MS2 analysis upon detection. Any identified precursors not matching the original confidence cut-offs (elution group PEP ≤ 0.02 or elution group q-value ≤ 0.05) were enabled for participation in MaxQuant.live's real-time retention-time alignment algorithm, but were prohibited from being sent for MS2.

PDAC samples, Figure 1c/d

The search results from the retention-time-calibration experiment were then filtered for use as an inclusion list via the data-processing pipeline available at github.com/SlavovLab/pSCoPE. Briefly, a set of difficult-to-identify precursor ions was selected by filtering the shotgun search results for peptides identified in 50% of the experiments or fewer at PEP $\leq$ .02. These peptides were then assigned to the top two priority tiers of an inclusion list such that each tier contained an equivalently difficult set of peptides to identify. The remaining high-confidence peptides from the shotgun results were placed on the middle priority tier, and high-confidence peptides solely identified via the DIA retention-time alignment run were placed on the low-priority tier. Precursor

identifications from the DIA retention-time-generating run with PEPs $\geq$ .02 were then selected for inclusion as precursor ions whose retention times could aid MaxQuant.live's real-time retention-time alignment algorithm, without being sent for MS2.

Scout experiment for BMDM nPOP-samples, Figures 2/3/4

The search results from the retention-time-calibration experiment were then filtered to meet the following criteria for use as an inclusion list: elution group PEP $\leq$ .02, elution group q-value $\leq$ .05, TMTPro labeling modifications (+ 304.2071 Da) on the n-terminus or lysine residues. These filtered precursors were then stratified into priority tiers by precursor intensity: precursors in the lowest intensity tertile were place on the bottom priority tier, precursors in the middle intensity tertile were placed on the middle priority tier, and precursors in the top intensity tertile were placed on the top priority tier. These precursors were enabled for both MS2 analysis upon detection and for participation in MaxQuant.Live's real-time-alignment algorithm. Precursors whose confidence of identification or labeling state did not pass the initial thresholds were included for participation in MaxQuant.live's retention-time alignment algorithm, but were not enabled for MS2 analysis. The subsequent prioritized inclusion list was used to analyze a $1\mu l$ injection of a 1x concentrated aliquot of carrier and reference sample via MaxQuant.live.

BMDM nPOP samples, Figures 2/3/4

The prioritized inclusion list for the BMDM samples was constructed by importing the search results from the DIA retention-time-calibration run into the R environment and subsetting the detected peptides into 4 tiers. Peptides and proteins of special biological interest or experimental utility were assembled from the following sources: all precursors identified at or below a PEP of 0.05 in the scout experiment (set α); proteins significantly correlated to PC1 (5% FDR) in a cross-condition PCA generated from the 20 initial shotgun analyses of single-cell BMDM samples (set β); proteins significantly correlated to PC1 (5% FDR) in a PCA of the LPS-treated single cells generated from the 20 initial shotgun analyses of BMDM samples (set γ); proteins significantly

correlated to PC1 (5% FDR) in a PCA of the untreated single cells generated from 20 initial shotgun analyses of BMDM samples (set δ); proteins with $|log_2$(fold changes)$| \geq 1$ between the low and high dextran uptake conditions of each treatment group analyzed by DIA which were found to be statistically significant, (set ϵ, statistical process described in differential protein analysis for DIA samples section, below); proteins with $|log_2$(fold changes)$| \geq 1$ between LPS-treated and untreated bulk BMDM samples analyzed via DIA which were found to be statistically significant, (set ζ, statistical process described in differential protein analysis for DIA samples section); proteins significantly correlated to PC1 (5% FDR, |Spearman correlation| > 0.35) from single-condition PCAs generated from prior mPOP-prepared single-cell BMDM analyses (set η); all precursors identified at 1% FDR in the 20 initial shotgun analyses of the single-cell BMDM samples (set θ); precursors identified in the retention-time-calibration run that were also contained in sets β, γ, and δ and identified in fewer than 50% of the corresponding shotgun analyses of single-cell BMDMs (set ι). To construct the top priority tier, peptides from the retention-time-estimation run were selected to correspond to: up to the top 125 most abundant precursors from set ι; the top 35 most abundant MEROPS-annotated precursors; all precursors in set θ that mapped to proteins in set ζ; up to the top 100 most abundant precursors that were in common between set θ and precursors derived from proteins in set η; up to the top 100 most abundant precursors in common between set θ and precursors derived from proteins in set ϵ; up to the top 5 most abundant precursors per protein in the intersection between set θ and precursors derived from proteins in set δ; up to the top 6 most abundant precursors per protein in the intersection between set θ and precursors derived from proteins in set γ; up to the top 4 most abundant precursors per protein in the intersection between set θ and precursors derived from proteins in set β.

The middle priority tier was composed of all remaining precursors from the set of peptides identified at 1% FDR in the accompanying single-cell shotgun analyses, set θ. The bottom tier was composed of all remaining peptides identified in the scout experiment, set α. A subsequent tier of peptides used only for retention-time alignment (i.e., not sent for MS2 analysis) was composed of all remaining peptides identified by the retention-time-calibration experiment.

Regarding precursor-intensity-dependent fill times, if a top tier precursor appeared in the bottom intensity tertile, it was allotted an MS2 fill time of 1000ms; if a top tier precursor appeared in the middle intensity tertile, it was allotted an MS2 fill time of 750ms; if a top-tier precursor appeared in the top intensity tertile, it was allotted an MS2 fill time of 500ms. These intensity tertiles were calculated across all filtered PSMs from the retention-time-calibration run.

mPOP-prepared BMDM troubleshooting samples

The search results from the DIA retention-time-calibration experiment were first filtered such that all remaining entries had an elution group PEP $< .05$ and an elution group Q value $< .05$., as well as at least one TMTPro modification (+ 304.2071 Da) on either the peptide n-terminus or lysine residue. The prioritized inclusion list for the mPOP BMDM samples was constructed such that the top tier contained the following types of precursors which intersected with the retention-time-calibration experiment identifications: precursors identified in less than 50% of the corresponding 13 shotgun experiments, precursors featuring a MEROPS-annotated cleavage site, precursors mapping to proteins of biological interest (TLRs,interleukin-associated proteins, lyosomal-associated membrane proteins, interferon-associated proteins, caspases, NF-kappaB-associated proteins, transcription factors, gasdermin, signal transducers, macrophage scavenger receptors, and proteins annotated to macrophage function). Precursors on the top priority tier were allocated fill-times dependent upon their precursor intensities in the following manner: precursors in the top abundance tertile were allocated a 600ms MS2 fill time, precursors in the middle abundance tertile were allocated a 750ms MS2 fill time, and precursors in the bottom abundance tertile were allocated a 900ms MS2 fill time. Precursors identified in the retention-time-calibration experiment that were part of a previous targeting experiment were placed on the middle priority tier along with the precursors that had been identified in the corresponding mPOP shotgun experiments. The bottom priority tier, although redundant to the top and middle tiers in composition, served to keep the instrument duty cycles full for optimal elution peak sampling. All precursors enabled for MS2, as well as all remaining precursors identified in the filtered retention-time-calibration run were enabled for

participation in MaxQuant.Live's real-time retention-time calibration algorithm.

Analysis of raw MS data

MEROPS database preparation and FASTA modification

The MEROPS database was downloaded from https://www.ebi.ac.uk/merops/download[26] and converted into a .csv for import into the R environment. All cleavage patterns consistent with trypsin (e.g., R or K as the P1 residue), were removed from the database. Then the SwissProt-annotated Mus musculus FASTA file was read into the R environment, and the sequence for each protein with an annotated MEROPS cleavage site was split between the P1 and P1' residues. Both halves of the MEROPS-cleaved peptide were then subjected to an in silico tryptic digestion such that the tryptic digest produced a fragment at least 6 amino acids long. The two semi-tryptic peptide halves were then added to the existing FASTA as separate entries, including annotations from the MEROPS database for the enzyme, cleavage residue number, and whether the peptide fragment contained the neo-c-terminus or neo-n-terminus.

DDA Data

Shotgun analyses of PDAC single-cell samples, Figure 1

Shotgun analyses of the PDAC samples were searched with MaxQuant (2.0.3.0) using a FASTA containing all entries from the human SwissProt database (`SwissProt_human_09042017`, 20,218 proteins). TMTPro 16plex was enabled as a fixed modification on n-termini and lysines via the reporter ion MS2 submenu. Methionine oxidation (+ 15.99492 Da) and protein n-terminal acetylation (+ 42.01056 Da) were enabled as variable modifications, and trypsin was selected for the in silico digestion with enzyme mode set to specific. Up to 2 missed cleavages were allowed per peptide with a minimum length of 7 amino acids. Second peptide identifications were disabled, calculate peak properties was enabled, and msScans was enabled as an output file. PSM FDR

and protein FDR were set to 1. False discovery rate (FDR) calculations were performed in the R programming environment by calculating the PEP threshold at which 1% of the entries were decoy identifications.

pSCoPE analyses of PDAC single-cell samples, Figure 1

In both data sets, the prioritized runs corresponding to fig1b/e and fig1c/d, the same search settings were used as in the accompanying shotgun data sets, with the exception of the FASTA database. In the case of the prioritized searches, reduced versions of the human SwissProt database were used which contained only those proteins whose peptides were on the inclusion list. `SwissProt_human_09042017_PDAC_lim2.fasta` (2,351 proteins) was used as the reference database for the prioritized data set corresponding to figure 1b/e and `sp_human_20211005_Consist_lim.fasta` (1,574 proteins) was used as the reference database for the prioritized data set corresponding to figure 1c/d. False discovery rate (FDR) calculations were performed in the R programming environment by calculating the PEP threshold at which 1% of the entries were decoy identifications.

TMT-labeled bulk BMDM sample analyses for differential protein analysis

TMT-labeled and mixed bulk samples of 24-hr LPS-treated and untreated BMDMs analyzed via DDA Bulk Instrument Methods 1 and 2 (wGH215 and wGH216, respectively) were searched with MaxQuant (v. 1.6.17.0) using a FASTA containing all entries from the SwissProt database for mus musculus, as well as MEROPS-annotated cleavage products generated as indicated previously (`musmusculus_SPonly_MEROPS_012221.fasta`), for a total of 27,117 protein entries. TMTPro 16plex was enabled as a fixed modification on n-termini and lysines via the reporter ion MS2 submenu. Methionine oxidation (+ 15.99492 Da) and protein n-terminal acetylation (+ 42.01056 Da) were enabled as variable modifications, and trypsin was selected with specific cleavage. Second peptide identifications were disabled, calculate peak properties was enabled, and msScans was enabled as an output. PSM FDR and protein FDR were set to 1.

Scout experiment for inclusion-list generation for BMDM nPOP samples

The raw file generated by this prioritized analysis was searched by MaxQuant (v 1.6.17.0) using a FASTA containing all entries from the murine SwissProt database (`musmusculus_SPonly_012221.fasta`, 17,056 proteins). TMTPro 16plex was enabled as a fixed modification on n-termini and lysines via the reporter ion MS2 submenu. Methionine oxidation (+ 15.99492 Da) and protein n-terminal acetylation (+ 42.01056 Da) were enabled as variable modifications, and trypsin was selected with specific cleavage. Second peptide identifications were disabled, calculate peak properties was enabled, and msScans was enabled as an output. PSM FDR and protein FDR were set to 1.

Shotgun analyses of BMDM nPOP samples, Figure 2/3/4

Shotgun analyses of the nPoP-prepared murine BMDM samples were searched with MaxQuant (2.0.3.0) using a FASTA containing all entries from the murine SwissProt database with additional entries for cleaved peptides consistent with the MEROPS database appended (`musmusculus_SPonly_MEROPS_012221.fasta`), for a total of 27,117 protein entries. Methionine oxidation (+ 15.99492 Da) and protein n-terminal acetylation (+ 42.01056 Da) were enabled as variable modifications, and trypsin was selected for the in silico digestion with enzyme mode set to specific. Up to 2 missed cleavages were allowed per peptide with a minimum length of 7 amino acids. Second peptide identifications were disabled, calculate peak properties was enabled, and msScans was enabled as an output file. PSM FDR and protein FDR were set to 1.

pSCoPE analyses of BMDM nPOP samples, Figure 2/3/4

The same search settings were used as in the accompanying shotgun data sets, with the exception of the FASTA database. For the prioritized samples, a reduced version of the murine SwissProt database with appended MEROPS entries was used which contained only those proteins whose peptides were on the inclusion list (`musmusculus_SPonly_MEROPS_012221_lim2.fasta`; 1,234 proteins). Subsequent to the MaxQuant search, the 20 shotgun-analyzed nPOP-prepared

SCoPE experiments, 40 pSCoPE-analyzed nPOP-prepared SCoPE experiments, and their preceding DIA-analyzed retention-time-calibration experiment were analyzed together by DART-ID[41] for retention-time-dependent PSM confidence updating. A DART-ID configuration file is included in the Massive repository associated with this publication.

Shotgun and pSCoPE analyses of BMDM mPOP troubleshooting samples

Shotgun and pSCoPE analyses of the murine BMDM single-cell samples prepared by mPOP were searched with MaxQuant (1.6.7.0) using a FASTA containing all entries from the murine SwissProt database with additional entries for cleaved peptides consistent with the MEROPS database appended (`musmusculus_SPonly_MEROPS_012221.fasta`), for a total of 27,117 protein entries. Methionine oxidation (+ 15.99492 Da) and protein n-terminal acetylation (+ 42.01056 Da) were enabled as variable modifications, and trypsin/P was selected for the in silico digestion with enzyme mode set to specific. Up to 2 missed cleavages were allowed per peptide with a minimum length of 7 amino acids. Second peptide identifications were disabled, calculate peak properties was enabled, and msScans was enabled as an output file. PSM FDR and protein FDR were set to 1.

DIA Data

PDAC samples, spectral library-generating search

$1\mu l$ injections of a 7.5x concentrated aliquot of the mixed carrier and reference sample were analyzed by DIA instrument methods 1 and 2, and a subsequent 2.5x concentrated aliquot of the mixed carrier and reference sample was injected and analyzed by DIA instrument method 1. These three sample analyses were used to construct a spectral library via Spectronaut (v. 15.6) for the analysis of a 1x mixed carrier and reference sample analyzed via DIA instrument method 1 (i.e., a retention-time-calibration experiment). The spectral library contained a total of 21,838 precursors, and was generated from a directDIA search using the `SwissProt_human_09042017.fasta` (20,218 proteins). Default search parameters were used, with the following exceptions: source specific

iRT calibration was enabled; Biognosys iRT kit was set to false; profiling strategy was set to template correlation profiling; Methionine oxidation (+ 15.99492 Da), protein n-terminal acetylation (+ 42.01056 Da) and n-terminal TMTPro labeling (+ 304.2071 Da) were enabled as variable modifications, while TMTPro labeling of lysines was enabled as a fixed modification. The spectral library produced by this search was named `20210925_234228_wAL00103.raw.kit`.

BMDM samples, spectral library-generating search

$1\mu l$ injections of both a 5x concentrated aliquot and a 1x concentrated aliquot of mixed carrier and reference samples were sequentially analyzed via DIA Instrument Method 1. The two sample analyses indicated above were used to construct a spectral library using Spectronaut (v. 15.1) via a directDIA search using the `musmusculus_SPonly_MEROPS_012221.fasta` (27,117 proteins). Methionine oxidation (+ 15.99492 Da), n-terminal acetylation (+ 42.01056 Da), and n-terminal TMTPro labeling (+ 304.2071 Da) were enabled as variable modifications, while TMTPro labeling of lysines was enabled as a fixed modification. The resulting spectral library contained a total of 11,701 precursors. Default search parameters were used, with the following exceptions: allow source specific iRT calibration was enabled, Biognosys' iRT-kit alignment peptides were specified as unused, and profiling strategy was set to template correlation profiling. The spectral library produced by this search was named
`20210809_120040_Priori_comb_080921.kit`.

PDAC samples, retention-time-calibration experiment, Figure 1b/e

Raw data was searched via Spectronaut (v. 15.6) with the PDAC-specific spectral library (`20210925_234228_wAL00103.raw.kit`, 21,838 precursors) discussed above to provide accurate retention times for the subsequent MaxQuant.live-enabled prioritized single-cell analyses. The reference FASTA for this spectral library was `SwissProt_human_09042017.fasta`

(20,218 proteins). The iRT-kit alignment peptides were specified as unused, template correlation profiling was selected as the profiling strategy, minimum q-value row selection was selected for the profiling row selection method, and allow source specific iRT Calibration was set to true. All other options kept as default

PDAC samples, retention-time-calibration experiment, Figure 1c/d

Raw data was searched via Spectronaut (v. 15.4) with the PDAC-specific spectral library (`20210925_234228_wAL00103.raw.kit`, 21,838 precursors) discussed above to provide accurate retention times for the subsequent MaxQuant.live-enabled prioritized analyses. The reference FASTA for this spectral library was `SwissProt_human_09042017.fasta` (20,218 proteins). The iRT-kit alignment peptides were specified as unused and allow source specific iRT Calibration was set to true. All other options were kept as default.

BMDM samples, pre-scout retention-time-calibration experiment, Figures 2/3/4

Raw data was searched via Spectronaut (v. 15.1) with the BMDM-specific spectral library: `20210809_120040_Priori_comb_080921.kit`. The reference FASTA for this spectral library was `musmusculus_SPonly_MEROPS_012221.fasta` (27,117 proteins). The iRT-kit alignment peptides were specified as unused, template correlation profiling was selected as the profiling strategy, minimum q-value row selection was selected for the profiling row selection method, allow source specific iRT calibration was set to true. All other options kept as default

BMDM nPOP samples, retention-time-calibration experiment, Figures 2/3/4

Raw data was searched via Spectronaut (v. 15.1) with the following spectral library: `20210809_120040_Priori_comb_080921.kit`. The reference FASTA for this spectral library was `musmusculus_SPonly_MEROPS_012221.fasta` (27,117 proteins). The iRT-kit alignment peptides were specified as unused, template correlation profiling was selected as the profiling strategy, minimum q-value row selection was selected for the profiling row selection

method, “allow source specific iRT calibration” was set to true. All other options kept as default

TMTPro-labeled bulk BMDM samples for differential protein analysis

TMT-labeled, unmixed bulk samples (wGH217/218/220) analyzed by DIA Instrument Method 3 were searched with Spectronaut (v. 14.1) in directDIA mode using the `musmusculus_SPonly_MEROPS_012221.fasta` containing 27,117 proteins. Methionine oxidation (+ 15.99492 Da), n-terminal acetylation (+ 42.01056 Da), and n-terminal TMTPro labeling (+ 304.2071 Da) were enabled as variable modifications, while TMTPro labeling of lysines was enabled as a fixed modification. Trypsin was specified as the enzyme for in silico digestion, and results were filtered to contain only precursors with TMTPro labeling modifications.

TMTPro-labeled bulk MEROPS validation samples for BMDM analysis, Figure 5

The Raw files from the MEROPS analyses were then searched with Spectronaut’s (v. 15.4) direct-DIA analysis feature, using a FASTA containing all entries from the SwissProt database for mus musculus, as well as MEROPS cleavage fragments generated as indicated previously, which contained 27,117 protein entries (`musmusculus_SPonly_MEROPS_012221.fasta`). Cysteine carbamidomethylation was set as a fixed modification and the following variable modifications were used: protein n-terminal acetylation (+ 42.01056 Da), methionine oxidation (+ 15.99492 Da), TMTPro modification (+ 304.2071 Da) of lysine and peptide n-termini. Trypsin enzymatic cleavage rules were enabled, allowing for a minimum peptide length of 7 and a maximum peptide length of 52. Up to 2 missed cleavages were allowed. All other search settings were left at their default values.

Label-free bulk endocytosis samples for BMDM analysis, Figure 4

Raw data from the bulk endocytosis sample analyses was searched with Spectronaut (v. 14.10) via DirectDIA, using a FASTA containing all entries from the SwissProt database for mus musculus, as well as selected isoforms and MEROPS-annotated cleavage products generated as indicated

previously, for a total of 33,996 protein entries (`Mouse_ONLYsp_plusMEROPS_v2.fasta`). Peptides with lengths between 6 and 52 amino acids, with up to 2 missed cleavages were permitted. Trypsin/P was selected for cleavage, and Protein n-term acetylation (+ 42.01056 Da) and methionine oxidation (+ 15.99492 Da) were enabled as variable modifications. A PEP Cut-off of 1 was selected, although downstream filtration (PEP $\leq$ .01) was performed in the differential protein analysis script. All other search settings were kept as default.

BMDM mPOP samples retention-time calibration experiment

The pre-prioritization retention-time-calibration experiment was searched by Spectronaut (v. 15.0) in directDIA mode using a FASTA containing the SwissProt database for mus musculus, as well as MEROPS cleavage fragments generated as indicated previously, containing 27,117 protein entries (`musmusculus_SPonly_MEROPS_012221`). Trypsin was specified as the enzyme for in silico digestion, TMTpro (+ 304.2071 Da) was selected as a fixed modification on lysines, and the following variable mods were used: protein n-terminal acetylation (+ 42.01056 Da), methionine oxidation (+ 15.99492 Da), and TMTPro modification of peptide n-termini. The results were then prefiltered in Spectronaut to only contain precursors with at least one TMTPro modification. All other search settings were kept as default.

Processing and normalizing single-cell MS data

Shotgun and pSCoPE PDAC analyses, Figure 1b/e and Figure 1c/d

Sets pertaining to Figure 1b/e and Figure 1c/d were processed separately. Single-cell MS data were processed via the SCoPE2 single-cell proteomics pipeline[5,42]. Peptides with precursor ion fractions below 50% or a mean RI intensity across the single cells greater than 10% of the intensity in the carrier channel were removed from the data set. Peptides were filtered at 1% FDR by determining the PEP threshold at which 1% of the entries were reverse matches. Cells with mean protein CVs greater than 0.4 were filtered out from the data set. Samples and precursors were then

filtered to have less than 99% missing data, before being log transformed and aggregated to protein-level abundance by taking the median abundance of the protein-specific peptides. The following intermediate data frames were generated for subsequent analysis from both data sets: the matrix of single cells by precursors, unfiltered for missingness; the matrix of unimputed protein abundances by single-cells, containing missing values. For the sets corresponding to Figure 1b/e, the following additional data matrices were produced: the complete matrix of protein abundances by single cells, containing imputed values; the batch-corrected complete matrix of protein abundances by single cells, containing imputed values; the re-normalized batch-corrected complete matrix of protein abundances by single cells, containing imputed values.

pSCoPE analyses of BMDM samples, Figures 2/3/4

Single-cell data from 40 prioritized analyses were processed via the single-cell pipeline[5,42]. Peptides with precursor ion fractions below 50% were removed from the data set, as were peptides with a mean intensity across the single cells greater than 2% of the intensity in carrier channel. Peptides were filtered at 1% FDR using the DART-ID[41] q-value column. Cells with mean protein CVs greater than 0.4 were filtered out from the data set. Samples and precursors were then filtered to have less than 99% missing data, before being log transformed and aggregated to protein-level abundance by taking the median abundance of the protein-specific peptides. The following intermediate data frames were generated for subsequent analysis: the matrix of single cells by precursors, unfiltered for missingness; the matrix of unimputed protein abundances by single-cells, containing missing values; the complete matrix of protein abundances by single cells, containing imputed values; the batch-corrected complete matrix of protein abundances by single cells, containing imputed values; the re-normalized batch-corrected complete matrix of protein abundances by single cells, containing imputed values.

Shotgun and pSCoPE analyses of BMDM samples, Figure S4

Single-cell data from 20 shotgun analyses and the first 20 prioritized analyses were processed via the single-cell pipeline[5,42] using the same parameters as indicated above for the pSCoPE-only BMDM analyses.

Data Analysis

Differential protein analysis for DIA samples

Differential protein abundance was assessed by modeling the distribution of noise as a function of average precursor intensity. To perform this analysis, a single sample was injected and analyzed twice by DIA, then searched with Spectronaut as described above. Precursor-level fold-changes between replicate injections of the same sample should cluster around 1:1, deviations from this expected 1:1 ratio reflect noise in the measurement, and can be used as a null distribution to test for differential protein abundance. However, because precursor quantitation is more accurate at higher absolute intensities, the null distribution of precursor fold-changes were split evenly into 15 bins with respect to the average precursor intensity of the pair. Fold-changes between experimental conditions were then calculated, and converted to a z-score using its corresponding null distribution of fold-changes (based on intensity). Lastly, the precursors for each protein were t-tested against a standard-normal null distribution of 10,000 values with mean = 0, and standard deviation = 1, then p-values were converted to q values using the Benjamini and Hochberg approach.

Figure 1, PDAC samples

Productive MS2 scans boxplot

The msms.txt files from the shotgun and pSCoPE MaxQuant search results were filtered at 1% FDR, by determining the PEP threshold at which 1% of the entries were reverse matches. All contaminant and reverse matches were then removed from the resulting data frames. The number

of remaining PSMs was tallied and divided by the total number of MS2 scans recorded per experiment, determined from the msmsScans.txt output of MaxQuant. This fraction was multiplied by 100 and presented in Fig. 1b as the percentage of productive MS2 scans per experiment.

Peptides/run boxplot

The evidence.txt files from the shotgun and pSCoPE MaxQuant search results were filtered at 1% FDR, by determining the PEP threshold at which 1% of the entries were reverse matches. All contaminant and reverse matches were then removed from the resulting data frames, and peptides with multiple charge states were collapsed to a single entry per experiment. The number of PSMs remaining was then tallied on a per experiment basis and presented as a boxplot in Fig. 1b.

Proteins/cell boxplot

Using the matrix of unimputed protein abundances by single-cell samples produced by the single-cell pipeline[5,42], the number of proteins per single-cell sample with detectable reporter ion intensities was tallied and presented in Fig. 1b.

Consistency heatmap

The precursors by single-cells matrix produced after PIF, mrri, FDR, and CV filtration was subsetted to contain only those precursors on the top two tiers of the consistency experiment inclusion list, which contained precursors identified in 50% or fewer of the shotgun experiments. The precursors were then aggregated to the peptide level, and their abundance measurements were binarized such that precursors with NA intensities for a single-cell sample were given a value of zero while precursors with detected reporter ion intensity were given a value of 1. The resulting binary dataframe was then presented as a heatmap with peptides on the y-axis and single cells on the x-axis to contrast data completeness for difficult-to-identify peptides placed on the top two tiers of a prioritized inclusion list. The set of peptides was filtered to only include entries with at least one cell with detected reporter ion intensity, as shown in Fig. 1c.

Peptide and protein-level data completeness contrast box plots

To produce the peptide-level data-completeness boxplot shown in Fig. 1d , the precursors by single-cells matrix produced after PIF, mrri, FDR, and CV filtration was subsetted into two dataframes: one containing precursors from the top two priority tiers of the prioritized inclusion list and one containing precursors that were enabled for MS2 analysis from any priority tier. Precursor abundances were aggregated to the peptide level by summing the relative intensities observed for multiple charge states on a per-sample basis; this sum was not used for relative quantitation, only for representing whether detectable reporter ion signal was observed for a peptide in a single-cell sample, as opposed to an NA value. The fraction of peptides with detected reporter ion signal per single cell was computed for these two priority categories before being multiplied by 100 to produce the percent data completeness per cell.

To generate the protein-level data completeness box plot shown in Fig. 1d, the matrix of protein abundances by single cells produced by the scp pipeline[5,42] was subsetted to include only the proteins whose precursors were specified for MS2 analysis on the prioritized inclusion list. The fraction of these proteins with detected reporter ion signal per single cell was computed and then multiplied by 100 to produce the percent data completeness per cell.

PDAC PCA color-coded by median protein-set abundance

PCA was performed on the imputed, batch-corrected, and normalized proteins by cells matrix, using the prcomp function in R. For PCAs color-coded by the median protein set abundance, the median abundance of all proteins mapping to a protein set was calculated on a per-cell basis, using the unimputed batch-corrected data matrix. The vector of single-cell protein set abundances was then z-scored, and the resulting vector was joined with the vector of principal component coordinates (the scores vectors) by sample ID. The protein sets presented in this analysis were selected from the results of PC-weight-based Protein Set Enrichment Analysis (PSEA), as described below.

Figures 2/3/4, BMDM samples

PCAs color-coded by median protein-set abundance

For PCAs color-coded by the median protein set abundance(Fig. 2a,Fig. 3a/c,Fig. S5), the procedure indicated in the preceding section was followed.

Protein set enrichment analysis (PSEA)

PSEA was performed using the vector of principal-component-associated protein weights produced by PCA analysis of the imputed, batch-corrected, and normalized cells x proteins matrix, generated from the prcomp function in R. The human gene set database was acquired from GOA.[43] The gene set database was filtered to remove entries corresponding to cellular components, in favor of entries annotated to molecular function and biological process. Gene-level annotation was used to map gene sets to the protein weights. The factor weights for all proteins matching a protein set were compared against the background distribution of protein weights using the two-tailed wilcoxon rank-based test of significance. The following filters were used to determine whether a statistical comparison was made: At least 5 proteins from a protein set must have been present in the data, at least 10% of the proteins within a protein set must have been present in the data, and protein sets must contain fewer than 200 entries. The median loading per protein set was then transformed to a z-score for interpretability. The p values were then converted to q values using the Benjamini and Hochberg approach, and results were filtered to 5% FDR. Single-condition PSEA: the PC-based PSEA performed in Fig. 2b was applied to each treatment group separately in Fig. 3a. For the PSEA performed on the PDAC samples shown in Fig. 1e, the minimum protein count was raised to 15, due to the higher protein coverage present in that set of experiments.

Endocytosis analysis, histogram, Figure 3b

The vectors of Dextran: AF568 MFI and event counts were retrieved from the Sony MA900 used to sort the dextran uptake subpopulations, and this data was filtered for MFIs greater than 1,000

and less than 50,000. The MFIs were then log_{10} transformed and plotted as a normalized histogram in Fig. 3b and for the LPS-treated samples and untreated samples, respectively.

Endocytosis analysis, volcano plot, Figure 3b

Differentially abundant proteins between the low and high-dextran uptake samples were identified via the DIA differential protein analysis script introduced earlier, and the results were plotted, such that proteins with $|\log_2(\text{FC})| > 3$ were annotated. The volcano plots associated with the LPS-treated samples and untreated samples are presented in Fig. 3c and Fig. S8, respectively.

Endocytosis analysis, PCA color-coded by endocytic proteins, Figure 3c

For each treatment condition (24hr LPS-treated or untreated), the set of proteins with statistically significant fold changes between the high and low dextran-uptake conditions ($|\log_2(\text{FC})| \geq 1$; fold-change q-value ≤ 0.01) was intersected with the set of quantified proteins for the respective set of single-cell samples. The median abundance per cell was calculated for the sets of proteins associated with low dextran uptake or high dextran uptake, each vector of median abundances was then z-scored and extreme values capped at a z-score of ± 2, and the single-condition PCAs were color-coded by these z-scored protein abundances. The figures associated with the LPS-treated samples and untreated samples are presented in Fig. 3c and Fig. S8, respectively.

Validation of MEROPS peptide quantitation, Figure 4a

MEROPS substrates which were n-terminally labeled with TMTPro in the bulk discovery experiments and whose cross-condition fold change was comparable between the bulk and single-cell experiments were taken to be validated measurements.

To benchmark the relative quantitation between the bulk discovery and single-cell samples, the treatment-condition-associated bulk samples detailed in the ”MEROPS Bulk validation experiments, BMDMs” section, above, were filtered to an elution group PEP ≤ 0.01, and entries

mapping to the same precursor species were condensed such that the observation with the highest intensity was taken to be representative. The LPS-treated and untreated samples were then joined by precursor, and the abundances column (sample) and row (precursor) normalized, by median and mean, respectively. The ratio of the normalized precursor abundances between the LPS-treated and untreated samples were then calculated. The matrix of batch-corrected unimputed protein abundances per single-cell from the prioritized BMDM analyses was then condensed to a representative abundance by treatment condition by taking the median protein abundance across all cells from a treatment group. The relative protein abundance ratio between treatment conditions was then computed, and the vector of fold changes was subset for the MEROPS cleavage products detected with TMT-Pro labeling of the neo-n-termini in the bulk experiments.

Biological annotation of MEROPS peptides, Figure 4b

Using the MaxQuant evidence.txt output from the DDA analysis of the bulk TMTPro-labeled duplex sample containing 24-hour LPS-treated (127C) and untreated (128N) BMDM samples (wGH215.raw), proteins which were differentially abundant between treatment conditions were identified in the following way: search results were filtered to contain precursors with PEPs ≤ 0.02 and PIFs > 0.8, and reverse matches and contaminants were filtered out; the reporter ion intensities for the two samples were column and row normalized by their means; the per-protein distributions of relative precursor abundances for each sample were subjected to a two-sided wilcoxon rank sum test; p-values were FDR corrected via the Benjamini and Hochberg approach and filtered to 1% FDR; the median relative abundance of the precursors mapping to a given protein were taken to reflect the relative abundance of that protein; differential proteins whose relative abundance ratio (LPS-treated:Untreated) was > 1 were annotated as marker proteins associated with LPS-treatment, or the untreated condition, otherwise.

A second set of marker proteins associated with pro-inflammatory M1-like macrophages or anti-inflammatory M2-like macrophages determined by transcriptomic analysis of monocytes, intermediate macrophages, fully differentiated macrophages, classically activated macrophages, and

alternatively activated macrophages was also used in this analysis[32]. Genes with a $\log_2$ M1-to-M2 ratio greater than zero in the publication-associated database were annotated to be M1-associated, while genes with an M1-to-M2 ratio less than zero were annotated to be M2-associated.

Actin L104 cleaved by cathepsin (fragment 1), citrate synthase H26 cleaved by cathepsin E (fragment 2), and Actin L288 cleaved by Cathepsin D (fragments 1 and 2) which had been validated via bulk analysis were then tested for significant associations with either the treatment-condition-associated protein panels or the M1/M2-associated protein panels using a permutation test.

The matrix of single cells by batch-corrected unimputed protein abundances was filtered to contain the four MEROPS cleavage products and proteins annotated to the supplied list of marker proteins (either treatment-condition specific or macrophage-polarization specific), and a protein-protein correlation matrix was produced from this filtered matrix. The median correlation was then calculated between each MEROPS cleavage product and the set of proteins annotated to either of the two reference conditions (LPS-treated or untreated; M2 or M1), and the difference between the median correlation associated with each treatment condition was recorded.

The same procedure was then repeated 10,000 times, permuting the column names of the cells by proteins matrix each time. The p-value of the original correlation distance was subsequently determined to be the fraction of times a correlation distance as extreme as the one initially observed was generated by chance alone. The set of p-values was then FDR corrected using the Benjamini and Hochberg approach. If the original p value was zero, meaning no value generated by chance was as extreme as the initially observed value, then a q-value of 10^{-5} was used.

Supplemental Figures

Fraction of inclusion-list precursors detected and analyzed, Figure S1

The MaxQuant.Live log files associated with the technical coverage and consistency experiments Fig. 1b/c/d were imported into the R environment and the lists of precursors detected by MaxQuant.live during the survey scan and subsequently sent for MS2 were extracted from the log files. The unique numeric precursor id was then matched to the associated inclusion list for each experiment to gen-

erate Fig. S1.

Quantification variability across single-cell and control samples, Figure S2

Within the single-cell pipeline[5,42], the CV (i.e. the standard deviation scaled by the mean) was computed for the relative abundances of all filtered precursors that mapped to a given leading razor protein on a per-sample basis (precursor-filtration metrics discussed in detail in the data filtration and normalization sections). The mean protein-level CV was then calculated on a per-sample basis, and a CV threshold was chosen which well separated the control samples from the single-cell samples. The distribution of CVs for single-cell and control samples associated with Fig. 1b are shown in Fig. S2a; the distribution of CVs for single-cell and control samples associated with Fig. 1c/d are shown in Fig. S2b; the distribution of CVs for single-cell and control samples associated with Fig. 2, Fig. 3, and Fig. 4 are shown in Fig. S2a

Peptide Properties and ID rates, Figure S3

The shotgun and prioritized search results for the technical consistency experiments Fig. 1c/d were imported into the R environment, and the set of precursors not identified at 1% FDR in the prioritized analyses was determined. The median spectral confidence of identification and number of matching fragments for each of the precursors in this set was then calculated across all shotgun experiments using the evidence.txt and msms.txt files, respectively, and plotted in Fig. S3.

BMDM technical comparison, Figure S4

The matrix of precursor abundances by single-cell samples for 20 shotgun analyses and 20 pSCoPE analyses was condensed to the peptide level by summing the relative intensities across charge states on a sample-specific basis. This was conducted as a means to determine peptides without detectable reporter ion signal in a given sample. The inclusion list was also condensed to the peptide level by associating a peptide sequence with the highest priority tier it appeared in. The fraction of peptides with detected reporter ion signal was calculated on a per sample and per priority-tier basis

and displayed in Fig. S4a. To calculate the percent data completeness on a per-protein level, the same procedure was followed for condensing precursors to proteins and for associating proteins with priority tiers.

To calculate the number of peptides with detectable reporter ion signal per single cell, the matrix of precursor abundances by single-cell samples for 20 shotgun analyses and 20 pSCoPE analyses was condensed to the peptide level as performed previously, and the number of non-NA values was tallied on a per-single-cell basis. The number of proteins with detectable reporter ion signal per single cell was calculated from the matrix of unimputed protein abundances per cell. These tallies are displayed in Fig. S4b.

To generate the histogram of representative precursor abundances and their corresponding fill times, the precursor intensities for all precursors identified in the Shotgun and pSCoPE single-cell BMDM analyses were split into tertiles. If a top tier precursor appeared in the bottom intensity tertile, it was allotted an MS2 fill time of 1000ms; if a top tier precursor appeared in the middle intensity tertile, it was allotted an MS2 fill time of 750ms; if a top-tier precursor appeared in the top intensity tertile, it was allotted an MS2 fill time of 500ms.

The matrix of precursor abundances by single-cell samples for 20 shotgun analyses and 20 pSCoPE analyses was subsetted to contain only those precursors that were allotted fill times of 750ms and 1000ms in the pSCoPE analyses. The percent data completeness was then calculated on a per-single-cell basis for the set of precursors allotted longer fill times present in the filtered matrix of precursors by single-cell samples. The results from this analyses are presented in Fig. S4c.

PCA from unimputed protein-level data, Figure S5

In order to assess whether the qualitative trends observed in the cross-condition PCA or the PSEA based on the PCA-derived protein weight vectors were compromised by imputation, PCA was performed on the correlation matrix generated by the batch-corrected, unimputed cells x proteins matrix, and the resulting PCA plot was color-coded by cell type or the median relative abundances of the proteins corresponding to Type I IFN signaling or phagosome maturation, Fig. S5.

PCA color-coded by precursor-level data completeness, Figure S6

In order to assess whether the sample separation observed in the cross-condition PCA (shown in Fig. 2) was driven by missing data, the single-cell data points were color-coded by the data completeness percentage. The post-data-filtration matrix of precursors by single-cells (filtration metrics described in the data processing and normalization section) was used to calculate the data completeness on a per-sample basis. The fraction of filtered precursors with observed reporter ion signal relative to the total number of filtered precursors detected across all experiments was then multiplied by 100 to generate the data completeness percentage.

Endocytosis Panel for Untreated BMDMs, Figure S8

This plot was constructed in the same manner as the main text figure corresponding to the 24-hr LPS-treated samples.

Instrument Methods

DIA Instrument Method 1

95-minute method with the following gradient characteristics: samples were loaded onto the column at 4% B; the gradient was then ramped to 8% at minute 12, 35% at minute 75, 95% at minute 77, 4% from minute 80.1 onward. MS1 scans had the following parameters: 140k resolution, 3e6 AGC target, 512 maximum injection time, and a scan range from 450 to 1258Th. Two sets of DIA windows were used: 21 20Th-wide windows (spanning the space from 450Th to 860Th) and 8 50Th-wide windows (spanning the space from 859.5Th to 1256Th). The DIA windows had the following characteristics: 35k resolution, AGC target of 5e5, maximum injection time determined automatically, fixed first mass of 200Th, NCE of 33, and a default charge state of 2. DIA windows spanned the space from 450Th to 1256Th and included a 0.5Th overlap.

DIA Instrument Method 2

All characteristics of DIA Instrument Method 1 were preserved, but instead of two sets of variable-width DIA windows, three were used: 21 15Th-wide windows (spanning the space from 450Th to 755Th), 8 20Th-wide windows (spanning the space from 754.5Th to 911Th), and 7 50Th windows (spanning the space from 910.5Th to 1257.5Th). DIA windows spanned the space from 450Th to 1257.5Th and included a 0.5Th overlap.

DIA Instrument Method 3, BMDM bulk samples

Samples were analyzed using a 105-minute method with the following gradient characteristics: samples were loaded onto the column at 4% B; the gradient was then ramped to 8% at minute 12, 35% at minute 75, 95% at minute 77, 4% from minute 80.1 onward. MS parameters were identical to DIA Instrument Method 1.

DIA Instrument Method 4, BMDM MEROPS samples

Samples were analyzed using a 160-minute method with the following characteristics: samples were loaded onto the column at 4% B; the gradient was then ramped to 7% B at 12 minutes, 32% B at 135 minutes, 95% B at 137 minutes, 4% B from minutes 140.1 to minute 160. MS1 scans were conducted using the following instrument settings: 70k resolution, 3e6 AGC target, 300ms maximum IT, with a scan range of 478 to 1470Th. The following DIA window scheme was used: 25 12.5Th-wide windows, 7 25Th-wide windows, and 8 62.5Th-wide windows. The following DIA window settings were used: 35k resolution, 3e6 AGC target, 110ms maximum fill time, and an NCE of 27.

DIA Instrument Method 6, BMDM Dextran Samples

Samples were analyzed using a 155-minute method with the following characteristics: samples were loaded onto the column at 4% B; the gradient was then ramped to 8% at minute 22; 29% at minute 130; 95% at minute 134; and 4% at minute 140.1. DIA cycles were made up of 1 survey scan from 400Th to 800Th, and 20 19.5Th windows spanning the space from 400Th to 799.5Th. Each MS1 scan had the following characteristics: 70k resolution, 3e6 AGC target, and 300ms maximum fill time. Each MS2 scan had the following characteristics: 17.5k resolution, 5e5 AGC target, 64ms maximum fill time, fixed first mass of 200Th, and an NCE of 27. Samples were resuspended in 0.1% formic acid post-stage-tipping and loaded in glass HPLC inserts (Thermo Fisher C4010-630) prior to analysis.

Shotgun Bulk Instrument Method 1

Samples were analyzed using a 95-minute method with the following characteristics: samples were loaded onto the column at 4% B; the gradient was then ramped to 8% B at minute 12, 35% B at minute 75, 95% at minute 77, and 4% from minute 80.1 to 95. The following MS1 settings were used: 70k resolution, 1e6 AGC target, 100ms maximum injection time, and a scan range of 450Th to 1600Th. MS2 scans were acquired with the following settings: 70k resolution, 5e4 AGC target, 300ms maximum injection time, loop count (i.e. top-n) of 7, Isolation window of 0.7Th with a 0.3Th offset, fixed first mass of 100 m/z, NCE of 33, and a centroid spectrum data type. The minimum AGC target was 2e4, apex triggering was disabled, and charge exclusion was enabled for unassigned charge states, as well as charge states greater than 6. The peptide match setting was disabled, exclude isotopes was enabled, and dynamic exclusion was set to 30 seconds. Voltage was set to 0 for the first 25 minutes, sweep gas was applied from minute 24.6 to 25 to dislodge any accumulated droplets from the capillary tip. From minute 25 to 80, voltage was set to 1.7kV, capillary temp to 250 °C, and the S-lens RF level to 80. From minute 94.20 to 94.60, sweep gas was applied to dislodge any accumulated droplets from the capillary tip.

Shotgun Bulk Instrument Method 2

Samples were analyzed using a 95-minute method with the following characteristics: samples were loaded onto the column at 4% B; the gradient was then ramped to 8% B at minute 12, 35% B at minute 75, 95% at minute 77, and 4% from minute 80.1 to 95. The following MS1 settings were used: 70k resolution, 1e6 AGC target, 100ms maximum injection time, and a scan range of 450Th to 1600Th. MS2 scans were acquired with the following settings: 35k resolution, 5e4 AGC target, 150ms maximum injection time, loop count (i.e. top-n) of 14, Isolation window of 0.7Th with a 0.3Th offset, fixed first mass of 100 m/z, NCE of 33, and a centroid spectrum data type. The minimum AGC target was 2e4, apex triggering was disabled, and charge exclusion was enabled for unassigned charge states, as well as charge states greater than 6. The peptide match setting was disabled, exclude isotopes was enabled, and dynamic exclusion was set to 30 seconds. Voltage was set to 0 for the first 25 minutes, sweep gas was applied from minute 24.6 to 25 to dislodge any accumulated droplets from the capillary tip. From minute 25 to 80, voltage was set to 1.7kV, capillary temp to 250 °C, and the S-lens RF level to 80. From minute 94.20 to 94.60, sweep gas was applied to dislodge any accumulated droplets from the capillary tip.

Shotgun Instrument Method 3, Figure 1b/e single-cell sets

Single-cell samples were analyzed using a 95-minute method with the following characteristics: samples were loaded onto the column at 4% B; the gradient was then ramped to 8% B at minute 12, 35% B at minute 75, 95% at minute 77, and 4% from minute 80.1 to 95. The following MS1 settings were used: 70k resolution, 1e6 AGC target, 100ms maximum injection time, and a scan range of 450Th to 1258 Th. MS2 scans were acquired with the following settings: 70k resolution, 1e6 AGC target, 300ms maximum injection time, loop count (i.e. top-n) of 7, Isolation window of 0.5 Th without an offset, fixed first mass of 100 m/z, NCE of 33, and the spectra were saved in centroid form. The minimum AGC target was 2e4, apex triggering was disabled, and charge exclusion was enabled for unassigned charge states, as well as charge states greater than 6. The

peptide match setting was disabled, exclude isotopes was enabled, and dynamic exclusion was set to 30 seconds. Voltage was set to 0 for the first 25 minutes, sweep gas was applied from minute 24.6 to 25 to dislodge any accumulated droplets from the capillary tip. From minute 25 to 88, voltage was set to 1.7kV, capillary temp to 250 °C, and the S-lens RF level to 80. From minute 88 to 95, voltage was set to 0kV. From minute 94.20 to 94.60, sweep gas was applied to dislodge any accumulated droplets from the capillary tip.

Shotgun Instrument Method 4, Figure 1c/d single-cell sets

This method differs from Shotgun Method 3 in its usage of a 0.7Th isolation window with a 0.5Th offset.

Shotgun Instrument Method 5, BMDM nPOP single-cell sets

This method differs from Shotgun Method 3 in that its MS2 scans were conducted at a resolution of 140k, an AGC target of 1e6, a maximum injection time of 500ms, with a loop count of 4

Shotgun Instrument Method 6: mPOP single-cell samples

Samples were analyzed using a 105-minute total run time with the following gradient stages: samples were loaded onto the column at 4% B; the gradient was then ramped to 8% B at 12 minutes, 35% B at 75 minutes, 95% B at 77 minutes, 4% B from minutes 80.1 to minute 105. The MS instrument settings matched Shotgun Instrument Method 5, with the exception of a maximum MS2 injection time of 600ms.

MaxQuant.Live parameters

Parameters for prioritized single cell sets, Figures 1/2/3/4

LC settings were identical to the matched shotgun analysis described above. Scan parameters implemented following the MQ.live listening scan guidelines: Two full MS - SIM scans were applied from minute 25 to 30 to trigger MaxQuant.live. Both MS - SIM scans had the following parameters in common: 70k resolution, 3e6 AGC target, and a 300ms maximum injection time. The first MS - SIM scan covered 908Th to 1070Th, the upper bound indicating a total acquisition time of 70 minutes to MaxQuant.live, discounting the initial 25-minute period of each run in which the voltage was off. The second MS - SIM scan covered the scan space from 909 to the m/z corresponding to the MaxQuant.live method index to call. The total Xcalibur MS method time was 95 minutes.

Global settings: Survey scan	**Fig. 1b**	**Fig. 1c,d**
ScanDataAsProfile	True	True
PositiveMode	True	True
MaxIT	100ms	100ms
Resolution	70,000	70,000
AgcTarget	1,000,000	1,000,000
MzRange	(450,1258)	(450,1258)
BoxCarScans	0	0
Global settings: TopN		
NumOfMS2Scans	0	0

RealtimeCorrection	**Fig. 1b**	**Fig. 1c,d**
MzTolerances	(4.5,5)	(4.5,5)
RetentionTimeTolerances	(0.01,2)	(0.01,2)
SigmaScaleFactorRt	3	3
PeptideHistoryLength	2	2
MinUsedCorrectionPeptides	15	15
IntensityPeakRatioThreshold	0.01	0.01
PeptideDetectionIsoPeaks	2	2
IsotopeTolerance	9	9
Ms2DetectionNeeded	False	False
Ms2ExcludeDetectedPeptides	False	False
Ms2MinNormIntensity	0.1	0.1
Ms2MzTolerance	20	20

TargetedMs2	**Fig. 1b**	**Fig. 1c,d**
BatMode	False	False
AutoPriority	True	True
DefaultPriority	0	0
MaxNumOfScans	1	1
WindowAndOffsetInDalton	False	False
ScanDataAsProfile	False	False
WindowSize	0.5	0.5
MzOffset	0	0
LowerMzBound	100	100
CollisionEnergy	33	33
LifeTime	2,100ms	2,100ms
Resolution	70,000	70,000
MaxIT	300ms	300ms
AgcTarget	1,000,000	1,000,000
PositiveMode	True	True

Prioritized acquisition parameters, nPOP BMDM single-cell samples

Global settings: Survey scan	**Scout sample**	**Single-cell sets**
ScanDataAsProfile	True	True
PositiveMode	True	True
MaxIT	100ms	100ms
Resolution	70,000	70,000
AgcTarget	1,000,000	1,000,000
MzRange	(450,1258)	(450,1258)
BoxCarScans	0	0
Global settings: TopN	**Scout sample**	**Single-cell sets**
NumOfMS2Scans	0	0

RealtimeCorrection	**Scout sample**	**Single-cell sets**
MzTolerances	(4.5,5)	(4.5,5)
RetentionTimeTolerances	(0.01,2)	(0.01,2)
SigmaScaleFactorRt	3	3
PeptideHistoryLength	2	2
MinUsedCorrectionPeptides	15	15
IntensityPeakRatioThreshold	1e-5	1e-5
PeptideDetectionIsoPeaks	2	2
IsotopeTolerance	9	9
Ms2DetectionNeeded	False	False
Ms2ExcludeDetectedPeptides	False	False
Ms2MinNormIntensity	0.1	01
Ms2MzTolerance	20	20

TargetedMs2	Scout sample	Single-cell sets
BatMode	False	False
AutoPriority	True	True
DefaultPriority	0	0
MaxNumOfScans	1	1
WindowAndOffsetInDalton	False	False
ScanDataAsProfile	False	False
WindowSize	0.5	0.5
MzOffset	0	0
LowerMzBound	100	100
CollisionEnergy	33	33
LifeTime	2,400ms	2,100ms
Resolution	140,000	140,000
MaxIT	500ms	500ms
AgcTarget	1,000,000	1,000,000
PositiveMode	True	True

Prioritized acquisition parameters, mPOP BMDM single-cell samples

All single-cell samples were resuspended in $1.2\mu l$ of 0.1% formic acid (Thermo Fisher 85178) and injected from glass HPLC inserts (Thermo Fisher C4010-630). LC settings were preserved from the complementary shotgun set. Scan parameters implemented following the MQ.live listening scan guidelines: Two Full MS - SIM scans were applied from minute 26 to 31 to trigger MaxQuant.live. Both MS - SIM scans had the following parameters in common: 70k resolution, 1e6 AGC target, and a 300ms maximum injection time. The first MS - SIM scan covered 908 to 1079Th, since the acquisition started at minute 26 and ended at minute 105. The second MS - SIM

scan covered the scan space from 909 to 1030Th, the upper bound indicating the MaxQuant.live method index to call. The total Xcalibur MS method time was 105 minutes. MaxQuant.live parameters contained in the associated table, below.

Global settings: Survey scan	**mPOP Single-cell sets**
ScanDataAsProfile	True
PositiveMode	True
MaxIT	100ms
Resolution	70,000
AgcTarget	1,000,000
MzRange	(450,1258)
BoxCarScans	0
Global settings: TopN	**mPOP Single-cell sets**
NumOfMS2Scans	0

RealtimeCorrection	**mPOP Single-cell sets**
MzTolerances	(4.5,5)
RetentionTimeTolerances	(0.01,1.5)
SigmaScaleFactorRt	3
PeptideHistoryLength	2
MinUsedCorrectionPeptides	15
IntensityPeakRatioThreshold	1e-5
PeptideDetectionIsoPeaks	2
IsotopeTolerance	9
Ms2DetectionNeeded	False
Ms2ExcludeDetectedPeptides	False
Ms2MinNormIntensity	0.1
Ms2MzTolerance	20

TargetedMs2	mPOP Single-cell sets
BatMode	False
AutoPriority	True
DefaultPriority	0
MaxNumOfScans	1
WindowAndOffsetInDalton	False
ScanDataAsProfile	NA
WindowSize	0.5
MzOffset	0
LowerMzBound	100
CollisionEnergy	33
LifeTime	2,400ms
Resolution	140,000
MaxIT	600ms
AgcTarget	1,000,000
PositiveMode	True

References

1. Murray, P. J. Macrophage polarization. *Annual review of physiology* **79,** 541–566 (2017).
2. Jha, A. K. *et al.* Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization. *Immunity* **42,** 419–430 (2015).
3. Specht, H. *et al.* Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biology* **22.** doi:10.1186/s13059-021-02267-5 (2021).
4. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass-spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biology* **19,** 161 (2018).
5. Petelski, A. A. *et al.* Multiplexed single-cell proteomics using SCoPE2. *Nature Protocols* **16,** 5398–5425 (2021).
6. Schoof, E. M. *et al.* Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nature Communications* **12** (2021).
7. Furtwängler, B. *et al.* Real-Time Search Assisted Acquisition on a Tribrid Mass Spectrometer Improves Coverage in Multiplexed Single-Cell Proteomics. en. *Mol. Cell. Proteomics.* doi:10.1016/j.mcpro.2022.100219 (2022).
8. Cong, Y. *et al.* Ultrasensitive single-cell proteomics workflow identifies > 1000 protein groups per mammalian cell. *Chemical Science* (2020).
9. Lombard-Banek, C. *et al.* In vivo subcellular mass spectrometry enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (X. laevis). *Angewandte Chemie International Edition* **60,** 12852–12858 (2021).
10. Singh, A. Towards resolving proteomes in single cells. en. *Nat. Methods* **18,** 856 (Aug. 2021).
11. Clark, N. M., Elmore, J. M. & Walley, J. W. To the proteome and beyond: advances in single-cell omics profiling for plant systems. en. *Plant Physiol.* (Sept. 2021).
12. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M.-C. & Yates III, J. R. Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews* **113,** 2343–2394 (2013).
13. Slavov, N. Driving Single Cell Proteomics Forward with Innovation. *Journal of Proteome Research* **20,** 4915–4918 (2021).
14. Picotti, P. & Aebersold, R. Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. *Nature methods* **9,** 555 (2012).
15. Picotti, P. *et al.* A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* **494,** 266–270 (2013).
16. Marx, V. Targeted proteomics. *Nature methods* **10,** 19 (2013).
17. Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S. & Coon, J. J. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Molecular & Cellular Proteomics* **11,** 1475–1488 (2012).

18. Bensimon, A., Heck, A. J. & Aebersold, R. Mass spectrometry–based proteomics and network biology. *Annual review of biochemistry* **81,** 379–405 (2012).

19. Erickson, B. K. *et al.* A strategy to combine sample multiplexing with targeted proteomics assays for high-throughput protein signature characterization. *Molecular cell* **65,** 361–370 (2017).

20. Manes, N. P. & Nita-Lazar, A. Application of targeted mass spectrometry in bottom-up proteomics for systems biology research. *Journal of proteomics* **189,** 75–90 (2018).

21. Schweppe, D. K. *et al.* Full-Featured, Real-Time Database Searching Platform Enables Fast and Accurate Multiplexed Quantitative Proteomics. en. *J. Proteome Res.* **19,** 2026–2034 (May 2020).

22. Furtwängler, B. *et al.* Real-Time Search Assisted Acquisition on a Tribrid Mass Spectrometer Improves Coverage in Multiplexed Single-Cell Proteomics. en. *bioRxiv,* 2021.08.16.456445 (Aug. 2021).

23. Wichmann, C. *et al.* MaxQuant.Live enables global targeting of more than 25,000 peptides. *Molecular & Cellular Proteomics* **18,** 982–994 (2019).

24. Leduc, A., Huffman, R. G. & Slavov, N. Droplet sample preparation for single-cell proteomics applied to the cell cycle. *bioRxiv 2021.04.24.441211.* doi:10.1101/2021.04.24.441211 (2021).

25. Shalek, A. K. *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. en. *Nature* **510.** ISSN: 1476-4687. doi:10.1038/nature13437. (2022) (June 2014).

26. Rawlings, N. D. *et al.* The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research* **46,** D624–D632. ISSN: 0305-1048 (Jan. 2018).

27. Lee, S. *et al.* Negative Self-Regulation of TLR9 Signaling by Its N-Terminal Proteolytic Cleavage Product. *Journal of immunology* **193,** 3726–3735. ISSN: 0022-1767 (Oct. 2014).

28. He, W.-t. *et al.* Gasdermin D is an executor of pyroptosis and required for interleukin-1β secretion. en. *Cell Research* **25,** 1285–1298. ISSN: 1748-7838 (Dec. 2015).

29. Vérollet, C. *et al.* Extracellular proteolysis in macrophage migration: losing grip for a breakthrough. *European journal of immunology* **41,** 2805–2813 (2011).

30. Wang, L., Main, K., Wang, H., Julien, O. & Dufour, A. Biochemical Tools for Tracking Proteolysis. en. *J. Proteome Res.* **20,** 5264–5279 (Dec. 2021).

31. Meyer, B. *et al.* Characterising proteolysis during SARS-CoV-2 infection identifies viral cleavage sites and cellular targets with therapeutic potential. en. *Nature Communications* **12.** ISSN: 2041-1723. doi:10.1038/s41467-021-25796-w. (2022) (Sept. 2021).

32. Martinez, F. O., Gordon, S., Locati, M. & Mantovani, A. Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression. en. *The Journal of Immunology* **177.** doi:10.4049/jimmunol.177.10.7303. (2022) (Nov. 2006).

33. Remes, P. M., Yip, P. & MacCoss, M. J. Highly Multiplex Targeted Proteomics Enabled by Real-Time Chromatographic Alignment. en. *Anal. Chem.* **92,** 11809–11817 (Sept. 2020).

34. Zhu, H. *et al.* PRM-LIVE with Trapped Ion Mobility Spectrometry and Its Application in Selectivity Profiling of Kinase Inhibitors. en. *Anal. Chem.* **93,** 13791–13799 (Oct. 2021).

35. Slavov, N. Scaling Up Single-Cell Proteomics. *Molecular & Cellular Proteomics* **21,** 100179. ISSN: 1535-9476 (2022).

36. Slavov, N. Increasing proteomics throughput. *Nature Biotechnology* **39,** 809–810 (2021).

37. Slavov, N. Learning from natural variation across the proteomes of single cells. *PLOS Biology* **20,** 1–4 (Jan. 2022).

38. Specht, H. & Slavov, N. Optimizing Accuracy and Depth of Protein Quantification in Experiments Using Isobaric Carriers. *Journal of Proteome Research* **20.** PMID: 33190502, 880–887 (2021).

39. Specht, H. *et al.* Automated sample preparation for high-throughput single-cell proteomics. *bioRxiv,* 399774 (Jan. 2018).

40. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. en. *Nature Protocols* **2,** 1896–1906. ISSN: 1754-2189, 1750-2799 (Aug. 2007).

41. Chen, A. T., Franks, A. & Slavov, N. DART-ID increases single-cell proteome coverage. *PLOS Computational Biology* **15,** 1–30 (July 2019).

42. Specht, H. *et al.* Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biology* **22,** 50. ISSN: 1474-760X (Jan. 2021).

43. Huntley, R. *et al.* The GOA database: Gene Ontology annotation updates for 2015. *Nucleic Acids Research* **43,** D1057–63 (2015).

Supplemental Figures

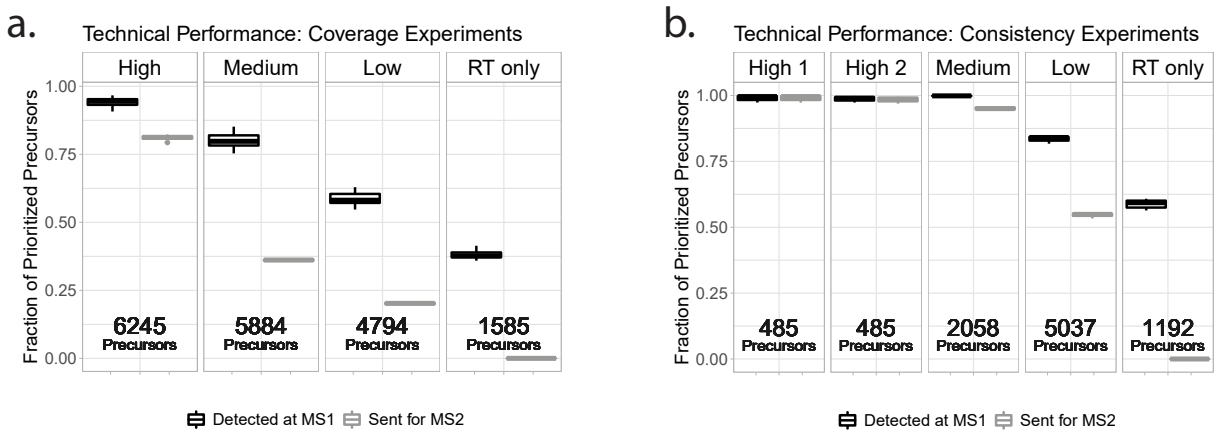


Figure S1 | Fraction of inclusion-list precursors detected and analyzed in pSCoPE runs
(**a**) MS1 detection and MS2 analysis rates for prioritized precursors in the benchmark experiments displayed in Fig. 1b. (**b**) MS1 detection and MS2 analysis rates for prioritized precursors in the benchmark experiments displayed in Fig. 1c,d. In both panels, the statistics are shown for each tier along with the number of precursors in the tier.

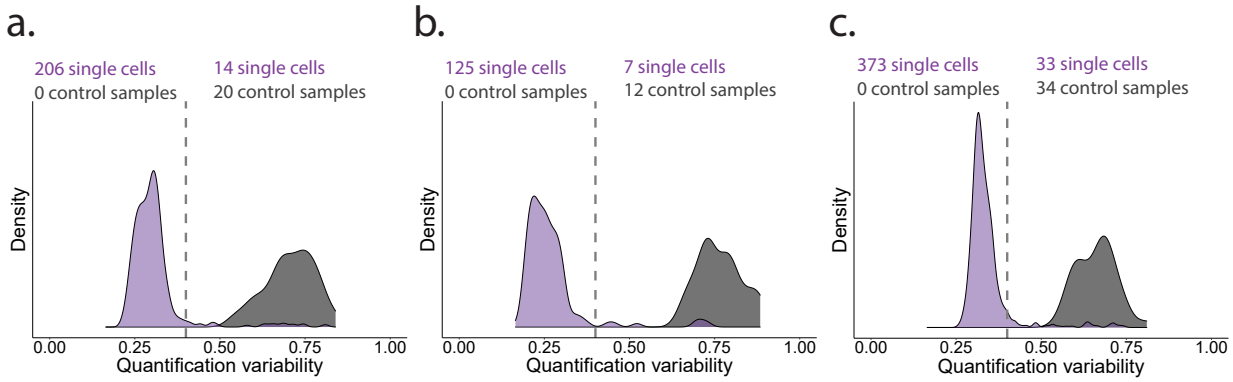


Figure S2 | Single-cell quality controls
The mean coefficient of variation (i.e. the standard deviation scaled by the mean) of all peptide-level relative abundances that map to the same leading razor protein is used to separate successfully prepared single cells from those that will not generate accurate data. By choosing a CV threshold that separates control samples (droplets which received all reagents but did not contain a single cell) from single cells, cells with noisier protein-level quantitation can be removed prior to further data processing. The cell single cell and control tallies above each figure represent the number of single cells or control wells that passed the CV threshold of 0.4. (**a**) contains the CV distributions for the single-cell samples associated with Fig. 1b. (**b**) contains the CV distributions for the single-cell samples associated with Fig. 1c/d. (**c**) contains the CV distributions for the single-cell samples associated with Fig. 2, Fig. 3, and Fig. 4.

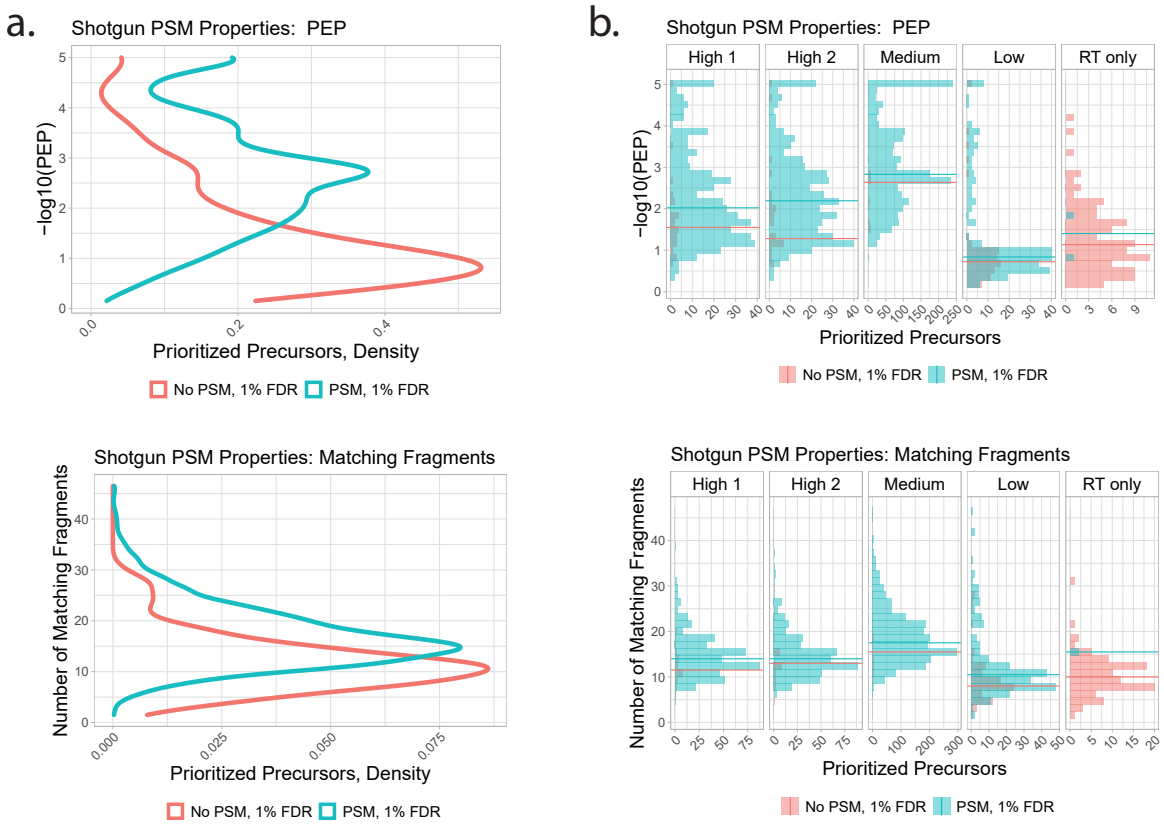


Figure S3 | Properties of peptides successfully identified in pSCoPE runs
The precursors from the inclusion list were split into those that resulted in confident PSMs and those that did not, and the properties of each set analyzed based on the shotgun runs used for making the inclusion lists. (**a**) Confidence of identification (quantified by the posterior error probability; PEP) and number of matching peptide fragments for successful and unsuccessful precursors. The data are shown for all prioritized peptides across all priority tiers. (**b**) The data from panel a are shown faceted by priority tier. All data shown are from the consistency experiments from Fig. 1c/d.

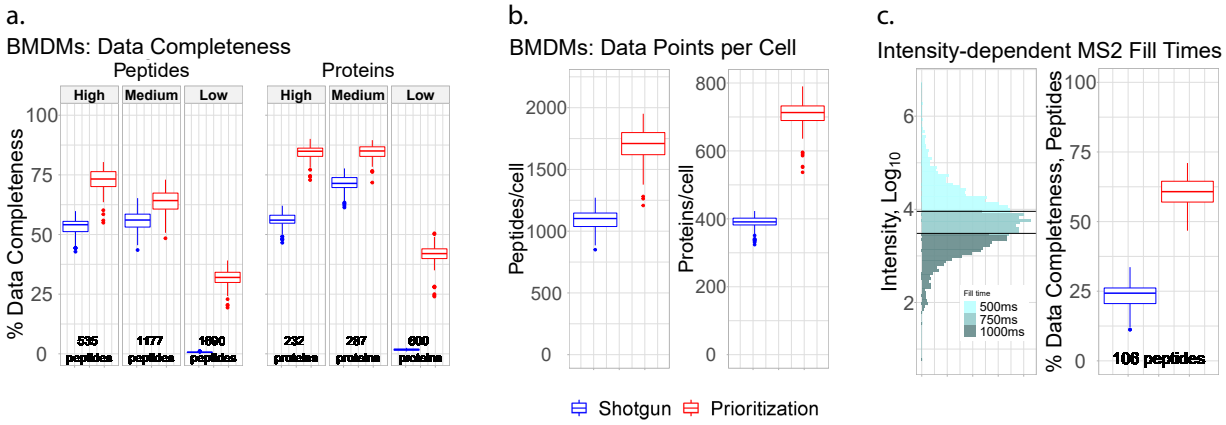


Figure S4 | Data completeness and proteome coverage for single BMDM analysed by shotgun or prioritized methods

(**a**) Percent data completeness tallied for peptides and proteins quantified across twenty shotgun and twenty pSCoPE experiments, faceted by priority tier. (**b**) Number of peptides and proteins per single-cell sample across twenty shotgun and twenty pSCoPE experiments. (**c**) Illustration of precursor-intensity-dependent MS2 fill times for precursors on the top priority tier. Percent data completeness contrast for precursors which were allotted increased fill times in the pSCoPE analyses.

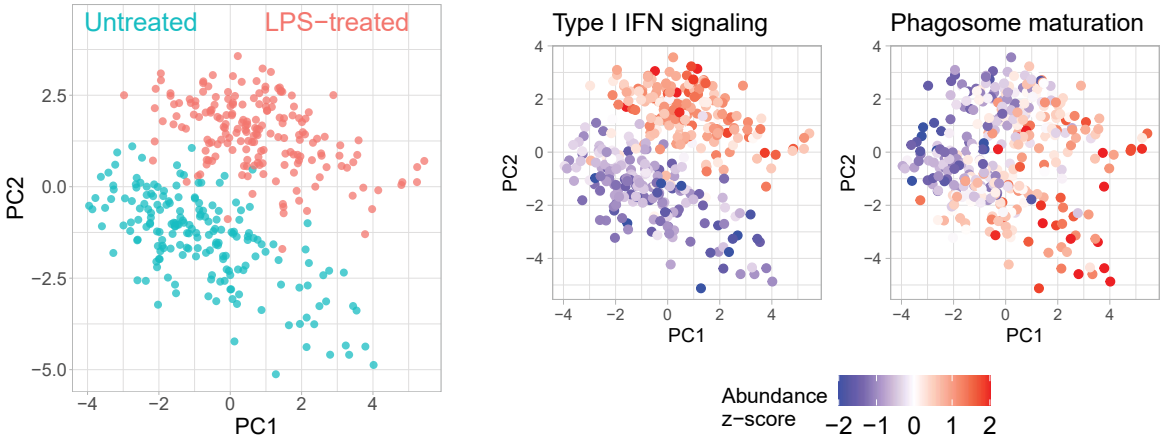


Figure S5 | PCA of single BMDM using only observed datapoints
To evaluate the robustness of our results to uncertainties stemming from missing data, we performed PCA of unimputed BMDM data. The single cells are color-coded by treatment condition, with adjoining PCA plots color-coded by the median relative abundance of proteins corresponding to type I interferon-mediated signaling and phagosome maturation.

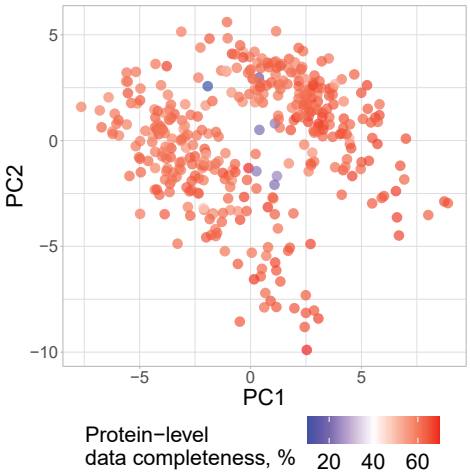


Figure S6 | PCA color-coded by protein-level data completeness
To evaluate the whether the biological conclusions we drew from our PC-weight-based PSEA could have been influenced by separation due to data completeness, we color-coded our cross-condition BMDM PCA by the percent data completeness on a per-cell basis.

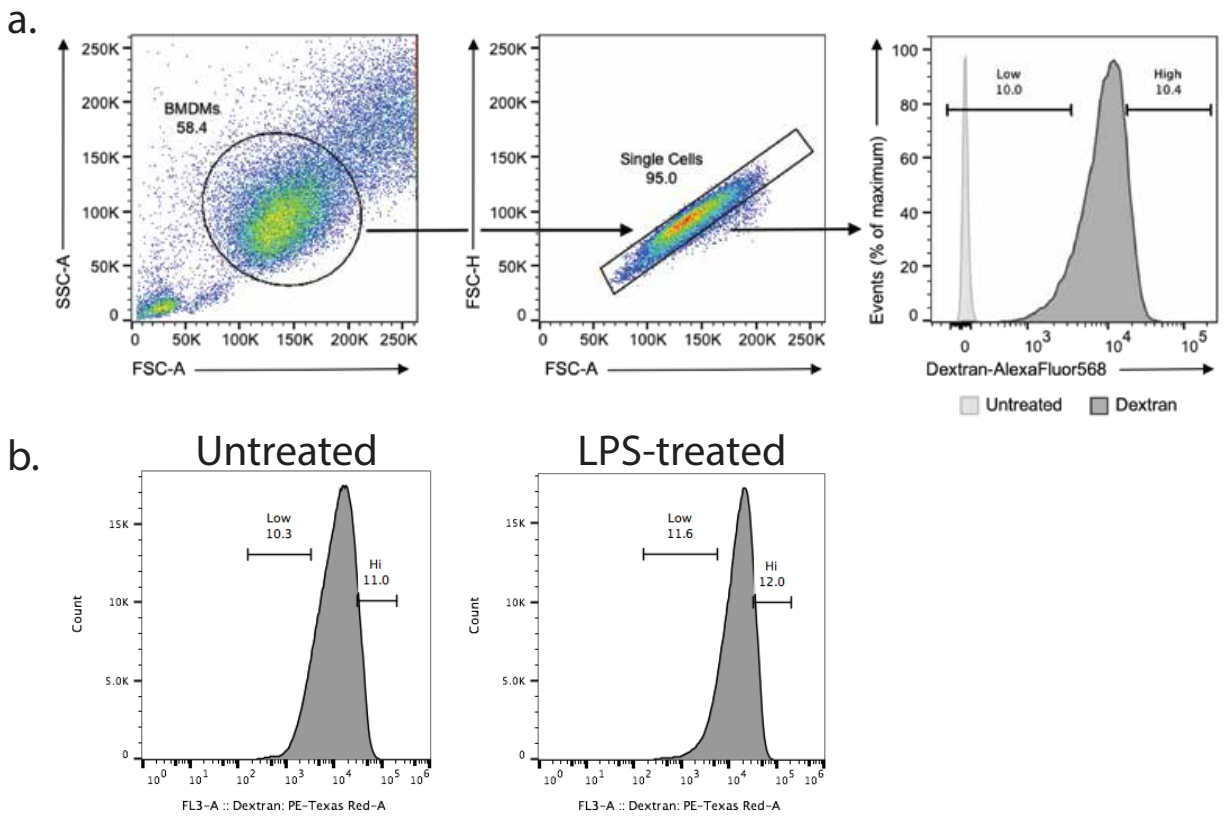


Figure S7 | FACS gating parameters and staining controls
(a.) FSC-A and SSC-A gates for sorted bone-marrow-derived macrophages and positive/negative staining populations. (b.) Dextran:PE-Texas Red gating parameters for isolating the most and least endocytic BMDM populations from each treatment group (untreated and LPS-treated).

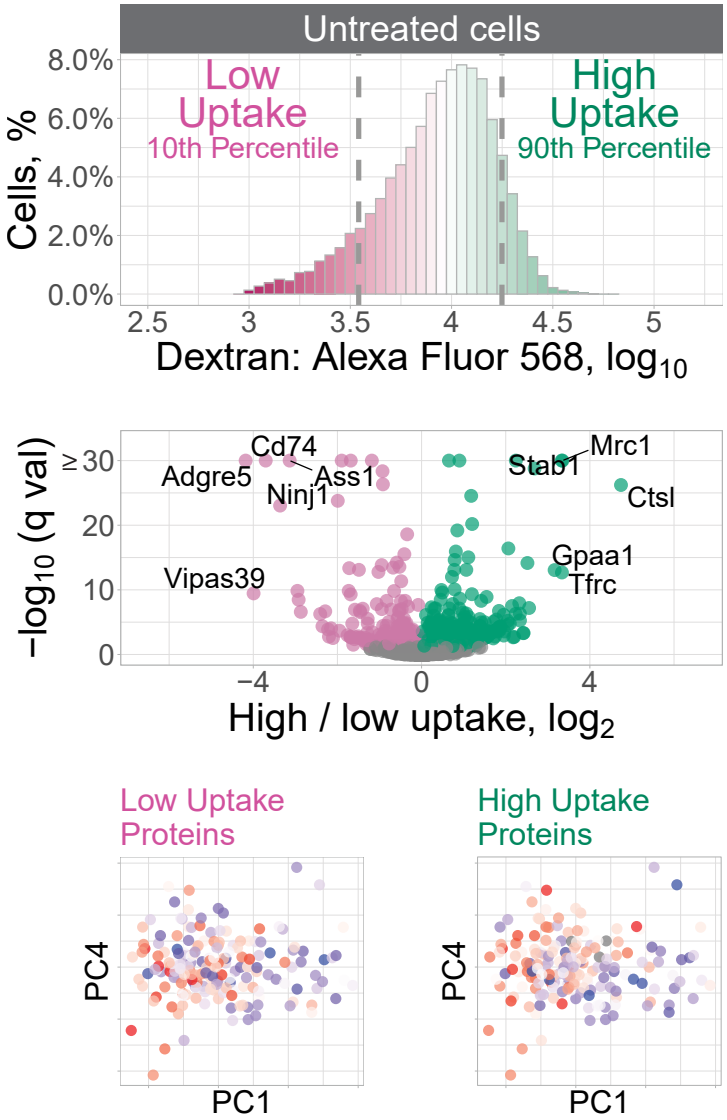


Figure S8 | Dextran uptake in untreated BMDM samples
The uptake of fluorescent dextran by the untreated macrophages was measured by FACS, and the cells with the lowest and highest uptake were isolated for protein analysis. The volcano plot displays the fold changes for deferentially abundant proteins and the associated statistical significance. The untreated macrophages were displayed in the space of their PCs and color-coded by the median abundance of the low-uptake or the high-uptake proteins. Both the low and the high-uptake proteins correlate inversely to PC1 (low-uptake: Spearman r = -0.29, $q \leq 6 \times 10^{-4}$; high-uptake: Spearman r = -0.37, $q \leq 4 \times 10^{-6}$).

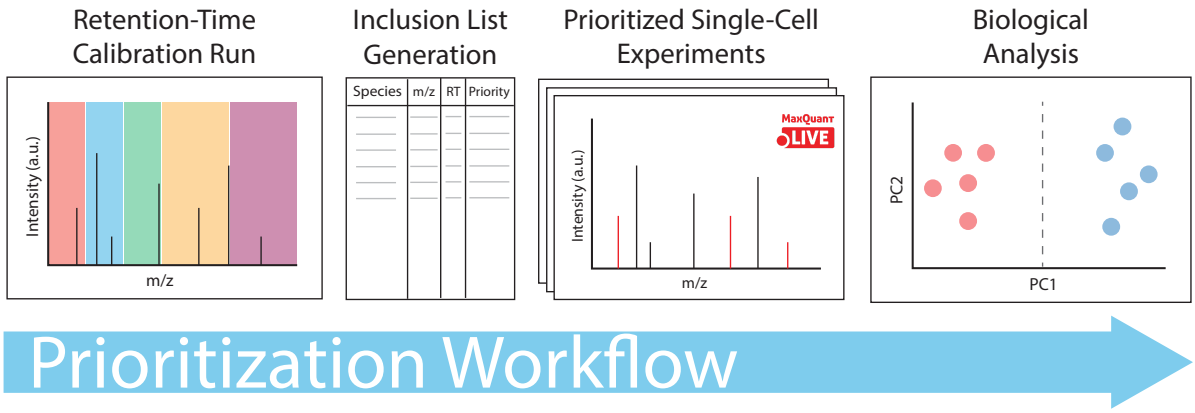


Figure S9 | pSCoPE workflow schematic
The pSCoPE workflow begins with the assembly of a number of precursors of experimental interest which can originate from prior DDA or DIA analyses of bulk samples, DDA analyses of previous SCoPE samples, or literature. A DIA method is then applied to a 1x injection of carrier and reference material to generate accurate retention times for the peptides of interest. The user then stratifies the peptides identified during the retention-time calibration run into priority tiers, with the top priority tier containing peptides of highest experimental interest. The prioritized inclusion list is imported into MaxQuant.Live for prioritized analysis of SCoPE samples.